

Inaugural-Dissertation zur Erlangung der Doktorwürde
der Tierärztlichen Fakultät der Ludwig-Maximilians-Universität
München

Pose Estimation in der Lahmheitsdiagnostik und
Bewegungsanalyse des Pferdes

von Ann-Kristin Feuser

aus Starnberg

München 2023

Aus dem Zentrum für Klinische Tiermedizin der Tierärztlichen Fakultät
der Ludwig-Maximilians-Universität München

Lehrstuhl für Pferdemedizin

Arbeit angefertigt unter der Leitung von: Univ.-Prof. Dr. Anna May

Gedruckt mit Genehmigung der Tierärztlichen Fakultät
der Ludwig-Maximilians-Universität München

Dekan: Univ.-Prof. Dr. Reinhard K. Straubinger, Ph.D.

Berichterstatter: Univ.-Prof. Dr. Anna May

Korreferent: Univ.-Prof. Dr. Susanne K. Lauer

Tag der Promotion: 11. Februar 2023

Für meine Eltern

Eva Feuser und Dr. Albert Feuser

INHALTSVERZEICHNIS

I.	EINLEITUNG	8
II.	LITERATURÜBERSICHT.....	11
1.	Lahmheit bei Pferden.....	11
1.1.	Definition von Lahmheit	11
1.2.	Einteilung von Lahmheiten	11
1.2.1.	Einteilung nach Art der Lahmheit	11
1.2.2.	Einteilung nach Schweregrad.....	12
1.2.3.	Einteilung in Vorder- und Hinterhandlahmheiten.....	13
2.	Klinische Lahmheitsuntersuchung	14
3.	Deep Learning und Pose Estimation	17
3.1.	Künstliche Intelligenz in der Human- und Tiermedizin.....	17
3.2.	Grundlagen zu Pose Estimation und Deep Learning.....	19
4.	Systeme für die Analyse von Lahmheiten	23
4.1.	Entwicklung.....	23
4.2.	Systeme auf kinetischer Grundlage.....	23
4.3.	Systeme auf kinematischer Grundlage	25
4.3.1.	Optical Motion Capture – Bewegungserfassung mittels Kameras.....	25
4.3.2.	Inertial Measurement-Systeme.....	27
4.3.3.	Radarsensoren	29
III.	EIGENE UNTERSUCHUNGEN.....	31
1.	Zielsetzung	32
2.	Publikation	32
	Introduction	34
	Materials and Methods	36
	Results	41
	Discussion	45
	Conclusions	48
	Appendix	49
	References	51
IV.	ERWEITERTE DISKUSSION	54

V.	ZUSAMMENFASSUNG.....	65
VI.	SUMMARY.....	67
VII.	LITERATURVERZEICHNIS	69
VIII.	ANHANG	82
IX.	DANKSAGUNG	83

ABKÜRZUNGSVERZEICHNIS

Aa.	Arteriae
AAEP	American Association of Equine Practitioners
al.	alia
ANN	Artificial Neural Network
CT	Computertomographie
engl.	englisch
ERU	Equine rezidivierende Uveitis
HZ	Hertz
IMU	Inertial Measurement Unit
KI	Künstliche Intelligenz
M	Musculus
MRT	Magnetresonanztomographie
n	Anzahl
OMC	Optical Motion Capture
Proc.	Processus
RADAR	Radio Detection and Ranging
ResNet	Residual Network
RMSE	root mean square error
T.	Tuber

I. EINLEITUNG

Die Entwicklung von Systemen zu Lahmheitsanalyse bei Pferden, die sich für den alltäglichen Gebrauch eignen, sowie deren Anwendung haben in den letzten Jahren stark zugenommen (van Weeren et al. 2017). In einer aktuellen Umfrage von Tierärztinnen und Tierärzten sowohl in europäischen als auch außereuropäischen Ländern haben über 50% der Befragten angegeben, derartige Systeme in ihrer Klinik oder Praxis für Pferde zu verwenden (Hardeman et al. 2022). Als Vorteile wurden die Möglichkeit zur Dokumentation, der Einsatz als Präventivmaßnahme und vor allem die Tatsache, dass die Anwendung zu einem Hinterfragen der eigenen Beurteilung von Lahmheiten und zu einer anderen Sichtweise geführt haben, genannt (Hardeman et al. 2022). Auch van Weeren et al. (2018) zeigt in einem Leitartikel die Tatsache auf, dass ein Erkennen der eigenen Grenzen und auch das Eingestehen von möglicherweise gemachten Fehlern helfen kann, seine Kompetenzen zu erweitern. Fuller et al. (2006) und Keegan (2019) konnten zeigen, dass vor allem bei geringgradigen Lahmheiten auch unter erfahrenen Tierärzten und Tierärztinnen häufig Uneinigkeit über das betroffene Bein besteht. Daher ist es nicht überraschend, dass ein Zusammenhang zwischen dem Ausbildungsstand junger Tierärzte und Tierärztinnen und deren fachlicher Kompetenz im Erkennen von Lahmheiten festgestellt werden konnte (Starke und May 2017). Starke und May (2017) haben gezeigt, dass Veterinärmedizinstudierende sowie frischgebackene Absolventinnen und Absolventen oft Schwierigkeiten haben, die lahme Gliedmaße zu identifizieren.

Lahmheiten repräsentieren in der Pferdemedizin das häufigste medizinische Problem und können jedes Pferd auf jedem Trainingsniveau betreffen (USDA 1998, Slater 2016). Eine unerkannte Lahmheit stellt ein erhebliches Problem für das Wohlbefinden des betroffenen Pferdes da, so dass BesitzerInnen und Tierärzte und Tierärztinnen in der Lage sein sollten, Veränderungen des Gangbildes so früh wie möglich zu erkennen. Sowohl Studien von Dyson und Pollard (2020) als auch Muller-Quirin et al. (2020) haben jedoch gezeigt, dass Lahmheiten beim eigenen Pferd von den BesitzerInnen oft nicht wahrgenommen werden und dass es insbesondere für unerfahrene ReiterInnen sehr schwierig ist, muskuloskelettale Schmerzen bei Pferden zu erkennen. Auch wenn das Wohl und die Unversehrtheit des Pferdes an erster Stelle stehen, sind auch die finanziellen Aspekte nicht zu

unterschätzen. In der Pferdebranche entstehen durch Lahmheiten erhebliche wirtschaftliche Verluste, sei es durch unterbrochene oder vorzeitig beendete Sportkarrieren, Tierarzt-, Medikamenten- oder Zusatzkosten, oder sogar den Tod (Seitzinger 2000).

Ein Aspekt, der in der Diskussion um die Anwendung von Analysesystemen immer wieder aufgeführt wird, ist die unterschiedliche Definition von „Lahmheit“ bei Pferden (van Weeren et al. 2017, van Weeren et al. 2018). Nach Baxter et al. (2020) ist eine Lahmheit ein Hinweis auf eine strukturelle oder funktionelle Störung in den Gliedmaßen oder dem Rücken, die sich sowohl im Stand als auch in der Bewegung äußern kann. Ross (2011) definiert Lahmheit als klinische Erscheinung, die mit Anzeichen von Entzündung einhergeht oder einen mechanischen Defekt darstellt. Diese führt zu einer Veränderung des Gangbildes, was in einem Hinken oder Humpeln resultiert. In der Definition von Back und Clayton (2013) wird Lahmheit wiederum als Veränderung des physiologischen Gangbildes aufgrund einer funktionellen oder strukturellen Störung im Bewegungsapparat erläutert. Diese Definitionen zeigen, dass eine klare Differenzierung zwischen einer messbaren Asymmetrie im Vergleich der rechten und linken Körperseite in Bewegung und dem Begriff „Lahmheit“ nicht möglich ist. Das Vorkommen perfekter Symmetrie gibt es in der Natur nicht (van Weeren et al. 2018), weshalb ein Anstreben dieser weder möglich noch sinnvoll ist. Dennoch bieten Analysesysteme die Möglichkeit, die Lahmheit eines Pferdes nicht nur in fünf Lahmheitsgraden einzuteilen, wie in der Lahmheitsskala der American Association of Equine Practitioners (AAEP) in den USA (American Association of Equine Practitioners 2022), sondern die Bewegung eines Pferdes genauer zu differenzieren. Ein weiterer Nachteil, der auch von Hardeman et al. (2022) aufgeführt wurden, ist die Erfassung von Analyseparametern primär in der sagittalen Ebene. Dieser Aspekt wird auch von van Weeren et al. (2018) aufgegriffen und macht deutlich, dass eine Weiterentwicklung der bestehenden Systeme notwendig ist, um mehr Parameter erfassen zu können. Weitere Faktoren, die eine Anwendung erschweren, sind der zeitliche Aufwand, den der Aufbau und die Anwendung einiger Systeme mit sich bringt, die ausschließliche Nutzbarkeit an einem Standort, sowie Anschaffungs- und Installationskosten und die komplexe Auswertung der Daten mancher Systeme (Hardeman et al. 2022).

Die Anwendung von Analysesystemen soll Tierärzte und Tierärztinnen nicht

ersetzen, aber sie können ein sinnvolles, zusätzliches Werkzeug sein, um bei Entscheidungen in der Lahmheitsuntersuchung zu helfen (van Weeren et al. 2017). Dem menschlichen Auge sind Grenzen gesetzt, was seine Wahrnehmung von geringgradiger Beschleunigung eines Objektes betrifft (Parkes et al. 2009) und auch eine gewisse Voreingenommenheit bei der Bewertung von Leitungsanästhesien konnte von Arkell et al. (2006) nachgewiesen werden. So können Analysesysteme sowie Untersuchungs- und Diagnostikprogramme eine Verbesserung und Standardisierung der Weiterbildung von Kolleginnen und Kollegen zum/zur Fachtierarzt/-ärztin oder der Ausbildung von Studierenden ermöglichen.

In Rahmen dieser Studie wurde der Einsatz von Pose Estimation (dt. räumliche Lageerkennung), einer markerlosen, nicht-invasiven Methode zur Aufzeichnung und Analyse der Bewegung von Menschen und Tieren, untersucht. Pose Estimation wird im Bereich von „Computer Vision“, der Verarbeitung und Analyse von Bilddateien und ihrer Inhalte durch einen Computer, eingesetzt. Es beschreibt das Erkennen und Markieren von Personen bzw. Tieren auf einem Foto oder Video (Andriluka et al. 2014). In Kombination mit der Nutzung von Deep Learning (dt. mehrschichtiges Lernen), einem Wissenschaftsbereich der Künstlichen Intelligenz (KI), bietet sich ein neuer Ansatz, um einige Nachteile von Lahmheitserkennungssystemen zu verbessern. Mit Hilfe eines neuronalen Netzwerkes ist es möglich, Muster in hochdimensionalen Datensätzen zu erkennen. Ziel dieser Studie war es, ein Programm zu entwickeln, das anatomische Referenzpunkte in einem sich bewegenden, mit einer Smartphone-Kamera gefilmten Pferd erkennen und diese Bewegungsdaten aufzeichnen kann. Mittels Berechnungen auf Grundlage dieser Aufzeichnungen war die Unterscheidung von gesunden sowie Pferden mit Vorder- und Hinterhandlahmheiten möglich. Basierend auf diesen Ergebnissen soll das Programm weiterentwickelt werden und durch die weitere Datenerhebung zu gesunden und lahmen Pferden ist es das langfristige Ziel, die Erkennung und Klassifizierung von Lahmheiten durch Algorithmen zuverlässig zu ermöglichen.

II. LITERATURÜBERSICHT

1. Lahmheit bei Pferden

1.1. Definition von Lahmheit

Veränderungen oder Unregelmäßigkeiten im physiologischen Gangbild eines Pferdes können unter dem Begriff „Lahmheit“ zusammengefasst werden. Ursache sind in der Regel strukturell oder funktionell bedingte pathologische Veränderungen im Bewegungsapparat (Ross 2011, Baxter et al. 2020). Es ist wichtig, dass eine medizinische Bewertung dieser Veränderungen im Gangbild anhand ihrer Ursache erfolgt. In den meisten Fällen werden sie durch Schmerzen verursacht und gehen mit den klinischen Zeichen der Entzündung einher, aber auch andere Gründe wie eine mechanische Einschränkung oder eine neurologische Dysfunktion können ursächlich sein (Baxter et al. 2020). Um die Lahmheit zu erkennen, zu lokalisieren und zu diagnostizieren und damit dem Pferd die bestmögliche Behandlung zukommen zu lassen, ist fundiertes anatomisches und biomechanisches Wissen in Kombination mit tiermedizinischer Erfahrung nötig (Ross 2011, Baxter et al. 2020).

1.2. Einteilung von Lahmheiten

1.2.1. Einteilung nach Art der Lahmheit

Grundsätzlich kann beim Pferd zwischen Stützbeinlahmheit, Hangbeinlahmheit und gemischten Formen unterschieden werden.

- 1) *Stützbeinlahmheit*: Sie ist die häufigste Lahmheitsform beim Pferd und wird durch eine verkürzte Standbeinphase im Vergleich zur kontralateralen Gliedmaße definiert (Ross 2011, Baxter et al. 2020, Dyson und Pollard 2020). Ursächlich können Veränderungen am Knochen, an den Gelenken, im Bandapparat oder im Huf sein (Baxter et al. 2020).
- 2) *Hangbeinlahmheit*: Einigen Autoren zufolge ist diese durch eine verkürzte Hangbeinphase charakterisiert und tritt in der Regel bei Pferden mit pathologischen Veränderungen an der proximalen Gliedmaße auf (Baxter et al. 2020). Andere bringen diese Art von Lahmheit ausschließlich mit einer mechanischen Einschränkung und damit mit bestimmten Erkrankungen wie

fibrotischer Myopathie, Hahnentritt (engl. Stringhalt) oder Patellafixation in Verbindung (Ross 2011).

- 3) *Gemischte Lahmheit*: Bei dieser Form ist neben der Standbeinphase auch die Hangbeinphase verkürzt; dabei können sowohl knöcherne als auch Weichteilstrukturen betroffen sein (Baxter et al. 2020). Nach Meinung von Ross (2011) ist dies die am häufigsten vorkommende Lahmheitsform und ist durch eine verkürzte Standbeinphase sowie eine verkürzte kraniale Hangbeinphase gekennzeichnet.

1.2.2. Einteilung nach Schweregrad

In den USA wird von der American Association of Equine Practitioners (AAEP) eine Skala verwendet, anhand derer die Lahmheit nach Sichtbarkeit in Schritt und Trab eingeteilt wird. Sie setzt sich aus fünf Lahmheitsgraden zusammen (American Association of Equine Practitioners 2022). Die gleiche Einteilung wird im Vereinigten Königreich verwendet, allerdings sind, wie in Tabelle 1 dargestellt, die fünf Grade dort in jeweils zwei Unterpunkte geteilt (Baxter et al. 2020). In Deutschland wird vermehrt die Einteilung der AAEP angewandt (Brehm et al. 2017).

Tabelle 1: Einteilung der Lahmheitsgrade (modifiziert nach (Wyn-Jones 1988, Baxter et al. 2020, American Association of Equine Practitioners 2022))

Einteilung nach AAEP		Einteilung UK	
Grad	Befunde	Grad	Befunde
0	Keine Lahmheit sichtbar.	0	Keine Lahmheit.
1	Die Lahmheit ist nur schwer erkennbar und unabhängig von den Gegebenheiten (z.B. unter dem Sattel, auf dem Zirkel, auf der Schrägen, auf hartem Boden) nicht kontinuierlich sichtbar.	1-2	Die Lahmheit ist im Schritt oder Trab nur schwer zu erkennen.
2	Die Lahmheit ist im Schritt und im Trab auf der Geraden schwer erkennbar, aber unter bestimmten Gegebenheiten (z.B. unter dem Sattel, auf dem Zirkel, auf der Schräge, auf hartem Boden) konstant sichtbar.	3-4	Die Lahmheit ist im Schritt nur schwer zu erkennen, im Trab gut sichtbar.
3	Die Lahmheit ist im Trab unter allen Gegebenheiten konstant sichtbar.	5-6	Die Lahmheit ist im Trab gut zu sehen.
4	Die Lahmheit ist im Schritt deutlich sichtbar.	7-8	Deutliche Lahmheit im Schritt, Traben ist kaum/nicht möglich.
5	Die Lahmheit führt zu einer minimalen Gewichtsbelastung in der Bewegung und/oder in Ruhe oder zu vollständiger Bewegungsunfähigkeit.	9-10	Das Pferd nimmt keine Last auf.

1.2.3. Einteilung in Vorder- und Hinterhandlahmheiten

Vorderhandlahmheiten

Pferde mit Lahmheit in der Vorhand zeigen ein charakteristisches Gangbild, das durch eine erhöhte Asymmetrie der Kopfbewegung in Relation zu Standbeinphase (Kopfnicken) charakterisiert ist (Buchner et al. 1996, Brehm et al. 2017). Um die Gewichtsbelastung auf dem lahmen Bein während dessen Standbeinphase zu verringern, nehmen sie in dieser den Kopf nach oben, wohingegen bei Belastung des gesunden Beines der Kopf nach unten genommen wird, um einen möglichst großen Anteil des Gewichtes auf die gesunde Körperseite zu verlagern. Diese Veränderung, sowohl in Bezug auf die Beschleunigung in der Bewegung als auch die Höhe des Kopfes in der Sagittalen, kann zur Diagnose einer Vorhandlahmheit verwendet werden (Buchner et al. 1996, Ross 2011, Baxter et al. 2020).

Hinterhandlahmheiten

Eine Lahmheit der Hintergliedmaße führt zu einer Asymmetrie in der vertikalen Bewegung des Beckens (Baxter et al. 2020) sowie zu einer Verschiebung des gesamten Beckens in Relation zu den Schritten der rechten und linken Hintergliedmaße. Außerdem kommt es zu einer Verschiebung der *Tubera coxae* in der Vertikalen im Verhältnis zueinander (May und Wyn-Jones 1987, Buchner et al. 1996, Kramer et al. 2000, Kramer et al. 2004). Wenn die Schmerzhaftigkeit beim Aufsetzen der Gliedmaße am größten ist, zeigt das Pferd ein vermindertes Absinken des Beckens auf der lahmen Seite, wohingegen ein vermindertes Abstoßen mit der Hintergliedmaße auf eine erhöhte Schmerzhaftigkeit in der zweiten Hälfte der Standbeinphase hindeutet (Baxter et al. 2020). Des Weiteren zeigen Pferde mit Lahmheit der Hintergliedmaße oft ein vermindertes Vorführen des lahmen Beines und eine verkürzte Standbeinphase (Buchner et al. 1995, Baxter et al. 2020).

Kompensatorische Lahmheit

In dem Versuch, die Gewichtslast auf der schmerzhaften Gliedmaße zu reduzieren, kann es zu einer vermeintlichen Lahmheit einer gesunden Gliedmaße kommen. Beispielweise kann eine ausgeprägte Vorhandlahmheit zu einer vermeintlichen Lahmheit der kontralateralen Hintergliedmaße führen (Baxter et al. 2020).

2. Klinische Lahmheitsuntersuchung

Jede Lahmheitsuntersuchung beginnt mit einer ausführlichen Anamnese und spezifischen Fragen zum zeitlichen Verlauf der Lahmheit. Neben dem Signalement, das Rasse, Geschlecht und Alter beinhaltet, sollte die Nutzung des Pferdes erfragt werden. Eine wichtige Grundlage für die weitere Untersuchung ist eine zeitliche Einordnung, also wann die Lahmheit erstmalig festgestellt wurde und ob im Rahmen der bestehenden Lahmheit bereits eine (tierärztliche) Behandlung erfolgt ist. Auch die Haltung des Pferdes, ob es in Boxenhaltung oder in einem (Aktiv-) Laufstall steht, und ob bereits Boxenruhe durch BesitzerInnen, PflegerInnen oder Tierärzte und Tierärztinnen initiiert wurde, ist ein wichtiger Teil der Anamnese. Außerdem können schon bekannte (orthopädische) Vorerkrankungen weiteren Aufschluss zur Ursache der Lahmheit geben (Ross 2011, Baxter et al. 2020). Im Rahmen jeder Lahmheitsuntersuchung sollte auch eine Allgemeinuntersuchung durchgeführt werden (Bartmann und Gehlen 2017, Davidson 2018).

Es folgt eine genaue Adspektion, während der das Pferd auf einer ebenen Fläche stehen soll. Diese kann dem Tierarzt oder der Tierärztin bereits wichtige Informationen über die Lahmheitsursache liefern. Das Entlasten eines Beines während der gesamten Untersuchung, sich wiederholendes Gewichtsverlagern von vorne nach hinten oder von rechts nach links oder Probleme, ein Bein mit dem vollen Körpergewicht zu belasten, können auf die betroffene Gliedmaße hinweisen. Weitere pathologische Auffälligkeiten wie asymmetrische Bemuskelung, Asymmetrien des knöchernen Skeletts vor allem der Hinterhand, Muskelatrophie oder ungleiches Hufwachstum können eine bereits länger bestehende Symptomatik anzeigen (Ross 2011, Davidson 2018, Baxter et al. 2020). Die Stellung der Gliedmaßen sollte sowohl von kranial und kaudal als auch von lateral begutachtet werden, um Hinweise auf eine ungleichmäßige Belastung der Hufe, knöcherner oder Weichteilstrukturen zu erhalten (Baxter et al. 2020).

Als nächstes wird das Pferd im Schritt und im Trab begutachtet. Als langsamer Viertakt ermöglicht der Schritt eine präzise Beurteilung des Bewegungsablaufes. Im Falle einer sehr geringgradigen Lahmheit kann es allerdings sein, dass sie nicht zu sehen ist, da die Schmerzgrenze in der Belastung nicht erreicht wird (Baxter et al. 2020). Pferde mit einer Hinterhandlahmheit können eine verkürzte Schrittlänge und das Unvermögen, über oder in die Spur des Vorderbeines zu treten, zeigen (Ross 2011). Obwohl die Bewegungen des Pferdes im Trab schneller sind, ist er aufgrund des Zweitaktes und seiner Diagonalität zur Überprüfung der Symmetrie in der Bewegung gut geeignet. Die linke und rechte Seite des Pferdes werden in Bezug auf Asymmetrie in der Belastung und der Schrittlänge miteinander verglichen (Davidson 2018, Baxter et al. 2020). Zusätzlich zur Adspektion des Pferdes auf einer geraden Linie sollte dies auch auf einer Kreisbahn erfolgen. In der Biegung muss es vermehrt die Innenseite belasten und aufgrund der Scherkräfte zwischen Huf und Boden kann es zur einer Verstärkung der Lahmheit, vor allem bei geringgradigen Lahmheiten, kommen (Hammarberg et al. 2016, Baxter et al. 2020). Das Laufen auf einer Kreisbahn führt immer zu einer asymmetrischen Belastung der außen- und innenliegenden Gliedmaßen, weshalb die erhobenen Befunde auf dem rechten und linken Zirkel miteinander verglichen werden müssen (Rhodin et al. 2013, Baxter et al. 2020).

Nach der Adspektion erfolgt eine ausführliche Palpation aller Beine, des Halses und des Rückens. Dieses Abtasten der Knochen- und darüber liegenden

Weichteilstrukturen ist wichtig, um Entzündungsanzeichen wie Schwellung, Wärme, Schmerzhaftigkeit oder Rötung zu erkennen (Ross 2011). An den Gliedmaßen werden die Gelenke im Seitenvergleich auf ihre Beweglichkeit, mögliche Schmerzreaktion beim Beugen sowie vermehrte Füllung überprüft (Baxter et al. 2020). Des Weiteren ist auf vermehrte Pulsation der *Aa. digitales palmares et plantares* zu achten und alle vier Hufe sollen mit einer Hufzange abgedrückt werden (Ross 2011, Davidson 2018).

Um die Lahmheitsursache weiter einzugrenzen, werden als nächstes Provokationsproben durchgeführt. In der Regel wird dabei die Gliedmaße für 30-60 Sekunden gebeugt und das Pferd anschließend auf einer geraden Linie vorgetrabt. Untersuchungen haben gezeigt, dass ein Beugen mit 100N zu wenig falschpositiven Ergebnissen führte (Verschooten und Verbeeck 1997), wohingegen andere Autoren auch 100-150N als geeignete Kraft angeben (Keg et al. 1997). Bezüglich der Zeitdauer sind sich die Autoren einig, dass 60 Sekunden einer ausreichenden Beugedauer entspricht (Keg et al. 1997, Verschooten und Verbeeck 1997). Alle Beugeproben einer Lahmheitsuntersuchung sollen von der gleichen Person durchgeführt werden, da es ansonsten schwierig ist, jede Beugeprobe mit exakt der gleichen Kraft und in der gleichen Art und Weise zu wiederholen (Keg et al. 1997, Verschooten und Verbeeck 1997, Baxter et al. 2020). Die Beugeprobe wird zuerst am kontralateralen, nicht-lahmen Bein durchgeführt (Ross 2011).

Es ist schwer, den exakten schmerzhaften Bereich mit einer Beugeprobe zu lokalisieren, da es fast unmöglich ist, nur auf eine einzelne anatomische Struktur Druck auszuüben (Ross 2011). Eine Differenzierung durch Unterteilung der Beugeproben in Beugung der distalen Gliedmaße, d.h. Huf-, Kron- und Fesselgelenk, und Beugung der proximalen Gliedmaße kann erfolgen. An der Vordergliedmaße können das Karpalgelenk weitgehend isoliert gebeugt, Ellbogen- und Schultergelenk jedoch nicht voneinander getrennt gebeugt oder gestreckt werden. Auch an der Hintergliedmaße können Tarsus sowie Knie- und Hüftgelenk aufgrund anatomischer Gegebenheiten nicht vollständig voneinander isoliert manipuliert werden (Ross 2011, Baxter et al. 2020). Die Spannsägenkonstruktion des Pferdes, bestehend aus dem *M. fibularis tertius* und dem *M. flexor digitalis superficialis*, erlaubt nur ein gleichsinniges Beugen und Strecken von Knie- und Sprunggelenk (Budras et al. 2012).

Mit Hilfe diagnostischer Anästhesien kann der Ursprung der Lahmheit genauer lokalisiert werden (Davidson 2018). In der Regel beginnt man mit einer perineuralen Infiltration, auch Leitungsanästhesie genannt, im Bereich der distalen Gliedmaße (Bassage und Ross 2011). Die genaue Reihenfolge und die Art der Anästhesien sind abhängig von den bisherigen Befunden und der Erfahrung des Tierarztes. Ein frühes Anästhesieren im proximalen Bereich kann eine genaue Lokalisation der Lahmheit verhindern (Bassage und Ross 2011, Baxter et al. 2020).

Im letzten Schritt der Lahmheitsuntersuchung muss entschieden werden, welche Art von bildgebender Diagnostik aufgrund der Befunde am zielführendsten ist und die meisten Informationen liefert. Abhängig von der darzustellenden anatomischen Struktur kann zwischen Röntgen, Ultraschall, Computertomographie, Magnetresonanztomographie oder Szintigraphie gewählt werden (Rantanen et al. 2011). Alle bildgebenden Verfahren haben jeweils spezifische Vor- und Nachteile, auf die aus Gründen des Umfangs und der Relevanz in Bezug auf die vorliegende Studie nicht weiter eingegangen wird.

3. Deep Learning und Pose Estimation

3.1. Künstliche Intelligenz in der Human- und Tiermedizin

Im Bereich der Bilderkennung und Diagnostik in der Humanmedizin werden Deep Learning-Algorithmen sehr erfolgreich eingesetzt. Vor allem in den letzten Jahren wurden auf diesem Gebiet große Fortschritte erzielt (Gulshan et al. 2016, Quellec et al. 2017, Yala et al. 2019). Ziel dabei ist es, dass die Analyse durch den Algorithmus möglichst der Vorgehensweise von HumanmedizinerInnen sowie Tierärzten und Tierärztinnen entspricht. Für das Training des Algorithmus wird eine große Menge an Bilddaten benötigt, wofür sich vor allem MRT- oder CT-Aufnahmen eignen, weil diese bereits digitalisiert und abgespeichert vorliegen. Im Rahmen von Routineuntersuchungen, beispielsweise in der Dermatologie, werden große Mengen an Bildern ohne pathologische Auffälligkeiten gesammelt, die für das Training verwendet werden können. Je größer und vielfältiger die Menge an Trainingsdaten ist, desto besser ist später der Algorithmus im Erkennen von Pathologien. Es gibt bereits Netze, die im Vergleich zu spezialisierten Dermatologen und Dermatologinnen ähnliche Leistungen bei der Erkennung von Hautkrebs erreichen (Esteva et al. 2017). Die Diagnosestellung in diesem Bereich erfolgt primär visuell, weshalb sich Deep Learning für diese Anwendung gut eignet.

Im Bereich der Onkologie werden Algorithmen in der Mammographie eingesetzt, indem sie Röntgenbilder analysieren und potenziell relevante Bereiche markieren. So müssen Radiologen und Radiologinnen nur noch einen Teil der Bilder selbst auswerten (Yala et al. 2019).

Auch in der Veterinärmedizin werden Künstliche Intelligenz (KI)-gestützte Programme und Deep Learning bereits vielseitig eingesetzt, sowohl in der Forschung als auch in der Bildgebung und Histopathologie (Bertram et al. 2019, Biercher et al. 2021, Aubreville et al. 2022). Bertram et al. (2019) und Aubreville et al. (2022) haben Mitosestadien in Mastzelltumoren untersucht, um ein Tumorstaging zu evaluieren. Im Fachbereich Radiologie hat eine Gruppe deutscher und englischer Kleintierradiologen und -radiologinnen in Zusammenarbeit ein Programm entwickelt, das bei der Auswertung von MRT-Aufnahmen von Hunden zwischen verschiedenen Bandscheiben- und Wirbelsäulenerkrankungen sowie gesunden Hunden differenzieren kann (Biercher et al. 2021). Im Bereich der Pferde-Ophthalmologie wurde eine Software mit Fotoaufnahmen von Pferdeaugen trainiert, um unterscheiden zu können, ob es sich um ein gesundes Auge handelt oder ob Hinweise auf ERU (Equine rezidivierende Uveitis) oder eine andere Augenerkrankung vorliegen (May et al. 2022). Im Gegensatz zur repetitiven Bilddatenerkennung sind Bewegungs- und Verhaltensstudien am Tier sehr zeitaufwändig, da in der Regel große Mengen an Videomaterial ausgewertet werden müssen oder sogar Personen in die direkte Überwachung vor Ort eingebunden sind. Die Möglichkeit, auch ohne Sensoren am Tier Bewegungen erfassen und auswerten zu können, ermöglicht auch Studien sehr kleiner Lebewesen wie *Drosophila*, ohne dass Filmmaterial manuell und damit sehr zeitaufwändig analysiert werden muss (Mathis et al. 2018). Um das Schmerzverhalten von Pferden zu untersuchen, wurde in Schweden in einer Studie deren Bewegungsverhalten innerhalb der Box während eines Klinikaufenthaltes aufgezeichnet. Anhand von gelabelten Standbildern, auf denen Nüstern, Widerrist und Schweif gekennzeichnet waren, wurde ein Algorithmus trainiert (Kil et al. 2020). So konnten Unterschiede im Verhalten eines kolikoperierten Pferdes über mehrere Tage aufgezeichnet und analysiert werden (Kil et al. 2020). Pose Estimation zur Lahmheitserkennung wurde in einer Studie von Wang et al. (2021) evaluiert, in der das Programm DeepLabCut© (DeepLabCut – a software package for animal pose estimation; <https://github.com/DeepLabCut/DeepLabCut>) mit Standbildern und Videos lahmere

und nicht lahmer Pferde trainiert und anschließend eine kleine Anzahl an Pferden (lahmfrei: n=3, lahm: n=6) für die Analyse verwendet wurden. Unterschiede in der Bewegung des Kopfes, der Beine und der Kruppe zwischen lahmen und nicht lahmen Pferden konnten graphisch dargestellt werden. In der Studie wurden jedoch keine Messdaten zur Berechnung der Lahmheit veröffentlicht oder Angaben zum Lahmheitsgrad in Relation zur graphischen Darstellung gemacht.

3.2. Grundlagen zu Pose Estimation und Deep Learning

Für die Bewegungsanalyse an Menschen und Tieren gibt es verschiedene Systeme, die sowohl über Videoaufnahmen als auch mittels Messsensoren Daten erfassen können. Nachteilig sind zum einen der Zeitaufwand und, je nach System, eine Einschränkung für das Untersuchungsobjekt aufgrund mehr oder weniger invasiver Mess-Techniken. Es müssen Marker oder Sensoren an der Person oder dem Tier befestigt werden, was zum einen die Mindestgröße des Untersuchungsobjektes festlegt und zum anderen zu einer Beeinträchtigung in der Bewegung führen kann. Außerdem muss vor Beginn der Analyse festgelegt werden, welche Körperabschnitte untersucht werden sollen und ein nachträgliches Hinzufügen von Analysepunkten ist nicht möglich (Mathis et al. 2018).

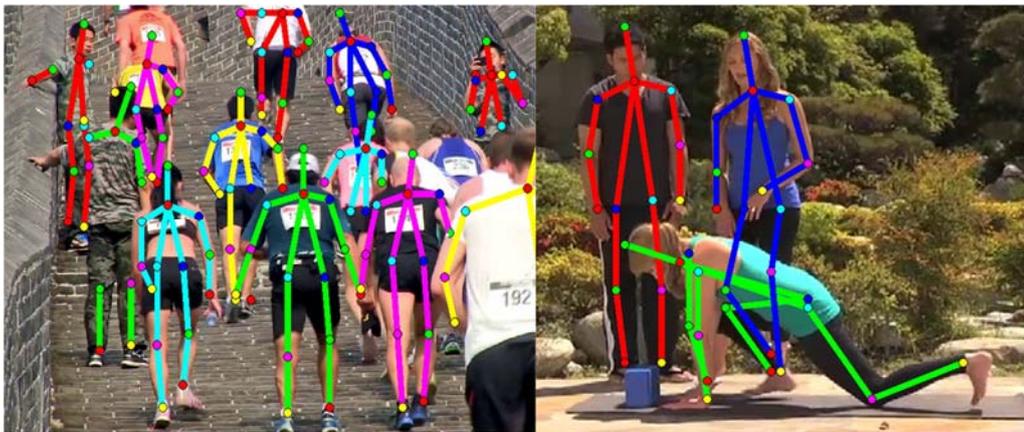
Die Entwicklung und Anwendung von Pose Estimation basiert auf Deep Learning, welches einen Teil des Wissenschaftsgebietes der KI repräsentiert. Bei dieser Art des maschinellen Lernens entsteht ein neuronales Netzwerk, das aus vielen einzelnen, miteinander verbundenen Schichten besteht, die als Layer bezeichnet werden. Jede Schicht ist für die Erkennung eines anderen Merkmales beziehungsweise Musters im Bild verantwortlich. Durch kontinuierliches Einfügen von Daten in das Netzwerk und das damit verbundene Trainieren ist es dem Algorithmus möglich, wiederkehrende Muster in hochdimensionalen Datensätzen zu erkennen. Im Unterschied zu anderen Programmen, deren Arbeitsweise durch den Programmierer und dessen genaue Festlegung von Funktionen geprägt ist, entwickeln sich das neuronale Netzwerk und die Filterkriterien, die es anwendet, durch den Algorithmus selbst (LeCun et al. 2015).

Für die Anwendung von Pose Estimation gibt es bereits verschiedene Algorithmen, die jeweils für unterschiedliche Zielgruppen geeignet sind. Im Bereich der menschlichen Bewegungserkennung liefern Programme wie Art-Track oder Open-Pose sehr gute Ergebnisse (Andriluka et al. 2014, Insafutdinov et al. 2017).

DeepLabCut© ist ein Software-Programm, das bereits sehr erfolgreich für Bewegungsstudien am Tier eingesetzt wurde (Nath et al. 2019). Als Entwicklungsgrundlage für DeepLabCut© wurde DeeperCut verwendet, der als einer der besten Algorithmen im Bereich der Pose Estimation gilt (Mathis et al. 2018). Einer der großen Vorteile von DeepLabCut© ist der verhältnismäßig kleine Datensatz an Bildern, der zum Trainieren des neuronalen Netzwerkes ausreicht. In diesem Datensatz werden in jedem einzelnen Bild die anatomischen Strukturen markiert, die für die Bewegungsanalyse von Interesse sind. Bereits ab einer Anzahl von ungefähr 200 Trainingsbildern kann man die Genauigkeit, mit der das Programm in der nachfolgenden Analyse Bilder labelt, mit menschlicher Präzision vergleichen. Im Gegensatz dazu benötigen andere Algorithmen eine deutlich höhere Anzahl an Bildern, wie zum Beispiel das MPII Human Pose mit rund 25.000 Bildern (Mathis et al. 2018). In Abbildung 1 sind zwei Beispielbilder aus dem MPII Human Pose Dataset dargestellt.

Abb. 1 Beispielbilder aus dem MPII Human Pose Data Set

(Quelle <https://pose.mpi-inf.mpg.de>, ©2014 Max-Planck-Gesellschaft zur Förderung der Wissenschaften e.V., München)



Es gibt verschiedene Arten von neuronalen Netzen, die als Artificial Neural Networks (ANNs) zusammengefasst werden können. ANNs bestehen aus einem zusammenhängenden Netz aus kleinen Recheneinheiten, sogenannten Neuronen und deren Verbindungen, die auch als Gewichte (engl. Weights) bezeichnet werden (Aghdam und Heravi 2017). Convolutional Neural Networks (CNNs), eine Untergruppe, sind neuronale Netze, die vor allem im Bereich des maschinellen Lernens eingesetzt werden (Lecun et al. 2010).

DeepLabCut© nutzt unter anderem Residual Networks (ResNets), eine Form von CNNs, für seine Funktionsweise (Mathis et al. 2018). CNNs werden im Bereich des

maschinellen Lernens eingesetzt und bestehen üblicherweise aus drei verschiedenen Arten von Schichten (engl. Layers), sogenannten Convolutional Layers, Pooling Layers und Fully Connected Layers (Sewak et al. 2018). Ein Layer setzt sich aus einzelnen künstlichen Neuronen zusammen (Aghdam und Heravi 2017). Convolutional Layer werden für die Erkennung bestimmter Merkmale in einem Bild verwendet, wobei jedes Layer für unterschiedliche Merkmale verantwortlich ist. Mittels Pooling Layers werden nicht benötigte Informationen herausgefiltert und ähnliche Features immer wieder zusammengefasst, was zu einer Komprimierung der Daten führt. Features sind die Eigenschaften der Eingabedaten, die für die Analyse von Interesse sind. Durch die Anwendung von Pooling Layers wird die Geschwindigkeit des neuronalen Netzwerkes erhöht und die Gefahr der Überanpassung (engl. Overfitting) reduziert (Aghdam und Heravi 2017, Sewak et al. 2018). Überanpassung beschreibt den Zustand eines Netzes, der entsteht, wenn es zu lange mit den gleichen Daten trainiert wird. Das Netz hat die Daten mehr oder weniger „auswendig“ gelernt und ist nicht mehr in der Lage, bekannte beziehungsweise erlernte Strukturen in unbekanntem Daten zu erkennen (Aghdam und Heravi 2017). Ein oder mehrere Fully Connected Layers verbinden alle Neurone des vorhergehenden Layer mit allen Neuronen des nachfolgenden Layer und schließen das Netzwerk ab (LeCun et al. 2015).

Ein Problem bei der Anwendung von CNNs ist, dass es mit zunehmender Tiefe des Netzwerkes, d.h. einer größer werdenden Anzahl an Layers, ab einem gewissen Punkt zu einer steigenden Fehlerrate kommt (He et al. 2016). CNNs werden über das sogenannte Zurückpropagieren (engl. Backpropagation) trainiert (Aghdam und Heravi 2017). Nach dem Einspeisen der Eingabedaten in das Netz wird der erwartete Ausgabewert mit dem tatsächlichen Ausgabewert am Ende des Trainings verglichen. Der Fehler des Netzes kann durch sogenannte „Loss Functions“ berechnet werden. Ziel ist es, sich dem Gradienten folgend einem minimalen Loss anzunähern, um damit den Fehlerwert zu minimieren. Gradienten beschreiben die Änderung einer definierten Größe, d.h. deren Anstieg oder Abnahme, auf einer definierten Strecke (The MathWorks Inc. 2022). Durch das Zurückpropagieren des Ausgabewertes werden die Verbindungen zwischen den Neuronen im Netz angepasst und das Netz auf diese Art trainiert. Innerhalb des Netzwerkes wird jeweils der vorherige Layer als Maß für die Anpassung des nachfolgenden Layers verwendet. Mit zunehmender Anzahl an Layers kommt es jedoch nicht zu einer

immer weiteren Verbesserung des Netzes, sondern zu einer steigenden Fehlerrate. Dieses Problem wird auch als „Phänomen des verschwindenden Gradienten“ bezeichnet. Im Training eines Netzwerkes ist das Erreichen eines möglichst kleinen Gradienten das Ziel, da dieser angibt, inwieweit sich die Ausgabewerte noch von den Eingabewerten unterscheiden und damit ein Maß für den Trainingszustand darstellt. Wenn jedoch einfach nur immer mehr Layer hinzugefügt werden, um das Netzwerk vermeintlich zu verbessern, kann es passieren, dass der Gradient allmählich gegen Null geht und kein weiteres Lernen des Netzwerkes stattfindet (Bengio et al. 1994, Glorot und Bengio 2010).

Dieses Problem kann mit der Anwendung von Residual Networks (ResNets) umgangen werden (He et al. 2016, Insafutdinov et al. 2016). Es werden mehrere Layer zu sogenannten Blocks zusammengefasst, was ein Überspringen mehrerer Schichten im Training ermöglicht. Ziel ist nicht mehr die Annäherung des Gradienten gegen 0 (wobei er den Nullwert nicht erreichen dürfte), sondern eine Darstellung des Zielwertes im Vergleich zum Ausgangswert, der dann die Differenz und damit den Trainingszustand wiedergibt (He et al. 2016). DeepLabCut© verwendet vortrainierte ResNets, die zuvor mittels ImageNet, einer Datenbank, die Daten für groß angelegte Bilderkennung anbietet, trainiert wurden (Mathis et al. 2018). Dadurch wird in der späteren Anwendung eine geringere Datenmenge für das Training des Algorithmus benötigt. Des Weiteren nutzt DeepLabCut© sogenannte Deconvolutional Layers, die genau gegensätzlich funktionieren wie Convolutional Layers (Zeiler et al. 2010). Daten, die zuvor durch das CNN herausgefiltert wurden und aufgrund deren Fehlen die Detailerkennung in der späteren Anwendung des Netzes verschlechtert werden würde, werden erneut gefiltert und gesammelt (Zeiler et al. 2010, Insafutdinov et al. 2016). Deconvolutional Layer vermehren die eingeführte Datenmenge im Netzwerk und helfen so, aus den Eigenschaften der Eingabedaten ein genaueres Bild zu generieren (Mathis et al. 2018).

Für den Einsatz von DeepLabCut© wird ein Trainings- und ein Testdatensatz an Bildern unterschieden (Nath et al. 2019). Aus den Videos des Trainingsdatensatzes werden Standbilder herausgenommen und die relevanten Referenzpunkte manuell markiert, d.h. gelabelt, wobei ein gesetzter Referenzpunkt einem Pixel im Bild entspricht (DeepLabCut 2022). Die genaue Position der Punkte stellt den erwarteten Ausgabewert des Netzes dar. Für die Bewegungsanalyse werden beliebige, gut

sichtbare, anatomische Punkte gewählt (Mathis et al. 2018). Nach Analyse des Testdatensatzes können einzelne Bilder ausgewählt, korrigiert und erneut für das Training verwendet werden, was die Wahrscheinlichkeit für ein korrektes Setzen der Punkte in der weiteren Anwendung erhöht (Nath et al. 2019). Auch im weiteren Verlauf der Anwendung können Referenzpunkte hinzugefügt werden (Nath et al. 2019). Als Parameter für den Trainingszustand des Algorithmus werden, wie oben beschrieben, Loss Functions verwendet. Eine Stabilisierung des RMSE (root mean square error) im Verlauf des Trainings zeigt, dass eine hohe Sicherheit im Testdatensatz besteht und eine weitere Verbesserung erst durch Hinzufügen neuer gelabelter Bilder erfolgen kann (Mathis et al. 2018).

4. Systeme für die Analyse von Lahmheiten

4.1. Entwicklung

In den letzten Jahren hat die Entwicklung von verschiedenen Systemen für die Analyse von Lahmheiten und des Gangbildes im Allgemeinen sehr stark zugenommen (van Weeren et al. 2017, van Weeren et al. 2018). Abhängig davon, ob sie auf kinetischen oder kinematischen Messtechniken basieren, können diese Systeme in zwei große Gruppen unterteilt werden. Die Kinetik beschreibt die Bewegung eines festen Körpers ausschließlich in Abhängigkeit der auf ihn einwirkenden Kräfte. Im Gegensatz dazu wird diese Bewegung in der Kinematik in räumlich-zeitlicher Abhängigkeit beschrieben und es werden sowohl die Zeit als auch die Strecke als messbare Parameter verwendet, ohne die Kraft zu berücksichtigen (Keegan 2007, Back und Clayton 2013, Bosch et al. 2018).

4.2. Systeme auf kinetischer Grundlage

Eines der ersten und immer noch verwendeten Systeme zur Lahmheitsanalyse ist eine Kraftmessplatte (Morris 1987). Mittels piezoelektrischer Sensoren, die die Bodenreaktionskraft eines Körpers erfassen, kann eine asymmetrische Verteilung des Körpergewichts auf die Gliedmaßen im Stand und in der Bewegung gemessen werden (Morris 1987, Byström et al. 2018). Nach dem 3. Newton'schen Gesetz übt jeder Körper auf den Boden, auf dem er steht, eine Kraft aus, die Gewichtskraft. Die Bodenreaktionskraft ist der Gewichtskraft genau entgegengesetzt und ihre Größe entspricht im Stand der Gewichtskraft. In der Bewegung setzt sie sich aus der Gewichtskraft und den Beschleunigungskräften zusammen (Budsberg et al. 1987, Back und Clayton 2013). Obwohl es möglich ist, sehr genaue Messdaten

aufzuzeichnen, ist die Lahmheitsanalyse mittels Kraftmessplatte sehr teuer und an einen festen Ort gebunden. Die Pferde müssen zuvor trainiert und an das Laufen auf dem Laufband gewöhnt werden. Außerdem können die Messungen nur in einer Klinik stattfinden, da die Kraftmessplatte in ein Betonfundament eingebettet ist. Um eine ausreichend große Anzahl an messbaren Tritten zu erzielen, muss das Pferd mehrfach über die Messplatte geführt werden (Keegan 2012). Nur wenn es das jeweilige zu untersuchende Bein exakt im Messbereich platziert, kann eine Analyse erfolgen (Keegan 2007, Oosterlinck et al. 2010, Back und Clayton 2013). Dennoch wird die Kraftmessplatte immer noch als Goldstandard in der Lahmheitsanalyse angesehen (Keegan 2012, Serra Bragança et al. 2018).

Eine Alternative ist eine Druckmessplatte, die, bei entsprechender Länge, mehrere aufeinanderfolgende Tritte analysieren kann (Oosterlinck et al. 2010). Mittels druckempfindlicher Polymer-Sensoren können die Druckverhältnisse im Verlauf der Standbeinphase bestimmt werden (van Heel et al. 2004). Es ist jedoch nicht möglich, die Bodenreaktionskraft zu untergliedern und durch das Einwirken von Scherkräften kann es zu einer Beeinflussung der erhobenen Daten kommen (Serra Bragança et al. 2018). Außerdem muss vor jeder Messung eine Kalibrierung erfolgen, was zeitaufwändig ist und die erzielten Messungen beeinflussen kann (Oosterlinck et al. 2010).

Ein weiteres System ist ein Hufschuh, der Bodenreaktionskräfte aufzeichnen kann. Da jedoch die Möglichkeit besteht, dass durch das zusätzliche Gewicht und die Größe des Schuhs die Bewegung des Pferdes beeinflusst wird, ist der Gebrauch für die Lahmheitsanalyse eingeschränkt (Keegan et al. 2004, Keegan 2007).

Mit einem speziell angefertigten Laufband, das in der Abteilung für Sportmedizin des Pferdes, Departement für Pferde, Tierspital der Universität Zürich, Schweiz verwendet wird, wurde eine Möglichkeit gefunden, einige Vorteile einer Kraftmessplatte und eines Laufbandes zu kombinieren. In das Laufband ist ein Kraftmesssystem integriert, welches die Bodenreaktionskraft einer Gliedmaße messen kann (Weishaupt et al. 2002, Byström et al. 2018). Auf diese Weise können viele aufeinanderfolgende Tritte von allen vier Beinen gleichzeitig aufgezeichnet und gemessen werden (Serra Bragança et al. 2018). Die Pferde müssen jedoch weiterhin zuvor an das Laufband gewöhnt werden und zusammen mit der Tatsache, dass eine solche Spezialanfertigung sehr teuer ist, ist es für den breiten klinischen Gebrauch daher ungeeignet (Keegan 2007, Keegan 2012, Serra Bragança et al.

2018).

4.3. Systeme auf kinematischer Grundlage

Es gibt mehrere Systeme, die auf kinematischen Grundprinzipien basieren und diese zur Lahmheitsanalyse nutzen. Hierbei wird die Bewegung eines Körpers oder einzelner Körperteile im dreidimensionalen Raum betrachtet. Die aufgezeichneten Parameter werden entweder zeitabhängig bestimmt (Geschwindigkeit) oder ortsabhängig, wenn Winkel oder zurückgelegte Distanzen gemessen werden (Keegan 2007, Keegan 2012, Serra Bragança et al. 2018). Die meisten der heute verwendeten kinematischen Systeme können einer von zwei großen Gruppen, Optical Motion Capture (OMC) und Inertial Measurement Unit (IMU) zugeordnet werden, neben wenigen anderen Systemen auf dem Markt.

4.3.1. Optical Motion Capture – Bewegungserfassung mittels Kameras

Bewegungserfassung (engl. Motion Capture) kann auf magnetischer, mechanischer oder optischer Grundlage erfolgen (Guerra Filho 2005). Magnetische Systeme besitzen Sensoren, die über Kabel mit der Recheneinheit verbunden sind und somit die Beweglichkeit des Untersuchungsobjekts einschränken, mechanische Systeme sind in spezielle Anzüge integriert, die am Körper anliegend die Bewegung direkt aufzeichnen (Guerra Filho 2005). Optical Motion Capture (OMC) ermöglicht präzise Bewegungsanalyse in Kombination mit vollständiger Bewegungsfreiheit und kann sowohl mit einem einzelnen Kamerasystem als auch mit mehreren Kameras verwendet werden. Bei der Anwendung eines monokularen Systems muss nach der Aufnahme eine dreidimensionale Rekonstruktion erfolgen, weshalb für die Bewegungsanalyse der Einsatz mehrerer Kameras empfohlen wird (Guerra Filho 2005). Des Weiteren ermöglicht OMC die Anwendung von markerbasierten und markerlosen Systemen (Guerra Filho 2005). Die ersten Studien, in denen eine Kamera zur Bewegungsanalyse beim Pferd verwendet wurde, stammen von Edward Muybridge um 1887. Mit Hilfe mehrerer Stoppuhren und eines elektromagnetischen Kreislaufes schaffte er es, die fotografische Belichtungszeit von einer halben Sekunde auf eine sechstausendstel Sekunde zu reduzieren (Back und Clayton 2013). Der technische Fortschritt der letzten Jahre hat einen großen Nutzen für die Kameras erbracht, die heute für Optical Motion Capture verwendet werden. In der Regel werden 3D-Infrarot-Kameras verwendet, die es mit einer Aufnahmeleistung von 100-300 Hz ermöglichen, bereits in einer Versuchsreihe

eine große Datenmenge zu erhalten (Serra Bragança et al. 2018). Für die Aufnahmen muss ein speziell gestalteter Gang, in dem die Pferde auf einer geraden Linie geführt werden können, oder ein Laufband verwendet werden. Die Kameras müssen in einer festen Position angebracht sein und regelmäßig neu kalibriert werden (Eichelberger et al. 2016). Die meisten OMC-Systeme arbeiten mit kugelförmigen Markern, die auf dem Fell des Pferdes fixiert werden (Barrey 1999, Back und Clayton 2013). Diese besitzen eine reflektierende Oberfläche und werden mit doppelseitigem Klebeband oder Kleber am Pferd fixiert oder an einer Gamasche befestigt (Rhodin et al. 2018, Serra Bragança et al. 2018, Hardeman et al. 2019). Im Rahmen einer solchen Messeinrichtung ist es möglich, mittels eines OMC-Systems Bewegungen im dreidimensionalen Raum sehr exakt aufzuzeichnen. Das kostspielige Equipment und die zeitaufwändige Vorbereitung und Platzierung der Kameras an einem Ort limitieren die Anwendung auf große Kliniken und Universitäten (Bosch et al. 2018). Marker, die an Kopf und Rumpf fixiert werden, ermöglichen das Erfassen von Minimal- und Maximalwerten in der vertikalen Bewegung der einzelnen Messpunkte (Kramer et al. 2004, Kelmer et al. 2005, Byström et al. 2018, Hardeman et al. 2019). Durch das Anbringen von Markern an den Gliedmaßen können weitere Parameter wie Schrittlänge erfasst werden (Byström et al. 2018).

Das System von QHorse, Qualisys AB (Göteborg, Schweden, www.qualisys.com/applications/equine-animal) ist sowohl im klinischen Gebrauch als auch in der Forschung in der Tierklinik Lüsche GmbH, Lüsche, Deutschland und in der Abteilung für Sportmedizin des Pferdes, Departement für Pferde, Tierspital der Universität Zürich, Schweiz verwendet wird (Byström et al. 2018, Hardeman et al. 2019). Die Anzahl an Markern und Kameras kann den Gegebenheiten angepasst werden, in der Studie von Hardeman et al. (2019) wurden 28 Kameras und neun Marker verwendet. Byström et al. (2018) haben 12 Kameras sowie 12 Marker für die Bewegungsanalyse eingesetzt. Ein weiteres System gibt es von Vicon® Motion Systems Limited (Unit 6, Oxford Industrial Park, Yarnton, Oxfordshire, UK, <https://www.vicon.com>), das mittels fünf fest installierter Infrarotkameras und zwei oder mehr reflektierender Marker die Bewegung eines Pferdes auf dem Laufband aufzeichnen kann (Keegan et al. 2004, Kramer et al. 2004, Kelmer et al. 2005). Die Marker können je nach Analyseschwerpunkt für die Vorhand am Kopf (*Proc. zygomaticus*) und am rechten Vorderhuf (Kelmer et al.

2005), für die Hinterhand mittig zwischen den *Tubera sacralia* und am rechten Hinterhuf fixiert werden, bei der gleichzeitigen Analyse von Vor- und Hinterhand ist eine Kombination dieser Punkte möglich (Kramer et al. 2004, Kelmer et al. 2005).

4.3.2. Inertial Measurement-Systeme

Inertiale Messeinheiten (engl. Inertial Measurement Units, IMUs) gehören zu der zweiten Gruppe der Messsysteme auf kinematischer Basis. Hier werden eine Kombination aus Beschleunigungssensoren und Gyroskopen für die Bewegungsanalyse verwendet (Keegan 2012, Bosch et al. 2018). Mittels eines Beschleunigungssensors wird die Beschleunigung entlang einer einzelnen oder mehrerer Achsen gemessen, wohingegen sich mit einem Gyroskop die Winkelgeschwindigkeit um eine Achse herum messen lässt (Keegan et al. 2004, Keegan 2007). Die Beschleunigungssensoren messen keine absolute Bewegung des Körpers im Raum, sondern die Beschleunigung senkrecht zu ihrer Oberfläche beziehungsweise ihren Oberflächen (Keegan et al. 2011). In der Regel senden die Sensoren ihre Messdaten an einen Rechner oder besitzen einen Zwischenspeicher, so dass eine kabellose Verwendung möglich ist. Die Fixierung am Pferd erfolgt über Riemen oder doppelseitiges Klebeband, die Anzahl der Sensoren und deren genaue Platzierung ist abhängig von dem jeweiligen System (Keegan et al. 2011). IMUs sind in ihrer Anwendung flexibler im Vergleich zu OMC-Systemen, da sie transportabel und damit fast überall einsetzbar sind. Nichtsdestotrotz ist für die Interpretation der Daten eine gewisse Expertise vonnöten und durch die Akkumulation von Driftfehlern können die Messdaten und damit die Ergebnisse der Untersuchung beeinflusst werden. Als Driftfehler bezeichnet man die Summe aller Messfehler, die im Rahmen einer Messfolge entstehen (Titterton und Weston 2004).

Equimetrix (keine Herstellerangaben verfügbar) ist ein einzelner Beschleunigungssensor, der an einem Bauchgurt am Pferd befestigt werden kann. In der Studie von Weishaupt et al. (2001) wird er als zweidimensionaler Sensor aufgeführt, bei López-Sanromán et al. (2012) wird ein dreidimensionaler Sensor genannt. Eine genaue Aussage bezüglich der Verfügbarkeit und der Sensoranzahl sowie Aufbau des Systems ist zum jetzigen Zeitpunkt nicht möglich, es konnten keine Herstellerangaben erfasst werden. Eine weitere Möglichkeit, mittels eines

einzelnen Sensors Hinterhandlahmheiten zu analysieren, ist die Anwendung eines iPhones. Der im Telefon eingebaute Sensor, bestehend aus Beschleunigungssensor und Gyroskop, kann die Minimal- und Maximalwerte der *Tubera sacralia* in der Vertikalen aufzeichnen (Marunova et al. 2021). Der Equinosis Q with Lameness Locator® (Equinosis LLC, Columbia, Missouri, USA, www.equinosis.com) verwendet zwei einachsige Beschleunigungssensoren, die am Genick und am *Sacrum* fixiert werden und die vertikale Beschleunigung des Kopfes und des *Sacrums* messen. Mittels eines einachsigen, piezoelektrischen Gyroskops, das an der rechten Vordergliedmaße an Huf oder Fesselkopf fixiert wird, kann die Winkelgeschwindigkeit an der Gliedmaße und damit die Standbeinphase bestimmt werden (Keegan et al. 2011, Leelamankong et al. 2020). Das Analyse-System von GaitSmart Pegasus (Codicote, Hitchin, Hertfordshire, UK; <https://www.gaitsmart.com/contact/>) besteht aus einem dreiachsigen Beschleunigungssensor und drei Gyroskopen, die mittels Gamaschen an den vier Gliedmaßen befestigt werden. Bis zum jetzigen Zeitpunkt ist der Autorin nur eine veröffentlichte Studie zu diesem System bekannt, die jedoch keine ausreichende Vergleichbarkeit mit einem der anderen Systeme zulässt (Roepstorff et al. 2013). Zwei Systeme der Firma XSens© Technologies B.V (Pantheon 6a, Enschede, Niederlande, <https://www.xsens.com>) wurden in verschiedenen Studien mit vier Beschleunigungssensoren (Church et al. 2009) beziehungsweise vier Beschleunigungssensoren und einem Gyroskop (Pfau et al. 2018) angewandt. Die Beschleunigungssensoren können an Genick, Widerrist, auf den *Tubera sacralia* und an je einem *Tuber coxae* fixiert werden, das Gyroskop wird zwischen den beiden *Tubera sacralia* mittels doppelseitigen Klebebandes befestigt (Church et al. 2009, Pfau et al. 2018). Das Equigait©-System (Equigait, United Kingdom, www.equigait.co.uk) besteht aus bis zu sechs Sensoren, die an Genick, Widerrist, den *Tubera sacralia*, an je einem *Tuber coxae* sowie zusätzlich im Lumbalbereich fixiert werden können. Der Autorin ist zum jetzigen Zeitpunkt keine Studie über den Einsatz von Equigait© im Vergleich zu anderen Systemen oder Tierärztinnen und Tierärzten bekannt. Mittels Beschleunigungssensoren, die an Kopf, Hals und Rumpf des Pferdes fixiert werden, können die Minimal- und Maximalwerte der Vertikalbewegung der einzelnen Messpunkte im Seitenvergleich gemessen werden (Keegan et al. 2011, Pfau et al. 2018). Das Equimoves©-System (Equimoves, Niederlande, <http://www.equimoves.nl/>), wie in Abbildung 2 gezeigt, besitzt acht Sensoren, von denen jeder einzelne sowohl einen Beschleunigungssensor als auch

ein Gyroskop enthält. Es werden jeweils vier der Sensoren an Genick, Widerrist, *Sternum* und *Sacrum* angebracht, die übrigen Sensoren werden an je einer Gliedmaße fixiert (Bosch et al. 2018). So können zum einen Asymmetrien in der Bewegung des Kopfes und des Rumpfes im Seitenvergleich bestimmt werden. Durch die zusätzlichen Sensoren an den Gliedmaßen ist es zum anderen möglich, Vor- und Rückfuhrwinkel der Röhrbeine im Verhältnis zum Karpus beziehungsweise Tarsus sowie Ad- und Abduktion zu messen, zusätzlich kann die Schrittlänge bestimmt werden (Bosch et al. 2018).

Abb. 2 Anwendung des Equimoves©-Systems am Pferd

(Quelle: <https://equimoves.nl/system-overview/applications/>, Foto by Christelle Althaus, ©2022 EquiMoves)



4.3.3. Radarsensoren

Am Weipers Centre, Glasgow Equine Hospital & Practice, University of Glasgow, Schottland wurde eine Studie zur Anwendung von Radarsystemen für die Bewegungsanalyse von Pferden durchgeführt (Shrestha et al. 2017). RADAR steht für „Radio Detection and Ranging“ und beschreibt das Orten und Erkennen von Objekten durch Radiowellen (Ludloff 1998). Das Radar strahlt über Antennen elektromagnetische Wellen in den Raum aus, die von Objekten reflektiert werden. Teile der Energie werden zurückgeworfen und von den Antennen registriert. Durch den Einsatz eines Dopplers kann die Frequenzverschiebung, die durch das sich bewegende Objekt entsteht, erfasst werden (Chen et al. 2014). Radar-Systeme werden bis jetzt mehrheitlich in der automatischen Zielerkennung eingesetzt, um zwischen Menschen und beispielsweise Tieren oder Fahrzeugen differenzieren zu können. Zur Bewegungsanalyse einzelner Spezies sind bisher wenig Daten

veröffentlicht. In der Studie von Shrestha et al. (2017) konnten Unterschiede in der Bewegung eines lahmen Pferdes im Vergleich zu einem lahmfreien Pferd dargestellt werden.

III. EIGENE UNTERSUCHUNGEN

1. Zielsetzung

Der Einsatz von Systemen zur Lahmheitsanalyse bei Pferden hat in den letzten Jahren deutlich zugenommen und ermöglicht, neben der Beurteilung durch den Tierarzt oder die Tierärztin, eine zusätzliche objektive Analyse des Gangbildes.

Ziel dieser Studie war es, den Einsatz von Pose Estimation für die Lahmheitsuntersuchung und Bewegungsanalyse bei Pferden zu evaluieren. Es wurde ein Programm entwickelt, das auf Grundlage von Deep Learning die Bewegung lahmer sowie nicht lahmer Pferde anhand von Videos erfassen kann und eine nachfolgende Analyse der Aufnahmen ermöglicht.

2. Publikation

Folgendes Manuskript wurde zur Publikation in Animals angenommen:

Submitted to section:

Equids

Equine Gait Analysis: Translating Science into Practice

Original Research / Research Article

The Usage of Artificial Intelligence for Lameness Detection in Horses – A Preliminary Study

Ann-Kristin Feuser¹, Stefan Gesell-May², Tobias Müller² and Anna May³

¹Equine Hospital in Parsdorf, Vaterstetten, Germany;

²anirec GmbH, artificial intelligence solutions in veterinary medicine;

³Equine Hospital, Ludwig Maximilians University, Munich, Germany 7

*) Corresponding author:

Prof. Dr. Anna May

Equine Hospital of Ludwig-Maximilians-Universitaet Munich

Sonnenstrasse 14

85764 Oberschleissheim

Germany

anna.may@pferd.vetmed.uni-muenchen.de



animals



Article

Artificial Intelligence for Lameness Detection in Horses—A Preliminary Study

Ann-Kristin Feuser, Stefan Gesell-May, Tobias Müller and Anna May

Special Issue

[Equine Gait Analysis: Translating Science into Practice](#)

Edited by

Prof. Dr. Hilary Clayton and Dr. Lindsay St. George



<https://doi.org/10.3390/ani12202804>



Article

Artificial Intelligence for Lameness Detection in Horses—A Preliminary Study

Ann-Kristin Feuser ¹, Stefan Gesell-May ², Tobias Müller ²  and Anna May ^{3,*}¹ Equine Hospital in Parsdorf, 85599 Vaterstetten, Germany² Anirec GmbH, Artificial Intelligence Solutions in Veterinary Medicine, 80539 Munich, Germany³ Equine Hospital, Ludwig Maximilians University, 85764 Oberschleissheim, Germany

* Correspondence: anna.may@pferd.vetmed.uni-muenchen.de

Simple Summary: In the expanding field of artificial intelligence, deep learning and smart-device-technology, a diagnostic software tool was developed, which can help distinguish between lame and sound horses and locate the affected limb. As lameness influences the welfare of horses and is often difficult to detect, this tool can help owners and veterinarians in the process of evaluation. The technology is based on pose estimation, which is already used in human and veterinary science to study movement of limbs or bodies without the need to fix any devices onto the object of interest. In this study, 22 horses with unilateral fore- or hindlimb lameness and a control group of eight sound horses were analysed with the program. Based on the results of the program, it was possible to differentiate between horses with fore- and hindlimb lameness and sound horses. Difficult light settings, such as direct sunlight or darkness, or very even-coloured coats, complicate the precise placement of reference points. The analysis and detection with software-generated movement trajectories using pose estimation is very promising but requires further development.



Citation: Feuser, A.-K.; Gesell-May, S.; Müller, T.; May, A. Artificial Intelligence for Lameness Detection in Horses—A Preliminary Study. *Animals* **2022**, *12*, 2804. <https://doi.org/10.3390/ani12202804>

Academic Editors: Hilary Clayton and Lindsay St. George

Received: 18 August 2022

Accepted: 12 October 2022

Published: 17 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Lameness in horses is a long-known issue influencing the welfare, as well as the use, of a horse. Nevertheless, the detection and classification of lameness mainly occurs on a subjective basis by the owner and the veterinarian. The aim of this study was the development of a lameness detection system based on pose estimation, which permits non-invasive and easily applicable gait analysis. The use of 58 reference points on easily detectable anatomical landmarks offers various possibilities for gait evaluation using a simple setup. For this study, three groups of horses were used: one training group, one analysis group of fore and hindlimb lame horses and a control group of sound horses. The first group was used to train the network; afterwards, horses with and without lameness were evaluated. The results show that forelimb lameness can be detected by visualising the trajectories of the reference points on the head and both forelimbs. In hindlimb lameness, the stifle showed promising results as a reference point, whereas the tuber coxae were deemed unsuitable as a reference point. The study presents a feasible application of pose estimation for lameness detection, but further development using a larger dataset is essential.

Keywords: artificial intelligence; deep learning; pose estimation; lameness; equine

1. Introduction

Lameness is a term that describes a horse's change in gait, usually caused by pain or mechanical restriction. There are substantial economic losses attributed to lameness in the equine industry, due to interrupted or truncated sports careers, costs of veterinary services, drugs and additional treatment costs, as well as death [1]. Lameness is one of the most common medical issue in equine veterinary medicine [2], and it can affect any horse at any level of training [3,4].

As undetected lameness poses a significant welfare issue for the affected horse, owners and veterinarians need to be capable of recognising changes of gait as early as possible.

Studies have shown that owners are often unable to recognise lameness in their own horses [5] and that identifying whether the horse experiences musculoskeletal pain resulting in lameness can be very difficult, especially for inexperienced riders [6]. On the clinical side, veterinary experience influences subjective lameness evaluation. Veterinary students and recent graduates often exhibit difficulties in identifying the affected leg [7]. Even amongst experienced veterinarians, there is often a lack of agreement on the affected leg in horses with subtle lameness cases [8,9]. Further limitations to subjective lameness evaluation are the inaccuracy of the human eye and the influence of bias due to the assessment and interpretation of lameness after diagnostic anaesthesia [10,11].

Over the years, many technology-assisted methods have been developed to objectively evaluate gait, movement and lameness in horses. These systems can be divided into two major groups, depending on whether they are based on kinetic or kinematic measuring techniques. Kinetics describes the movement of a rigid body, depending only on the action of forces. In contrast, kinematic analysis characterises the spatio-temporal movement of a rigid body, using time and distance as measurable parameters, without considering the forces [12–14].

One of the first kinetic instruments for analysing lameness, which is still used in research and clinical cases [15,16], is the force plate [17]. By recording the ground reaction forces from a lame horse, asymmetrical distribution of body weight on the legs can be measured [18]. Though offering very precise data, lameness analysis with the force plate is expensive, time consuming and only applicable in institutions where this measuring platform is available [12,13,19]. Nevertheless, it is still seen as the gold standard in equine lameness evaluation [20,21]. Other options include a force-measuring horseshoe, which can record ground reaction forces. However, the additional weight and size of the shoe potentially influences the movement of the horse, which reduces its value in lameness evaluation [13,22]. The instrumented treadmill located at the University of Zurich, Switzerland [15,18], offers the possibility to measure the ground reaction forces from several consecutive strides and from all four limbs [21]. Still, horses need to be trained to walk on the treadmill, which is time-consuming. In addition, because of its custom-made, relatively expensive characteristics, the treadmill is not suitable for broad clinical use in the field [13,20,21].

Most of the kinematic lameness evaluation systems can be assigned to one of two groups: optical motion capture (OMC) and inertial measurement unit (IMU). OMC systems use infrared cameras with a recording speed between 100–300 Hz, allowing the collection of a large amount of three-dimensional (3D) coordinate data [21]. Most OMC systems capture data using retro-reflective, spherical markers that are attached to the skin over anatomical locations of interest [12,23,24]. In this setup, an OMC system enables precise recording of 3D movement. However, the cost-intensive nature of the equipment and the time-consuming setup largely limits the use of OMC systems to large clinics and universities [14]. In contrast, the functionality of IMUs is based on gyroscopes and accelerometers [14,20]. Usually both sensors work wirelessly and are attached to certain body segments of a horse, using straps or double-sided tape [25]. The number of sensors and the exact placement differ across IMU systems. While a gyroscope measures the angular velocity around an axis, accelerometers measure the velocity and acceleration along a single axis or multiple axes [13,22]. Even though IMUs are portable, they are still relatively cost-intensive and require a certain level of expertise for data collection, analysis and interpretation. Furthermore, the accumulation of drift errors, which are the sum of all minor measuring errors during one analysis, can influence the results and thereby the outcome of the examination [26].

In the last few years, there has been increasing development of these systems [27,28]. Considering the fact that they require markers or inertial sensors, which need to be fixed onto the object of interest, the studied body parts must be defined beforehand [29].

In this study, we attempt to combine pose estimation with lameness evaluation in horses. This offers a new approach that ameliorates some of the disadvantages of other objective lameness detection systems. The use of pose estimation offers a non-invasive

way to track and record movements for further analysis. The development and use of pose estimation are based on deep learning. As part of the broad scientific field of artificial intelligence (AI), deep learning creates a neural network of multiple layers which relate to each other. By constantly incorporating new data into the network, it can be trained to recognise patterns in high-dimensional data. The significant difference in comparison to other computer programs is the fact that the filtering criteria of these layers are built autonomously from the algorithm itself, instead of by a software engineer [30].

The aim of this study was to evaluate the usability of pose estimation for detecting and marking specific anatomical reference points, using cell-phone videos of horses being lunged on a circle line. A secondary aim was to determine whether pose estimation can be used to differentiate between sound horses and horses with fore- and hindlimb lameness. We hypothesise that, using reference points on the head and forelimbs, it is possible to distinguish between a forelimb-lame and a non-lame horse. Furthermore, we hypothesize that a differentiation between hindlimb-lame horses and non-lame horses by using the stifle and the tuber coxae as reference points is feasible.

2. Materials and Methods

2.1. Technology

2.1.1. Deep Learning

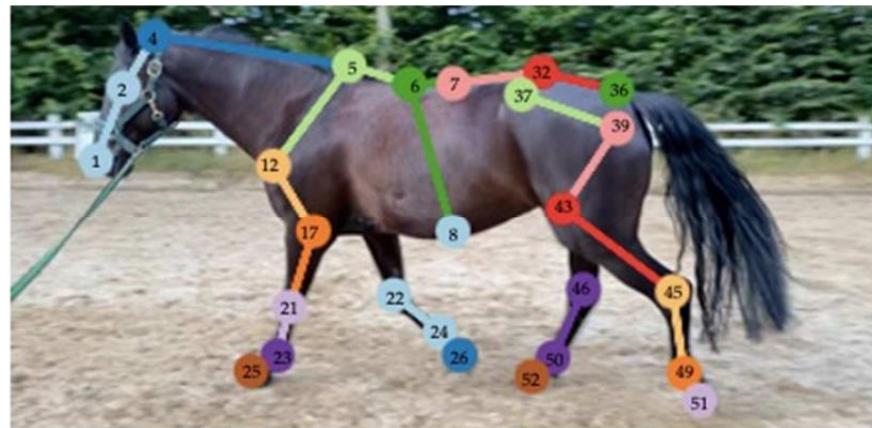
In veterinary science, deep learning is already used in many areas. It offers the possibility to improve behavioural studies, for example of drosophila flies or mice [29], or to aid in developing a pain detection model for stabled horses [31]. Other fields of application are image recognition in radiology, such as the automatic classification of canine thoracic radiographs [32], or in equine ophthalmology, integrated in a diagnostic application with a focus on equine uveitis [33].

2.1.2. Pose Estimation

Pose estimation allows for the tracking and recording of the movement of humans, animals, or objects without the need to fix any markers or sensors directly onto the subject of interest [29]. For the study of human poses, several well-described programs such as ArtTrack (Saarbrücken, Germany) or Open-Pose already exist [34,35]. After showing promising results in prior studies with pose estimation on animals, the DeepLabCut (2.2rc3 and 2.2.0.6: <https://github.com/DeepLabCut/DeepLabCut/tree/v2.2.0.6>; accessed on 10 August 2022) program was used in this study [36]. DeepLabCut is a deep convolutional network based on DeeperCut, which is considered one of the best algorithms for pose estimation. In contrast to other pose estimation tools, such as the MPII Human Pose dataset, with approximately 25,000 datasets, DeepLabCut only requires a relatively small number of 200 training images to train a network [29,37]. The functioning of DeepLabCut is based on two main elements. On the one hand, it uses pre-trained residual neural networks (ResNets), which are trained beforehand on ImageNet (resnet_50: http://download.tensorflow.org/models/resnet_v1_50_2016_08_28.tar.gz; accessed on 10 August 2022), a database that provides images for large-scale object recognition models. On the other hand, it is based on deconvolutional layers, which help to increase the visual information inserted into the network and reach spatial probability densities. After being trained with only a small number of labelled images (~200), the algorithm can predict and mark body parts with accuracy comparable to humans [29].

2.1.3. Reference Point Selection

For the pose estimation, 58 reference points, as listed in Figure 1, were determined. Selection criteria were identifiable anatomical landmarks on the horse, with some of these already used and proven in other lameness detection systems [14,38]. There were four markers on the head, four markers on the neck and trunk, 11 on each forelimb from the shoulder down to the hoof and 14 on each hindlimb between the tubera sacrale and the hooves. Each reference point corresponded to one pixel in one picture.



Head	Forelimb—proximal	Forelimb—distal	Hindlimb—proximal	Hindlimb—distal
1 Nostril	9 Spina scapulae left	19 Os carpi accessorium left	31 Croup middle	45 Tarsus left
2 Eye left	10 Spina scapulae right	20 Os carpi accessorium right	32 T. sacrale left	46 Tarsus right
3 Eye right	11 Tub. supraglenoidale left	21 Carpus left	33 T. sacrale right	47 Calcaneus left
4 Poll	12 Tub. supraglenoidale right	22 Carpus right	34 Kink left	48 Calcaneus right
Neck and trunk	13 Shoulder joint left	23 Fetlock left	35 Kink right	49 Fetlock left
5 Withers	14 Shoulder joint right	24 Fetlock right	36 Tail root	50 Fetlock right
6 Lowest back	15 Elbow hock left	25 Coronary band dorsal left	37 T. coxae left	51 Coronary band dorsal left
7 T18/L1	16 Elbow hock right	26 Coronary band dorsal right	38 T. coxae right	52 Coronary band dorsal right
8 Abdomen	17 Elbow joint left	27 Coronary band palmar left	39 Coxofemoral joint left	53 Coronary band plantar left
	18 Elbow joint right	28 Coronary band palmar right	40 Coxofemoral joint right	54 Coronary band plantar right
		29 Hoof tip left	41 T. ischiadicum left	55 Hoof pad left
		30 Hoof tip right	42 T. ischiadicum right	56 Hoof pad right
			43 Stifle joint left	57 Hoof tip left
			44 Stifle joint right	58 Hoof tip right

Figure 1. Reference points. Different combinations of reference points can be chosen in the program and offer multiple variations for gait analysis; the picture only shows a selection of the reference points which are enlarged in the image for better visibility. In the program, one reference point corresponds to one pixel. The accurate anatomical locations corresponding to the reference points of the program are listed in Table A1.

2.2. Collection of Data in Investigated Groups

All horses used in this study were assigned to one of three groups: one training group, one analysis group for lame horses and one analysis group for non-lame horses. Detailed information regarding all three groups is summarised in Table A2. Ethical approval for

this study was obtained from the ethics committee of Ludwig Maximilians University, Munich, Germany.

Every horse of the three groups received a full orthopaedic lameness examination [39,40] by an orthopaedic specialist (German specialists for equine medicine), including flexion tests. All horses were examined on hard and soft ground in walk and trot on the straight line and on the circle. Horses with any sign of visible gait asymmetry, a positive flexion test or any pathological results in the lameness examination were excluded.

Lameness results were graded according to the AAEP lameness scale by the American Association of Equine Practitioners on a scale from 1 to 5.

All horses of the training group ($n = 65$) were filmed in various environmental surroundings, which included eight different indoor and 14 different outdoor riding arenas with varying sand and soil surfaces. In order to obtain high recognition probabilities on the labelled reference points, diversity in the coat colour of the horses and environmental backgrounds was necessary. Furthermore, care was taken to film in different weather conditions, such as under sunlight or clouded skies, and during different times of the day to obtain a broad spectrum of different video settings. Horses were recorded in walk and trot from the front, the back (11 s in walk and 7 s in trot, respectively), and from both sides on a straight line (12 s in walk and 7 s in trot, respectively). Horses were also recorded on a circle line with an approximate diameter of 12 m on soft ground (1 min in walk and trot) on both hands.

All horses included in the lame group were privately owned horses presented for lameness examination in the Equine Hospital in Parsdorf, Vaterstetten, Germany. In total, 22 horses were examined and included. Permission for the collection and use of data was obtained from the owners beforehand, and detailed information about the lameness history of the horses was documented. As part of the routine lameness examination in this clinic, the horses were first filmed in walk and trot on both hands for one minute on a 12 m diameter circle on soft ground. After performing flexion tests on concrete and examining gait on firm, as well as on soft, ground, horses were subjected to diagnostic anaesthesia. Depending on the results of the examination and the identified anatomical area, the horses underwent diagnostic imaging (radiographs, ultrasound, computed tomography) and treatment based on the diagnosis. The recorded lameness grades varied from 1 to 4 (AAEP). Horses with a lameness degree $\geq 4/5$ were excluded from the study, as well as horses that showed lameness on more than one leg.

The non-lame group represents the reference group and consisted of eight horses. All horses were privately owned by one owner/farm. The horses were filmed in walk and trot on a left (CL) and right (CR) circle line for one minute in each gait. Two additional horses were excluded due to positive flexion tests after lameness had been detected during lunging. All video-recordings were taken with an iPhone 11 (Apple), with the resolution set to 1080 p and 30 fps.

2.3. Training the Artificial Intelligence Tool Using Deep Learning

2.3.1. Data Processing and Training

For training the neural network, 454 still frames from 215 videos of the training group were extracted and the predetermined points of interest (reference points, as defined in Section 2.1.3) were labelled manually. To provide high diversity in the training data, attention was paid to select still frames with different limb positioning combined with varying overlay of limbs. Multiple intermediary trainings were conducted to find a suitable network configuration for the neural network. Additionally, frames with predicted poses that had a significant number of outliers were determined and labelled manually to improve the performance of the network. For the final training set of 454 labelled still frames, the ResNet50 network base architecture was utilised. Five percent of the images were reserved for evaluation during training. These images were used to survey the training status of the algorithm. As this application only had access to a limited amount of training data, the evaluation ratio was left at this default value. All hyperparameters related to the neural

network and training process were set to the default values of DeepLabCut. This was to ensure that the neural network in this study was based on the stable results of DeepLabCut, using pre-trained and tested networks [29].

Initial tests were conducted using full resolution images (1920×1080 pixels) to preserve as much of the details as possible, but stable results could not be achieved. By reducing the resolution of the input images, a significant improvement in training was reached. In the end, a resolution of 768×432 pixels, which is 40% of the resolution of the original images, was chosen. This represents a balance of reduced image size without losing too much detail. The latest neural network was trained with 550,000 iterations with a resulting loss of 0.0013 of the training data. This low value indicates that the model fit the training data well. During training, the intention is to reach a preferably low value which must not become zero. This would reveal that the algorithm has learned the data by heart.

However, a comparison of training and evaluation data with respect to error probability showed that there was an average error of 2.6 pixels for training data, compared to as many as 8.22 pixels for evaluation data. Given the resolution of 432 pixels in the vertical axis, this error can make a difference of up to $\sim 1.9\%$ between training and evaluation data. Removing outliers with a likelihood below 60% in the predicted points led to an average error for training data of 2.59 pixels and 6.14 pixels for evaluation data. The small difference in error values for the training data shows its already-high certainty, combined with a distinctly lower certainty on unseen evaluation data. For the setup in this study, the threshold for the exclusion of data was set at a certainty of 60% to obtain high reliability for reference point detection, combined with a low error rate.

2.3.2. Data Analysis and Measurements and Mathematical Calculations in Trot Videos

For the following analysis, only the trot data were used. Each video included one minute of filming time with an average number of 74 strides per video for Warmbloods and 84 strides for German Riding Ponies. All horses of the second group were subdivided in two categories: A = forelimb-lame, B = hindlimb-lame.

Forelimb Lameness

The movement pattern of forelimb lame horses is marked by certain, distinguishable alterations. When trotting, a forelimb-lame horse demonstrates a typical, iterative head nod compared to a sound horse [39–41]. In an attempt to shift weight away from the painful leg, a left forelimb-lame horse lowers its head when stepping on the sound right leg and lifts the head up when loading onto the lame left leg [40,41]. Thus, to detect forelimb lameness in this study, the movement of the two forelimbs in comparison with the motion of the head was recorded. Reference points on the forelimbs and the neck were chosen. Reference points 17 (Elbow joint left) and 21 (Carpus left) were used for CL, and 18 (Elbow joint right) and 22 (Carpus right) were used for CR. Reference point 4 (poll) shows the movement of the head during trotting on both circles. To be able to distinguish between the left and right stance phase, points 19 (Os carpi accessorium left), 20 (Os carpi accessorium right), 45 (Tarsus left) and 46 (Tarsus right) were selected. For each horse, the recorded trajectory of the reference points from CL and CR were extracted from the program in csv-files and presented in charts. These data were analysed visually.

Hindlimb Lameness

Horses with hindlimb lameness show significant changes in their kinematic pattern [42,43]. In this study, two separate analysis parameters were investigated based on these known changes.

Stifle Reference Point

Horses with hindlimb lameness often present with a decreased protraction of the lame limb [39,42,43]. To compare the step length of both hindlimbs, the horizontal movement of points 43 (Stifle left) and 44 (Stifle right) on CL and CR was recorded and measured. It was

estimated that horses with a hindlimb lameness show a shortened stride on the lame leg and, therefore, show a smaller difference between the measured minima and maxima of the stifle point on the lame side.

Tuber coxae reference point

As an approved reference point [41,44], the movement of the tuber coxae along the vertical axis was analysed. Studies have demonstrated that hindlimb-lame horses show an increased vertical displacement of the tuber coxae on the lame side [41,44,45]. Thus, it was estimated that horses with hindlimb lameness show a larger difference between the measured minima and maxima on the affected side.

For each horse, the recorded trajectory of the reference points from the CL and CR were extracted from the program in csv files and transferred into an Excel file (Microsoft Excel, Version 16.63.1). To avoid false results due to inaccurate placement of markers by the program, the maximum 5% (95–100%) and the minimum 5% (0–5%) of the recorded frames were excluded from the analysis. The maxima represent the highest measured values (90–95%) and the minima the lowest measured values (5–10%) of the stifle point and the tuber coxae points.

For the analysis of the stifle point, \overline{Max}_{St} (mean value of the stifle maxima) and \overline{Min}_{St} (mean value of the stifle minima) for every horse were calculated for the left and the right circle. The differences represent the length of the horizontal distance along which the stifle point is recorded during trotting on each circle:

$$DS_{St}(CL) = |\overline{Max}_{St}(CL) - \overline{Min}_{St}(CL)|$$

$$DS_{St}(CR) = |\overline{Max}_{St}(CR) - \overline{Min}_{St}(CR)|$$

For the analysis of the tuber coxae point, \overline{Max}_{Tcox} (mean value of the tuber coxae maxima) and \overline{Min}_{Tcox} (mean value of the tuber coxae minima) were calculated for both circles. The differences represent the length of the vertical distance between the highest and lowest tuber coxae values during movement on each circle:

$$DT_{Tcox}(CL) = |\overline{Max}_{Tcox}(CL) - \overline{Min}_{Tcox}(CL)|$$

$$DT_{Tcox}(CR) = |\overline{Max}_{Tcox}(CR) - \overline{Min}_{Tcox}(CR)|$$

In the next step the difference for the Stifle as a reference point was calculated to compare the CL and CR:

$$D_{St} = |DS_{St}(CL) - DS_{St}(CR)|$$

The values for the tuber coxae measurements were calculated the same way for comparison of CL and CR:

$$D_{Tcox} = |DT_{Tcox}(CL) - DT_{Tcox}(CR)|$$

Mean values \overline{D}_{St} were calculated by summing up the DS_{St} of the individual horses, which should be compared, and dividing them by the number of included horses.

Mean values \overline{D}_{Tcox} were calculated the same way with DS_{Tcox} .

2.3.3. Statistical Analysis

Diagnostic test properties based on the AI system in comparison to the clinical assessment (reference) were separately assessed for forelimb lameness, hindlimb lameness using the stifle reference point, and hindlimb lameness using the tuber coxae reference point, using 2×2 tables. Estimates for diagnostic sensitivity (SE) were calculated as the proportion of clinically lame horses that were correctly classified based on the AI results. Specificity (SP) was calculated as the proportion of clinically healthy horses that were cor-

rectly classified based on the AI results. Accuracy (ACC) was calculated as the proportion of correct (positive + negative) classifications based on the AI results. Positive predictive values (PPV), describing the probability that the AI positive result is correct, and negative predictive values (NPV), describing the probability that the AI negative result is correct, were evaluated. The agreement beyond chance (κ), a statistical value for quantifying inter-rater reliability, was used in this study to measure agreement between clinical scoring of the horses and classification based on the AI. Kappa scores were calculated on the basis of a 3×3 -table, including forelimb lameness, hindlimb lameness (only using stifle reference point data) and the non-lame control group. Finally, an overall accuracy (OA) was calculated as the percentage of all correctly classified horses based on the AI results [46].

3. Results

Of the 22 horses of the lame group, 13 horses were detected with forelimb lameness and nine horses with hindlimb lameness. The results of their analysis, together with the eight horses of the third group, are presented below.

3.1. Forelimb Lameness

In total, seven horses were diagnosed as left-forelimb-lame and six as right-forelimb-lame. The lameness degrees ranged from AAEP 1–2/5 in ten horses and AAEP 3–4/5 in three horses. As shown in Figure 2a), the upward and downward movement (“head nod”) of the poll reference point was visually correlated with the loading of the lame and the non-lame limb, respectively. The non-lame horses did not show any signs of repetitive up-and-down motion of the head, as illustrated in Figure 2b).

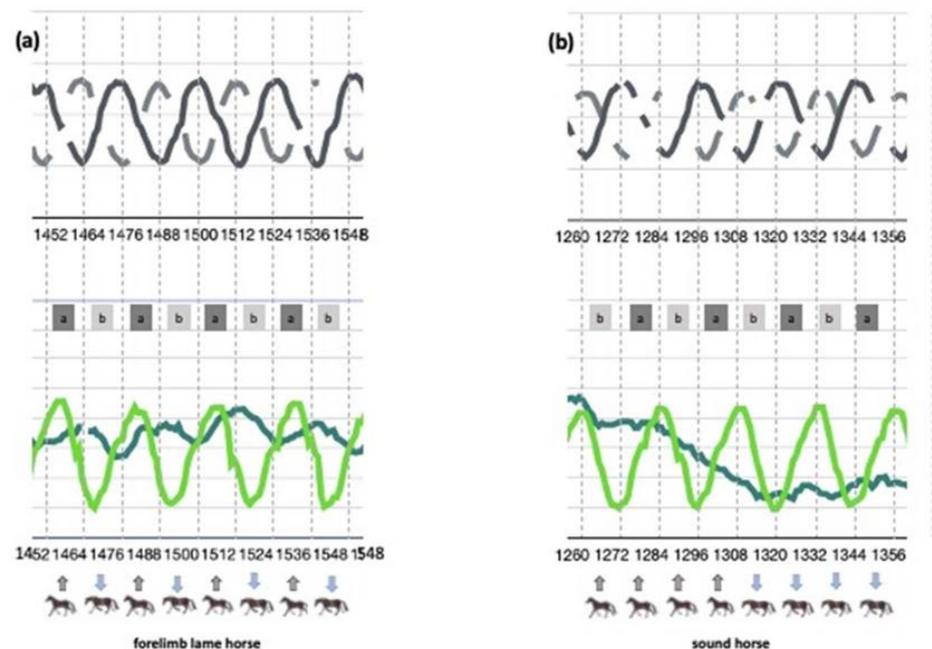


Figure 2. Graphical presentation of forelimb lameness in one representative horse (no. 11) (a) compared to a representative non-lame horse (no. 7), (b) on a left circle. a in the square = stance phase left forelimb, b in the square = stance phase right forelimb. Grey arrows indicate upward head movement, blue arrows indicate downward head movement. Upper graphs: grey lines show movements of left and right forelimb, with maximum values identifying the protracted foreleg = beginning of the stance phase (stride identification). Lower graphs: dark green line shows head movement, light green line shows movement of left forelimb; the numbers represent the frames of the video in the extracted sequence.

3.2. Hindlimb Lameness

The lameness degrees ranged from AAEP 1–2/5 in four horses and AAEP 3–4/5 in five horses. Five horses were lame on the left hindlimb, four horses were lame on the right hindlimb.

3.2.1. Stifle Reference Point

For every hindlimb-lame and every non-lame horse, the difference D_{St} was calculated. Results are presented in Tables 1 and 2. The median score of all D_{St} of the non-lame group was $\bar{D}_{St}(\text{non-lame}) = 0.55$. To verify detectability of hindlimb lameness with the stifle as reference point, a correlation between the lameness grade and the calculated D_{St} was constructed. After all videos were analysed, horses 2, 4, 7 and 9, were all classified with severe lameness and showed a clear difference in the calculated D_{St} compared to the median \bar{D}_{St} of the sound group. For horses 3, 5 and 8, graded with subtle lameness, a smaller difference in the calculated D_{St} compared to the median \bar{D}_{St} of the sound group could be illustrated. Therefore, a relation between the degree of lameness and the calculated D_{St} could be shown in all horses, except for horse 1.

Table 1. Stifle reference point—Hindlimb-lame horses.

Horse	Lameness	Degree of Lameness (1–5)		CL/CR	$DS_{St}(\text{CL})$ $DS_{St}(\text{CR})$	Difference $D_{St} = DS_{St}(\text{CL}) - DS_{St}(\text{CR}) $	Classified Lame Based on AI
		1–2	3–4				
1	LH		X	CL CR	42.50 44.17	1.67	No
2	RH		X	CL CR	42.17 34.32	7.85	Yes
3	RH	X		CL CR	31.16 29.69	1.47	Yes
4	LH		X	CL CR	47.68 54.61	6.93	Yes
5	RH	X		CL CR	43.32 42.09	1.23	Yes
6	LH	X		CL CR	36.20 38.21	2.01	Yes
7	LH		X	CL CR	48.03 51.12	3.09	Yes
8	LH	X		CL CR	47.36 49.60	2.24	Yes
9	RH		X	CL CR	49.90 38.55	11.35	Yes

RH = Right hindlimb, LH = Left hindlimb, CL = Circle left, CR = Circle right.

Table 2. Stifle reference point—Non-lame horses.

Horse	CL/CR	$DS_{St}(CL)$ $DS_{St}(CR)$	Difference $D_{St} = DS_{St}(CL) - DS_{St}(CR) $	Classified Sound Based on AI
1	CL	38.27	0.51	Yes
	CR	37.76		
2	CL	35.82	0.87	Yes
	CR	34.95		
3	CL	40.44	0.69	Yes
	CR	39.75		
4	CL	46.58	0.07	Yes
	CR	46.51		
5	CL	46.09	0.16	Yes
	CR	45.93		
6	CL	42.35	0.82	Yes
	CR	41.53		
7	CL	37.43	1.24	No
	CR	36.19		
8	CL	40.18	0.	Yes
	CR	40.18		

In the control group, with a calculated median $\overline{D}_{St} = 0.55$, all horses only showed small divergences in the comparison between CL and CR, except horse number 7.

3.2.2. Tuber Coxae Reference Point

For every hindlimb-lame and every non-lame horse, the difference D_{Tcox} was calculated. The results of the calculated D_{Tcox} for every hindlimb-lame horse are presented in Table 3, with the non-lame group in Table 4. The median score of all D_{Tcox} of the control group was $\overline{D}_{Tcox}(\text{non-lame}) = 1.30$. In three out of nine lame horses (horse 3, 5 and 9), the calculated D_{Tcox} corresponded with the lameness, as a larger difference between the measured minima and maxima on the lame side can be shown. In horses 1, 2, 4, 6, 7 and 8, D_{Tcox} indicated lameness on the contralateral non-lame limb. Comparing the median values of the detected lame, the non-detected lame and the non-lame horses, ($\overline{D}_{Tcox}(\text{lame}) = 1.21$, $\overline{D}_{Tcox}(\text{non-detected lame}) = 3.08$ and $\overline{D}_{Tcox}(\text{non-lame}) = 1.30$, respectively); therefore, no correlation between lameness, lameness grade and the absence of lameness could be drawn.

The mean values for SE, SP, ACC, PPV and NPV according to the analysis of the tuber coxae point of nine hindlimb-lame horses and eight non-lame horses are presented in Table 5. In comparison to the clinical assessment, the classification based on AI calculation was perfect (100% SE and SP) for forelimb lameness, close to 90% for hindlimb lameness when using the stifle reference point, but poor for hindlimb lameness when using the tuber coxae reference point (Table 5). The agreement beyond chance (κ) was $\kappa = 0.92573$. Due to the unreliable results and the inapplicability of tuber coxae as a reference point, it was excluded in this setup. An overall accuracy (OA) of 95.3% could be reached (Table A1).

Table 3. Tuber coxae reference point—Hindlimb-lame horses.

Horse	Lameness	Degree of Lameness (1–5)		CL/CR	$DS_{Tcox}(CL)$ $DS_{Tcox}(CR)$	Difference $D_{Tcox} = DT_{Tcox}(CL)$ $- DT_{Tcox}(CR) $	Classified Lane Based on AI
		1–2	3–4				
1	LH		X	CL CR	11.29 19.21	7.92	No
2	RH		X	CL CR	13.18 12.17	1.01	No
3	RH	X		CL CR	11.81 14.62	2.81	Yes
4	LH		X	CL CR	15.68 20.89	5.21	No
5	RH	X		CL CR	9.22 9.95	0.73	Yes
6	LH	X		CL CR	11.53 12.13	0.60	No
7	LH		X	CL CR	13.69 15.02	1.33	No
8	LH	X		CL CR	7.98 10.36	2.38	No
9	RH		X	CL CR	11.18 11.27	0.09	Yes

Table 4. Tuber coxae reference point—Non-lame horses.

Horse	CL/CR	$DS_{Tcox}(CL)$ $DS_{Tcox}(CR)$	Difference $D_{Tcox} = DT_{Tcox}(CL)$ $- DT_{Tcox}(CR) $	Classified Sound Based on AI
1	CL CR	11.13 11.82	0.69	Yes
2	CL CR	12.06 11.55	0.51	Yes
3	CL CR	14.28 19.06	4.78	No
4	CL CR	13.99 14.49	0.50	Yes
5	CL CR	11.38 11.81	0.43	Yes
6	CL CR	9.96 10.64	0.68	Yes
7	CL CR	8.45 9.59	1.14	No
8	CL CR	8.15 9.79	1.64	No

Table 5. Diagnostic test characteristics SE, SP, ACC, PPV and NPV of forelimb and hindlimb classification based on AI calculations when compared to the full clinical assessment (reference) in a study of 22 horses with lameness and eight horses without lameness (calculations of table contents based on Tables A3–A6)).

Test	True Positive	False Positive	False Negative	True Negative	SE (%)	SP (%)	AC (%)	PPV (%)	NPV (%)
Forelimb AI	13	0	0	8	100	100	100	100	100
Hindlimb AI stifle	8	1	1	7	88.9	87.5	88.2	88.9	87.5
Hindlimb AI tuber coxae	3	3	6	5	33.3	62.5	47.1	50	45.4

4. Discussion

In this study, the usability of an AI-based program and its capacity, based on the implementation of pose estimation, to detect specific anatomical landmarks of horses was evaluated. Calculations were made based on these data to differentiate between non-lame and unilateral fore- and hindlimb lame horses. Furthermore, the assessments made based on the program were compared to clinical lameness examination.

We believe that the use of a smartphone application in a real-world, equestrian setting would provide a great advantage to the standard lameness examination. Video analysis is non-invasive, and videos can be obtained at any chosen location with no equipment needed, except for a cell phone camera [29]. The ground surface and training facilities can therefore be those to which the horse is accustomed. This is particularly relevant, as studies have shown adaptations in equine movement and gait when, for example, a treadmill is used [12,47]. Videos obtained using a smartphone are easy to transfer via the internet and can be exchanged with veterinary colleagues all over the globe. Deep learning software is a tool which can help to detect fore- and hindlimb lameness in horses. By applying pose estimation to videos of horses filmed on a circle line and further evaluating the generated data, it is possible to detect lameness without additional hardware.

4.1. Forelimb Lameness

With the application of the reference points on the forelimbs and the head, forelimb lameness was detectable in this study. The data revealed head nodding as a result of increased weightbearing on the non-lame limb during stance. By contrast, horses within the non-lame control group did not show any consistent head movement asymmetry in rhythm with the steps onto the right or left forelimbs. A sensitivity and specificity of 100% shows that, by viewing the graphical charts, it is possible to differentiate a forelimb lame from a non-lame horse with this application. The next step will be a further development of the program to classify the extracted parameters of head and limb movement in relation to the stride time. This will allow calculation of the measured values and the collection of more specific data.

4.2. Hindlimb Lameness

For analysing hindlimb lameness in this setup, different equine anatomical landmarks on the hindlimbs were considered as reference points. In the pre-evaluation, reference points on the tuber coxae and stifle proved to be the most promising in the detection of hindlimb lameness. The tuber coxae have been used as a reference point in various locomotion studies [41,44,45], while the stifle has not been evaluated previously with portable systems in the horse, as it is not feasible to fix an accelerometer onto this point. To the authors' knowledge, it has been used as a reference point only in studies with OMC [42,48].

4.2.1. Stifle

In this study, a correlation between the degree of lameness and the calculated D_{St} could be shown in eight out of nine horses. Horse 1 displayed a slight difference between CL and CR, which did not correspond to its lameness grade (3–4). This horse was a dark-brown Warmblood with a very even-coloured coat. As mentioned below, the colour of the horses, especially when showing little or no variance, influences the accuracy of the reference points and, consequently, the results. Horse 7 of the control group was filmed during sunset in an outdoor riding arena and part of the arena was still covered in sunshine. This can affect the quality of the video with the sunbeams causing a glare effect. As mentioned above, the error rate for data evaluation was higher compared to the training data when these effects were present. Given the resolution of 432 pixels in the vertical axis, this error can make a difference of up to ~1.9%. Consequently, the reference points cannot be detected correctly in a few frames per circle, which results in a higher percentage of inaccurate placement. A sensitivity and specificity of almost 90% when using the stifle reference point provides promising results in this first setup. Using more labelled data will help to improve and stabilise the placement of the markers despite disadvantageous light conditions and horses with less well-defined anatomical landmarks.

4.2.2. Tuber Coxae

On the other hand, the tuber coxae point was not suitable for use with videos of horses on a circle line. Comparing the median values between the horses detected as being lame, the horses not detected as being lame and the non-lame horses, no correlation between lameness, lameness grade and the absence of lameness could be drawn. Other studies have shown that left and right tuber coxae should be compared at the same time to detect asymmetry [42,44,49]. As videos of horses on a circle line only show one side of the horse, a direct comparison using this setup was not possible. Furthermore, the large divergence of the calculated values in the control group confirms the fact that the tuber coxae are not suitable as a reference point for this purpose in the given setup.

Depending on the choice of reference points, the AI-based classification showed high to perfect agreement with the clinical assessment. The use of pose estimation reduces some of the limitations that contemporary lameness analysis systems must cope with. The EquiMoves system[®] (www.equimoves.nl, accessed on 10 August 2022) uses four sensors on the trunk and one sensor on each limb. It detects upper-body movement asymmetries in horses. In comparison with other systems that employ fewer IMU sensors, it is possible to determine stride length and certain limb angles for pro- and retraction and for ad- and abduction [14]. Nonetheless, the sensors must be fixed onto the horse, and the number of reference points is limited compared to the program evaluated in this study. Another IMU system is the Equinosis Q Lameness Locator[®], (Equinosis LLC, Columbia, MO, USA) which uses two accelerometers on the poll and tuber sacrale to measure the vertical maxima and minima of the head and pelvis during movement. A gyroscope attached to the right forelimb detects the stance phase to differentiate between movements of the left and right sides [25,50]. OMC systems such as QHorse from Qualisys Motion Capture Systems[®] (Qualisys AB, Motion Capture Systems, Göteborg, Sweden) allow marker fixation on different anatomical landmarks of the horse. With the need for a relatively large space to set up the cameras, evaluation and analysis of horses by this method are limited to large clinics and universities, reducing the flexibility and broad use of this system [18,51]. The use of pose estimation for equine gait analysis offers the possibility to record and analyse the movement of almost unlimited anatomical structures on a horse once the program has been adequately trained. Reference points can be selected before and after recording the horse and videos can be taken anywhere, with only a cellphone camera needed on site.

4.3. Limitations

There are some limitations in this study. Sample sizes were small, and larger studies on a broader range of patients are needed to derive robust estimates for SE and SP. To this

point, a differentiation of the anatomical origin of lameness is not possible due to small study groups and a limited amount of data. With improvement and advanced training of the program, further studies on the comparison of different causes of lameness are planned.

Using this software on a smartphone device, filming must be standardised, as multiple factors can affect the quality of the videos. As mentioned before, bright sunlight and shade lower the quality of the videos. This problem has also been discussed in other studies [29]. Consequently, the DeepLabCut software has been trained to learn how to robustly extract body parts, even with a cluttered and varying background, inhomogeneous illumination, or camera distortion [36]. In our study, evening light or bright sunshine made filming more difficult, and the analysed data became more imprecise. To evaluate the performance of the tool with videos that were not taken under perfect conditions, different light settings were considered. The horses were filmed inside equestrian arenas with windows and other light sources in different locations, as well as in outside riding arenas with different backgrounds (trees, fields, grass, traffic). Nonetheless, the diversity of videos used to train the AI system needs to be increased.

To find the most suitable filming position, 215 videos were evaluated. It showed that filming the horse, trotting on a straight line, from in front, behind, or from the side, did not offer enough steps for evaluation. However, videos filmed from the inner circle provided good consistency and a sufficient number of strides for analysis. In a complete lameness examination, horses should be evaluated on a straight line and on a circle line [39]. There are differences in motion of the torso and the pelvic area when horses' motions on a straight line and on a circle line are compared [39,52]. With further development and improvement of the program, it should be possible to analyse shorter video sequences on a straight line.

Irregular movements (horses shaking their heads, vocalising or becoming distracted and showing horizontal or vertical head movements) or other horses in the vicinity decreased correct positioning of reference points by the program. This effect did not have much impact on the results, as the chosen videos of horses on a circle line provided sufficient data to evaluate the lameness, despite data outliers.

When the coat or hoof colour of the horse resembled the background, the sand or the ground, it was difficult to recognise the anatomical markers and their locations became imprecise, so they could not be used. The anatomical structures were less prominent in horses that were completely black or white, especially when they were filmed in direct sunlight, so that labelling became demanding or even impossible in some cases, and they had to be excluded from the study. Apart from these rare cases, coat colour did not cause any selection bias; there was variation of colour in all three categories and a large colour spectrum was covered in non-lame and lame horses. The error rate increased when horses were over-weight or had a long winter coat that made anatomical structures less visible. By excluding the maximum and minimum 5% of the measured values, these small errors could be removed from the data. While the reference points were difficult to evaluate under the above circumstances, markers on the "edge" of the horse, as well as on easily visible anatomical structures, such as the nostril, eye or coronary band, were reproducible.

Another limitation was the quality of footing. Deep sand was unstable, causing horses to stumble or show irregular movements that could resemble lameness. This complicates any lameness examination and is not unique to this study. This needs to be considered with regard to the future use of the tool when videos taken by owners or inexperienced veterinarians will be used. As the volume of labelled data grows, the reliability of the program is expected to increase.

Evaluation of error values for training data showed that excluding outliers with a certainty below 60% only reduced the average error from 2.6 pixels down to 2.59 pixels, indicating that it is unlikely to improve with more training on the current model with the same data. It also shows that the network has high uncertainty on unseen evaluation data, which could be solved by having a greater variety of labelled images in the dataset. With additional augmentation through modification of the images, for example, by adding noise or changing colours or brightness, stability in difficult situations could be improved.

Additionally, with more data and different hyperparameters this error can be reduced in future iterations of the neural network.

4.4. Outlook for the Future

Pose estimation has the potential to improve gait analysis and lameness diagnostics in equine medicine and veterinary science. It can be applied to various gait or training assessments and can be used in various species such as horses, dogs, cats and dairy cattle. Studies have shown that dairy farmers do not recognise lameness in their cattle, even though it has a large impact on animal welfare, milk yield and, therefore, emerging costs [53,54]. With the help of this new, easily applicable pose estimation program, objective lameness evaluation can be efficiently executed, offering various possibilities for veterinary students and veterinarians to improve their abilities to assess horses' movements and, therefore, improve welfare for the affected animals [31,55].

Studies have shown that the quality of lameness examination improves with years of work experience, as veterinarians expand their skills and become better in detecting lameness [7]. In addition to these years of training, this tool may serve as a valuable system to improve learning quality and to refine and improve the veterinarian's ability to evaluate equine gait. Experienced veterinarians can use it for confirmation during daily clinical work and to keep records for retrospective evaluation of treatment. With increasingly more data being assessed and used to train the pose estimation tool, it may be possible to detect subtle gait changes, such as mild lameness or ataxia. Another possible use for the tool could be to compare different trainers or training methods. For example, gait analysis using all reference points to show swinging back movements or different swing-phase trajectories could be quantified to assess training efficacy.

5. Conclusions

This study demonstrated the feasibility of obtaining accurate measurements and data that match the clinical presentation in moderately lame horses (grade 3–4/5 AAEP). For horses that were only slightly lame (grade 1–2/5 AAEP), the smartphone app provided less distinct measurements, a sign that the program needs more labelled data and training to become more accurate and reliable. Furthermore, extended studies on the feasibility of the different reference points must be obtained, but these preliminary results are regarded as promising with regard to proof of concept.

Author Contributions: Conceptualization, A.-K.F., S.G.-M., A.M. and T.M.; methodology, S.G.-M. and T.M.; software, T.M.; validation, A.M., A.-K.F. and S.G.-M.; formal analysis, T.M.; investigation, A.-K.F.; resources, A.M., S.G.-M. and A.-K.F.; data curation, T.M. and A.-K.F.; writing—original draft preparation, A.-K.F. and A.M.; writing—review and editing, A.M.; visualization, A.-K.F., A.M. and T.M.; supervision, A.M.; project administration, A.M. and S.G.-M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The animal study protocol was approved by the Ethics Committee of Ludwig Maximilians University, Munich, Germany. (approval AZ 322-18-08-2022, August 2022).

Informed Consent Statement: Informed consent was obtained from the owners of the horses.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors would like to thank Marcus Doherr and Mathias Raths for valuable comments and help with statistics.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Reference points of the program with the correct anatomical location.

Reference Point in the Program	Anatomical Location	Reference Point in the Program	Anatomical Location
1. Nostril	nostril	30. Hoof tip right	hoof tip right forelimb
2. Eye left	left eye	31. Croup middle	midpoint between left and right tuber sacrale
3. Eye right	right eye	32. T. sacrale left	left tuber sacrale
4. Poll	poll	33. T. sacrale right	right tuber sacrale
5. Withers	withers	34. Kink left	midpoint between left tuber coxae and left tuber sacrale (view from behind)
6. Lowest back	lowest part of the dorsal line	35. Kink right	midpoint between right tuber coxae and right tuber sacrale (view from behind)
7. T18/L1	position of the 18th thoracic vertebra/first lumbar vertebra	36. Tail root	tail root
8. Abdomen	deepest part of the abdomen	37. T. coxae left	left tuber coxae
9. Spina scapulae left	scapular spine left	38. T. coxae right	right tuber coxae
10. Spina scapulae right	scapular spine right	39. Coxofemoral joint left	left coxofemoral joint
11. Tub. supraglenoidale left	supraglenoid tubercle left	40. Coxofemoral joint right	right coxofemoral joint
12. Tub. supraglenoidale right	supraglenoid tubercle right	41. T. ischiadicum left	left ischial tuberosity
13. Shoulder joint left	left shoulder joint	42. T. ischiadicum right	right ischial tuberosity
14. Shoulder joint right	right shoulder joint	43. Stifle joint left	left stifle joint
15. Elbow hock left	left elbow hock	44. Stifle joint right	right stifle joint
16. Elbow hock right	right elbow hock	45. Tarsus left	left tarsus
17. Elbow joint left	left elbow joint	46. Tarsus right	right tarsus
18. Elbow joint right	right elbow joint	47. Calcaneus left	left calcaneus
19. Os carpi accessorium left	left accessory carpal bone	48. Calcaneus right	right calcaneus
20. Os carpi accessorium right	right accessory carpal bone	49. Fetlock left	fetlock left hindlimb
21. Carpus left	left carpus	50. Fetlock right	fetlock right hindlimb
22. Carpus right	right carpus	51. Coronary band dorsal left	dorsal part of the coronet band left hindlimb
23. Fetlock left	fetlock left forelimb	52. Coronary band dorsal right	dorsal part of the coronet band right hindlimb
24. Fetlock right	fetlock right forelimb	53. Coronary band plantar left	plantar part of the coronet band left hindlimb
25. Coronary band dorsal left	dorsal part of the coronet band left forelimb	54. Coronary band plantar right	plantar part of the coronet band right hindlimb
26. Coronary band dorsal right	dorsal part of the coronet band right forelimb	55. Hoof pad left	heel bulb left hindlimb
27. Coronary band palmar left	palmar part of the coronet band left forelimb	56. Hoof pad right	heel bulb right hindlimb
28. Coronary band palmar right	palmar part of the coronet band left forelimb	57. Hoof tip left	hoof tip left hindlimb
29. Hoof tip left	hoof tip left forelimb	58. Hoof tip right	hoof tip right hindlimb

Table A2. Horses of Groups 1–3 (classified into sex, median age, median height, breed and colour).

		Group 1	Group 2	Group 3
Total Number		65	22	8
Sex	Mare	24	13	3
	Gelding	41	9	5
Median Age (in years)		13.8	11.6	12.4
Median Height (in meter)		1.60	1.61	1.62
Breeds	Warmblood	31	16	6
	Quarter Horse	7		
	PRE	5		
	Lusitano	3		
	Friese	1		
	Pinto	2		
	Knabstrupper	1		
	Arabian	1	1	
	Lewitzer	1		
	Haflinger	1		
	German Riding Pony	12	5	2
Colours	Black	8	1	
	Dark Bay	10	7	3
	Bay	11	6	3
	Chestnut	15	5	2
	Flaxen Chestnut	3		
	Buckskin	1		
	Palomino	3		
	Grey	4		
	White	4	2	
	Tobiano	5		
Leopard	1	1		

Table A3. 3 × 3-Table and statistical evaluation of κ (without reference point tuber coxae).

	Classified by AI Non-Lame	Classified by AI Forelimb-Lame	Classified by AI Hindlimb-Lame Stifle	Total
Clinically non-lame	20	0	1	21
Clinically forelimb-lame	0	13	0	13
Clinically hindlimb-lame stifle	1	0	8	9
Total	21	13	9	43

Appendix B

Table A4. Statistical classification of horses with and without forelimb lameness.

Forelimb Lameness	Clinically Forelimb-Lame	Clinically Non-Lame	Total	
AI classified as forelimb-lame	13	0	13	Positive predictive value 1
AI classified as non-lame	0	8	8	Negative predictive value 1
Total	13	8	21	
AI diagnostic test evaluation	Sensitivity of AI 1	Specificity of AI 1	Accuracy of AI 1	

Table A5. Stifle reference point—Statistical classification of horses with and without hindlimb lameness.

Hindlimb Lameness Stifle	Clinically Hindlimb-Lame	Clinically Non-Lame	Total	
AI classified as hindlimb-lame	8	1	9	Positive predictive value 0.88888889
AI classified as non-lame	1	7	8	Negative predictive value 0.875
Total	9	8	17	
AI diagnostic test evaluation	Sensitivity of AI 0.88888889	Specificity of AI 0.875	Accuracy of AI 0.882352941	

Table A6. Tuber coxae reference point—Statistical classification of horses with and without hindlimb lameness.

Hindlimb Lameness Tuber Coxae	Clinically Hindlimb-Lame	Clinically Non-Lame	Total	
AI classified as hindlimb-lame	3	3	6	Positive predictive value 0.5
AI classified as non-lame	6	5	11	Negative predictive value 0.454545455
Total	9	8	17	
AI diagnostic test evaluation	Sensitivity of AI 0.333333333	Specificity of AI 0.625	Accuracy of AI 0.470588235	

References

- Seitzinger, A.H. A comparison of the economic costs of equine lameness, colic, and equine protozoal myeloencephalitis (EPM). In Proceedings of the 9th International Symposium on Veterinary Epidemiology and Economics, Breckenridge, CO, USA, 6–11 August 2000; pp. 1–4.
- Nielsen, T.D.; Dean, R.S.; Robinson, N.J.; Massey, A.; Brennan, M.L. Survey of the UK veterinary profession: Common species and conditions nominated by veterinarians in practice. *Vet. Rec.* **2014**, *174*, 324. [[CrossRef](#)] [[PubMed](#)]
- USDA. *Part I: Baseline Reference of 1998 Equine Health and Management*; USDA: Washington, DC, USA, 1998; p. N280.898.

4. Slater, J. *National Equine Health Survey (NEHS) 2016*; Blue Cross for Pets: Burford, UK, 2016.
5. Muller-Quirin, J.; Dittmann, M.T.; Roepstorff, C.; Arpagaus, S.; Latif, S.N.; Weishaupt, M.A. Riding Soundness-Comparison of Subjective With Objective Lameness Assessments of Owner-Sound Horses at Trot on a Treadmill. *J. Equine Vet. Sci.* **2020**, *95*, 103314. [[CrossRef](#)] [[PubMed](#)]
6. Dyson, S.; Pollard, D. Application of a Ridden Horse Pain Ethogram and Its Relationship with Gait in a Convenience Sample of 60 Riding Horses. *Animals* **2020**, *10*, 1044. [[CrossRef](#)] [[PubMed](#)]
7. Starke, S.D.; May, S.A. Veterinary student competence in equine lameness recognition and assessment: A mixed methods study. *Vet. Rec.* **2017**, *181*, 168. [[CrossRef](#)]
8. Keegan, K.G.; Dent, E.V.; Wilson, D.A.; Janicek, J.; Kramer, J.; Lacarrubba, A.; Walsh, D.M.; Cassells, M.W.; Esther, T.M.; Schiltz, P.; et al. Repeatability of subjective evaluation of lameness in horses. *Equine Vet. J.* **2010**, *42*, 92–97. [[CrossRef](#)]
9. Fuller, C.J.; Bladon, B.M.; Driver, A.J.; Barr, A.R. The intra- and inter-assessor reliability of measurement of functional outcome by lameness scoring in horses. *Vet. J.* **2006**, *171*, 281–286. [[CrossRef](#)]
10. Parkes, R.S.; Weller, R.; Groth, A.M.; May, S.; Pfau, T. Evidence of the development of ‘domain-restricted’ expertise in the recognition of asymmetric motion characteristics of hindlimb lameness in the horse. *Equine Vet. J.* **2009**, *41*, 112–117. [[CrossRef](#)]
11. Arkell, M.; Archer, R.M.; Guitian, F.J.; May, S.A. Evidence of bias affecting the interpretation of the results of local anaesthetic nerve blocks when assessing lameness in horses. *Vet. Rec.* **2006**, *159*, 346–349. [[CrossRef](#)]
12. Back, W.; Clayton, H.M. 1. History. In *Equine Locomotion*, 2nd ed.; van Weeren, P.R., Ed.; Saunders Elsevier: Edinburgh, UK; New York, NY, USA, 2013; pp. 1–30.
13. Keegan, K.G. Evidence-based lameness detection and quantification. *Vet. Clin. North Am. Equine Pract.* **2007**, *23*, 403–423. [[CrossRef](#)]
14. Bosch, S.; Serra Bragança, F.; Marin-Perianu, M.; Marin-Perianu, R.; van der Zwaag, B.J.; Voskamp, J.; Back, W.; van Weeren, R.; Havinga, P. EquiMoves: A Wireless Networked Inertial Measurement System for Objective Examination of Horse Gait. *Sensors* **2018**, *18*, 850. [[CrossRef](#)]
15. Weishaupt, M.A.; Hogg, H.P.; Wiestner, T.; Denoth, J.; Stussi, E.; Auer, J.A. Instrumented treadmill for measuring vertical ground reaction forces in horses. *Am. J. Vet. Res.* **2002**, *63*, 520–527. [[CrossRef](#)] [[PubMed](#)]
16. Back, W.; Clayton, H.M. 9. Gait Adaption in Lameness. In *Equine Locomotion*; Buchner, H.H., Ed.; Saunders Elsevier: Edinburgh, UK; New York, NY, USA, 2013; pp. 175–197.
17. Morris, E.; Seeherman, H. Redistribution of ground reaction forces in experimentally induced equine carpal lameness. In *Equine Exercise Physiology*; Wiley: Hoboken, NJ, USA, 1987; pp. 553–563.
18. Byström, A.; Egenvall, A.; Roepstorff, L.; Rhodin, M.; Bragança, F.S.; Hernlund, E.; van Weeren, R.; Weishaupt, M.A.; Clayton, H.M. Biomechanical findings in horses showing asymmetrical vertical excursions of the withers at walk. *PLoS ONE* **2018**, *13*, e0204548. [[CrossRef](#)] [[PubMed](#)]
19. Oosterlinck, M.; Pille, F.; Huppés, T.; Gasthuys, F.; Back, W. Comparison of pressure plate and force plate gait kinetics in sound Warmbloods at walk and trot. *Vet. J.* **2010**, *186*, 347–351. [[CrossRef](#)] [[PubMed](#)]
20. Keegan, K.G. Objective measures of lameness evaluation. In Proceedings of the American College of Veterinary Surgeons Symposium, National Harbor, MD, USA, 1–3 November 2012; pp. 127–131.
21. Serra Bragança, F.M.; Rhodin, M.; van Weeren, P.R. On the brink of daily clinical application of objective gait analysis: What evidence do we have so far from studies using an induced lameness model? *Vet. J.* **2018**, *234*, 11–23. [[CrossRef](#)]
22. Keegan, K.G.; Yonezawa, Y.; Pai, P.F.; Wilson, D.A.; Kramer, J. Evaluation of a sensor-based system of motion analysis for detection and quantification of forelimb and hind limb lameness in horses. *Am. J. Vet. Res.* **2004**, *65*, 665–670. [[CrossRef](#)] [[PubMed](#)]
23. Barrey, E. Methods, applications and limitations of gait analysis in horses. *Vet. J.* **1999**, *157*, 7–22. [[CrossRef](#)]
24. Rhodin, M.; Persson-Sjodin, E.; Egenvall, A.; Serra Bragança, F.M.; Pfau, T.; Roepstorff, L.; Weishaupt, M.A.; Thomsen, M.H.; van Weeren, P.R.; Hernlund, E. Vertical movement symmetry of the withers in horses with induced forelimb and hindlimb lameness at trot. *Equine Vet. J.* **2018**, *50*, 818–824. [[CrossRef](#)]
25. Keegan, K.G.; Kramer, J.; Yonezawa, Y.; Maki, H.; Pai, P.F.; Dent, E.V.; Kellerman, T.E.; Wilson, D.A.; Reed, S.K. Assessment of repeatability of a wireless, inertial sensor-based lameness evaluation system for horses. *Am. J. Vet. Res.* **2011**, *72*, 1156–1163. [[CrossRef](#)]
26. Titterton, D.; Weston, J. 4 Gyroscope Technology 1. In *Strapdown Inertial Navigation Technology*; Institution of Engineering and Technology: London, UK, 2004; pp. 59–112.
27. van Weeren, P.R.; Pfau, T.; Rhodin, M.; Roepstorff, L.; Serra Bragança, F.; Weishaupt, M.A. Do we have to redefine lameness in the era of quantitative gait analysis? *Equine Vet. J.* **2017**, *49*, 567–569. [[CrossRef](#)]
28. van Weeren, P.R.; Pfau, T.; Rhodin, M.; Roepstorff, L.; Serra Bragança, F.; Weishaupt, M.A. What is lameness and what (or who) is the gold standard to detect it? *Equine Vet. J.* **2018**, *50*, 549–551. [[CrossRef](#)]
29. Mathis, A.; Mamidanna, P.; Cury, K.M.; Abe, T.; Murthy, V.N.; Mathis, M.W.; Bethge, M. DeepLabCut: Markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* **2018**, *21*, 1281–1289. [[CrossRef](#)] [[PubMed](#)]
30. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)] [[PubMed](#)]
31. Kil, N.; Ertelt, K.; Auer, U. Development and Validation of an Automated Video Tracking Model for Stabled Horses. *Animals* **2020**, *10*, 2258. [[CrossRef](#)] [[PubMed](#)]

32. Banzato, T.; Wodzinski, M.; Burti, S.; Osti, V.L.; Rossoni, V.; Atzori, M.; Zotti, A. Automatic classification of canine thoracic radiographs using deep learning. *Sci. Rep.* **2021**, *11*, 3964. [[CrossRef](#)] [[PubMed](#)]
33. May, A.; Gesell-May, S.; Muller, T.; Ertel, W. Artificial intelligence as a tool to aid in the differentiation of equine ophthalmic diseases with an emphasis on equine uveitis. *Equine Vet. J.* **2022**, *54*, 847–855. [[CrossRef](#)]
34. Insafutdinov, E.; Andriluka, M.; Pishchulin, L. ArtTrack: Articulated Multi-Person Tracking in the Wild. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 6457–6465. [[CrossRef](#)]
35. Cao, Z.; Simon, T.; Wei, S.-E. Realtime Multi-person 2D Pose Estimation Using Part Affinity Fields. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017.
36. Nath, T.; Mathis, A.; Chen, A.C.; Patel, A.; Bethge, M.; Mathis, M.W. Using DeepLabCut for 3D markerless pose estimation across species and behaviors. *Nat. Protoc.* **2019**, *14*, 2152–2176. [[CrossRef](#)]
37. Andriluka, M.; Pishchulin, L.; Gehler, P. 2D Human Pose Estimation: New Benchmark and State of the Art Analysis. In Proceedings of the IEEE Conference on computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014.
38. Keegan, K.G.; Wilson, D.A.; Kramer, J.; Reed, S.K.; Yonezawa, Y.; Maki, H.; Pai, P.F.; Lopes, M.A. Comparison of a body-mounted inertial sensor system-based method with subjective evaluation for detection of lameness in horses. *Am. J. Vet. Res.* **2013**, *74*, 17–24. [[CrossRef](#)]
39. Baxter, G.M.; Adams, O.R.; Stashak, T.S. *Adams and Stashak's Lameness in Horses*, 6th ed.; Wiley-Blackwell: Chichester, UK; Ames, IA, USA, 2011; p. xxviii. 1242p.
40. Ross, M.W.; Dyson, S.J. *Diagnosis and Management of Lameness in the Horse*, 2nd ed.; Saunders: St. Louis, MO, USA, 2011.
41. Buchner, H.H.; Savelberg, H.H.; Schamhardt, H.C.; Barneveld, A. Head and trunk movement adaptations in horses with experimentally induced fore- or hindlimb lameness. *Equine Vet. J.* **1996**, *28*, 71–76. [[CrossRef](#)]
42. Kramer, J.; Keegan, K.G.; Wilson, D.A.; Smith, B.K.; Wilson, D.J. Kinematics of the hind limb in trotting horses after induced lameness of the distal intertarsal and tarsometatarsal joints and intra-articular administration of anesthetic. *Am. J. Vet. Res.* **2000**, *61*, 1031–1036. [[CrossRef](#)]
43. Buchner, H.H.; Savelberg, H.H.; Schamhardt, H.C.; Barneveld, A. Limb movement adaptations in horses with experimentally induced fore- or hindlimb lameness. *Equine Vet. J.* **1996**, *28*, 63–70. [[CrossRef](#)]
44. May, S.A.; Wyn-Jones, G. Identification of hindleg lameness. *Equine Vet. J.* **1987**, *19*, 185–188. [[CrossRef](#)] [[PubMed](#)]
45. Church, E.E.; Walker, A.M.; Wilson, A.M.; Pfau, T. Evaluation of discriminant analysis based on dorsoventral symmetry indices to quantify hindlimb lameness during over ground locomotion in the horse. *Equine Vet. J.* **2009**, *41*, 304–308. [[CrossRef](#)] [[PubMed](#)]
46. Altman, D.G. *Practical Statistics for Medical Research*; Chapman & Hall/CRC: London, UK, 1999; p. XII. 611p.
47. Buchner, H.H.; Savelberg, H.H.; Schamhardt, H.C.; Merckens, H.W.; Barneveld, A. Kinematics of treadmill versus overground locomotion in horses. *Vet. Q.* **1994**, *16* (Suppl. 2), S87–S90. [[CrossRef](#)]
48. Audigié, F.; Pourcelot, P.; Degueurce, C.; Geiger, D.; Denoix, J.M. Kinematic analysis of the symmetry of limb movements in lame trotting horses. *Equine Vet. J. Suppl.* **2001**, *33*, 128–134. [[CrossRef](#)]
49. Kramer, J.; Keegan, K.G.; Kelmer, G.; Wilson, D.A. Objective determination of pelvic movement during hind limb lameness by use of a signal decomposition method and pelvic height differences. *Am. J. Vet. Res.* **2004**, *65*, 741–747. [[CrossRef](#)] [[PubMed](#)]
50. Leelamankong, P.; Estrada, R.; Mählmann, K.; Rungsri, P.; Lischer, C. Agreement among equine veterinarians and between equine veterinarians and inertial sensor system during clinical examination of hindlimb lameness in horses. *Equine Vet. J.* **2020**, *52*, 326–331. [[CrossRef](#)]
51. Hardeman, A.M.; Serra Bragança, F.M.; Swagemakers, J.H.; van Weeren, P.R.; Roepstorff, L. Variation in gait parameters used for objective lameness assessment in sound horses at the trot on the straight line and the lunge. *Equine Vet. J.* **2019**, *51*, 831–839. [[CrossRef](#)] [[PubMed](#)]
52. Rhodin, M.; Pfau, T.; Roepstorff, L.; Egenvall, A. Effect of lungeing on head and pelvic movement asymmetry in horses with induced lameness. *Vet. J.* **2013**, *198* (Suppl. 1), e39–e45. [[CrossRef](#)]
53. Whay, H.R.; Main, D.C.; Green, L.E.; Webster, A.J. Assessment of the welfare of dairy cattle using animal-based measurements: Direct observations and investigation of farm records. *Vet. Rec.* **2003**, *153*, 197–202. [[CrossRef](#)]
54. Whay, H.R.; Shearer, J.K. The Impact of Lameness on Welfare of the Dairy Cow. *Vet. Clin. N. Am. Food Anim. Pract.* **2017**, *33*, 153–164. [[CrossRef](#)]
55. Haubro Andersen, P.; Bech Glerup, K.; Wathan, J. Can a Machine Learn to See Horse Pain?: An Interdisciplinary Approach Towards Automated Decoding of Facial Expressions of Pain in the Horse. *Animals* **2018**, *11*, 1643. [[CrossRef](#)]

IV. ERWEITERTE DISKUSSION

Für die Lahmheitsuntersuchung beim Pferd wird ein weitgehend standardisierter Ablauf verwendet (Davidson 2018, Baxter et al. 2020). Dieser soll eine möglichst objektive Beurteilung von Lahmheiten ermöglichen, was durch die Anwendung einer weitgehend einheitlichen Skala zur Klassifizierung von Lahmheiten unterstrichen wird (American Association of Equine Practitioners 2022). Den Grad der Lahmheit eines Pferdes auf einer Skala von 0-5 (siehe Tabelle 1) anzugeben, wird jedoch immer wieder als unzureichend bewertet, da eine eindeutige, auch für nicht anwesende Personen nachvollziehbare Einteilung mit nur fünf Ziffern kaum möglich ist (Ross 2011, Brehm et al. 2017, Baxter et al. 2020). Fuller et al. (2006) haben gezeigt, dass sich die Anwendung einer numerischen Klassifizierung von 0-5 für eine Einzelperson eignet, jedoch die Übertragung auf andere Personen fehlerbehaftet ist. Die im Vereinigten Königreich angewandte Einteilung, Lahmheit auf einer Skala von 1-10 anzugeben, ermöglicht eine genauere Differenzierung des Lahmheitsgrades in Zahlen. Es werden jedoch in der Textbeschreibung jeweils zwei Ziffern zusammengefasst (siehe Tabelle 1), was eine genaue Klassifizierung erschwert (Wyn-Jones 1988, Baxter et al. 2020).

Bereits May und Wyn-Jones (1987) konnten darlegen, dass es einen Zusammenhang zwischen der Berufserfahrung eines Tierarztes oder einer Tierärztin und der Fähigkeit, Lahmheiten zu erkennen, gibt. Studierende und junge Absolventinnen und Absolventen haben häufiger Probleme, die lahme Gliedmaße zu benennen (Barstow et al. 2014, Starke und May 2017). Dennoch sollte beachtet werden, dass dem menschlichen Auge auch mit Training Grenzen in der Wahrnehmung gesetzt sind, wenn es um geringgradige Änderungen in der Beschleunigung beziehungsweise Bewegung eines Objektes geht (Haarmeier und Thier 2006, Parkes et al. 2009). Auch unter erfahrenen Tierärzten und Tierärztinnen gibt es vor allem bei geringgradigen Lahmheiten Diskrepanzen bezüglich der Identifikation des betroffenen Beines (Fuller et al. 2006, Keegan 2019). Ebenso hat sich gezeigt, dass die Voreingenommenheit eines einzelnen Tierarztes oder einer Tierärztin die Auswertung der Leitungsanästhesien beeinflussen kann (Arkell et al. 2006). Auch die als „positiv“ gewertete Anästhesie bei einer Verbesserung der Lahmheit um 70-80% lässt einen gewissen Interpretationsspielraum offen (Bassage und Ross 2011, Brehm et al. 2017).

Die Beurteilung von Lahmheiten mit der Hilfe von Pose Estimation kann für Tierärztinnen und Tierärzte eine Unterstützung darstellen. Zum einen ist es als Trainingsoption verwendbar, um Lahmheiten schneller erkennen zu können und auch bei geringgradigen Lahmheiten mehr Sicherheit zu bekommen. Zum anderen bietet es in den ersten Berufsjahren die Möglichkeit, im Rahmen der Lahmheitsuntersuchung eine „Zweitmeinung“ zu erhalten und die eigenen erhobenen Befunde zu verifizieren. Außerdem bieten Videoaufzeichnungen der Lahmheit eines Pferdes und deren Vergleich im Verlauf eine gute Dokumentationsmöglichkeit auch über einen längeren Zeitraum (Fuller et al. 2006). Die Verwendung eines Smartphones, um Lahmheiten ohne kostspielige Ausrüstung vor Ort beurteilen zu können, bietet einige Vorteile. Die Videoanalyse ist berührungslos und nicht invasiv und die Videos können an jedem beliebigen Ort aufgenommen werden. Wie Buchner et al. (1994) zeigen konnten, verändern sich Bewegung und Gangart eines Pferdes, wenn beispielsweise ein Laufband benutzt wird. Mit der Anwendung von Pose Estimation kann das Pferd im eigenen Stall in gewohnter Trainingseinrichtung und Bodenbeschaffenheit gefilmt werden. Außerdem lassen sich die Videos leicht über das Internet übertragen und vereinfachen den Austausch unter Tierärztkollegen und -kolleginnen.

In der Humanmedizin konnte in mehreren Fachbereichen ein erfolgreicher Einsatz von Deep Learning nachgewiesen werden (Gulshan et al. 2016, Esteva et al. 2017, Yala et al. 2019). Die Möglichkeit, den Algorithmus anhand eines großen Datensatzes zu trainieren, ist durch die Routineuntersuchungen, vor allem in der Bildgebung, sehr gut machbar. Ein Vorteil hierbei ist das Vorhandensein einer großen Anzahl gemeinsamer Analyse Kriterien in Kombination mit hoher Variabilität in ihrer Ausprägung (Yala et al. 2019). Die Relevanz eines großen, möglichst vielfältigen Datensatzes, um die Sicherheit des Algorithmus zu verbessern und genaue Ergebnisse in der Analyse zu erhalten, hat sich auch in der vorliegenden Studie gezeigt. In der Veterinärmedizin werden KI-gestützte Programmen vor allem in der Forschung sowie in der Bildgebung oder Histopathologie eingesetzt (Mathis et al. 2018, Bertram et al. 2019, Aubreville et al. 2022). Das Erfassen, Sammeln und Kategorisieren von Daten und CT- oder MRT-Bildern im Rahmen der (Routine-) Diagnostik kann helfen, größere Datensätze für das Training von Algorithmen auch in der Tiermedizin zu erhalten und so den Einsatz von KI voranzubringen. Kil et al. (2020) konnten zeigen, dass

Pose Estimation auch bei Pferden in einer Box angewandt werden kann. Die Verwendung einer einzelnen Kamera limitiert bei der Anwendung von OMC die vollständige Aufzeichnung des Objektes, weshalb Guerra Filho (2005) für die Bewegungsanalyse die Anwendung mehrerer Kameras empfiehlt. Der Einsatz diverser Kamerasysteme sowie das Aufzeichnen von Referenzpunkten am Kopf und im Gesicht des Pferdes würden eine tiefere Analyse des Verhaltens beispielsweise in Kombination mit der Horse Grimace Scale nach Dalla Costa et al. (2014) ermöglichen. Auch eine zusätzliche Anwendung der Horse Grimace Scale im Rahmen einer Lahmheitsuntersuchung ist dabei denkbar. Pose Estimation zur Lahmheitsanalyse wurde auch in der Studie von Wang et al. (2021) untersucht, die Unterschiede zwischen lahmen und nicht lahmen Pferden graphisch darstellt. Es wurden jedoch keine Angaben zum Lahmheitsgrad in Relation zur graphischen Darstellung gemacht oder Daten zur Berechnung der Lahmheit angegeben (Wang et al. 2021). Ohne diese ist keine genaue Bewegungsanalyse möglich.

Mit der Anwendung von Pose Estimation besteht die Möglichkeit, auf eine nicht-invasive Art und Weise Bewegung zu erfassen, aufzuzeichnen und zu messen. Nach dem Trainieren der KI mit zuvor gelabelten Bildern, das heißt Bilder, in denen entsprechende anatomische Strukturen manuell markiert wurden, ist es für den Algorithmus möglich, Körperteile auf neuen, ihm unbekanntem Bildern zu erkennen und zu kennzeichnen. Um für die Anwendung von Pose Estimation die am besten geeignete Aufnahmeposition zu finden, wurden in der vorliegenden Studie für das Training des Programmes insgesamt 215 Videos ausgewertet. Es zeigte sich, dass das Filmen des Pferdes im Trab auf einer geraden Linie von vorne, von hinten oder von der Seite nicht genügend Schritte für eine Auswertung bietet. Videos, die aus dem Innenkreis (Longierposition) gefilmt wurden, ermöglichen die Analyse einer ausreichenden Anzahl gleichmäßiger Schritte. Im Rahmen der Lahmheitsuntersuchung sollten die Pferde sowohl auf der Geraden als auch auf einer Kreislinie untersucht werden (Ross 2011, Baxter et al. 2020). In der Bewegung des Rumpfes und des Beckenbereichs zeigen sich Unterschiede, wenn ein Pferd auf einer geraden Linie und auf einer Kreisbahn evaluiert wird (Rhodin et al. 2013, Baxter et al. 2020). Jung (2021) konnte zeigen, dass sich die Kreisbahn sowohl für eine subjektive als auch für eine objektive Evaluation von Lahmheiten mittels akzelerometrisch erfasster Messdaten eignet. Mit der Weiterentwicklung des Programmes soll deshalb auch eine Analyse von kürzeren Videosequenzen auf

geraden Linien ermöglicht werden.

Im Rahmen der Analyse hat sich gezeigt, dass unregelmäßige Bewegungen (Kopfschütteln, Wiehern, plötzliches Erschrecken) oder andere Pferde, die sehr dicht am gefilmten Tier vorbeilaufen und so kurzzeitig auf dem Video zu sehen sind, die korrekte Positionierung der Referenzpunkte durch das Programm erschweren. Dies hat in der späteren Analyse zu keiner Beeinträchtigung geführt, da aufgrund der Kreisvideos ausreichend Daten zur Bewertung des Gangbildes zur Verfügung standen. Für eine spezifischere Analyse wurden diese Messfehler minimiert, indem 5% der Maximalwerte (95-100%) und 5% der Minimalwerte (0-5%) ausgeschlossen wurden. Die Maxima repräsentieren die größten gemessenen Werte (90-95%), die Minima die kleinsten gemessenen Werte (5-10%) für den einzelnen Referenzpunkt. Dennoch ist eine weitere Verbesserung mit einer möglichst stabilen Pose Estimation Ziel in der weiteren Entwicklung der Software.

Um eine möglichst breite Anwendung des Programmes zu ermöglichen, wurden Pferde zum einen in vielen verschiedenen Reithallen mit unterschiedlich positionierten Fenstern und Lichtquellen gefilmt, zum anderen in diversen Außenbereichen und Reitplätzen mit unterschiedlichen Hintergründen (Bäume, Felder, Gras, Verkehr). Dieser Aspekt ist auch für das spätere Training des Programms wichtig, damit eine Beurteilung durch die KI nicht anhand von Charakteristika der Umgebung vorgenommen wird. Eine möglichst große Vielfalt unterschiedlicher Hintergründe auf den gelabelten Bildern ist notwendig, um den Algorithmus optimal trainieren zu können (Nath et al. 2019). Eines der Ziele war es somit, die Software so zu trainieren, dass sie lernt, Körperteile auch bei unübersichtlichem und wechselndem Hintergrund, inhomogener Beleuchtung oder instabiler Kameraführung möglichst genau zu bestimmen. Im Rahmen der Studie hat sich gezeigt, dass Aufnahmen im Abendlicht oder bei starkem Sonnenschein das Filmen und die spätere Analyse erschwerten und die ermittelten Daten ungenauer wurden. Dieses Problem wurde bereits bei Mathis et al. (2018) diskutiert.

Eine weitere Einschränkung auf einigen Reitanlagen war die Qualität des Bodens, auf dem die Pferde getrabt wurden. Bei sehr tiefem Sand oder insgesamt sehr unebenem Boden wurde ein vermehrtes Stolpern der Pferde festgestellt. Außerdem waren sie in ihren Bewegungen sehr ungleichmäßig, was dann entweder einer Lahmheit ähneln konnte oder auch eine sehr geringgradige Lahmheit verdeckt hat.

Dies ist jedoch nicht ausschließlich eine Beeinträchtigung bei der Anwendung dieses Programmes, sondern erschwert Lahmheitsuntersuchungen im Allgemeinen. Dieser Faktor muss bei der zukünftigen Nutzung des Tools berücksichtigt werden, da das Programm auch dazu dienen soll, Videos zu analysieren, die von BesitzerInnen oder unerfahrenen Tierärzte und Tierärztinnen aufgenommen wurden und diese sich auf die Genauigkeit der Analyse verlassen. Es ist damit zu rechnen, dass ein weiteres Training mit mehr Daten die Genauigkeit des Programmes und damit auch dessen Zuverlässigkeit erhöht. Um das Risiko der Überanpassung zu reduzieren, ist hierfür das Einführen neuer gelabelter Daten notwendig.

Neben den Herausforderungen bei den Aufnahmen der Videos war auch das Setzen der Marker teilweise mit Schwierigkeiten verbunden. Ein Faktor ist die Farbe des Fells und/oder der Hufe, die bei einigen Pferden dem Hintergrund, dem Sand oder dem Boden farblich sehr ähnlich waren. Wenn diese Farben für das Programm nicht unterscheidbar und somit die anatomischen Referenzpunkte für den Algorithmus nicht eindeutig zu erkennen waren, wurden Marker ungenau gesetzt und entsprechende Videosequenzen konnten nicht verwendet werden. Bei komplett schwarzen oder weißen Pferden sowie bei einer sehr gleichmäßigen Fellfarbe waren anatomische Strukturen schlecht zu erkennen. Dies hat sich auch bei der Analyse des Kniegelenkes als Referenzpunkt gezeigt (Referenzpunkt Knie, Pferd 1 mit Lahmheit), als ein dunkelbraunes Warmblut ohne Farbnuancen im Fell evaluiert wurde. Die Situation wurde teilweise noch zusätzlich durch den Einfall von direktem Sonnenlicht erschwert, so dass eine Kennzeichnung sehr anspruchsvoll war und zu ungenaueren Analyseergebnissen führte (Referenzpunkt Knie, Pferd 7 Kontrollgruppe). In einigen Ausnahmefällen (insbesondere bei sehr hellem Licht) war ein Auswerten der Videos nicht möglich, so dass diese Pferde von der Studie ausgeschlossen werden mussten. Sowohl für das Training des Programms als auch in der nachfolgenden Analyse wurde grundsätzlich auf ein breites Farbspektrum in Bezug auf die Fellfarbe geachtet. Aufgrund der Ergebnisse im Rahmen der Studie wird dies auch im weiteren Training der Software von übergeordneter Bedeutung sein. Bei der Auswertung der Filmaufnahmen konnte bei adipösen Pferden oder bei Pferden mit sehr langem Winterfell eine erhöhte Fehlerquote beim Setzen der Marker erfasst werden. Als Ursache dafür zeigt sich auch hier das erschwerte Erkennen der anatomischen Strukturen durch das Programm. Durch den oben

erwähnten Ausschluss von 5% der erfassten Maximal- und Minimalwerte konnten diese Messfehler aus den Daten entfernt werden.

Für die Analyse der Vorhand-Lahmheiten mittels Pose Estimation wurden Referenzpunkte an den Vordergliedmaßen und am Kopf markiert. Am Kopf wurde der Punkt 4 (Genick) und an den Beinen die Punkte 17 (Ellbogengelenk links) und 21 (Karpus links) für den linken Zirkel sowie 18 (Ellbogengelenk rechts) und 22 (Karpus rechts) für den rechten Zirkel gewählt. Um die Standbeinphase bestimmen zu können, wurden die Punkte 19 (*Os carpi accessorium* links) und 20 (*Os carpi accessorium* rechts) sowie 45 (Tarsus links) und 46 (Tarsus rechts) markiert. Anhand der Ergebnisse und der graphischen Darstellung konnte eine Lahmheit eindeutig nachgewiesen werden. Die Kurven zeigten das Kopfnicken als Folge einer stärkeren Gewichtsbelastung der nicht lahmen Gliedmaße, wenn der Huf auf den Boden gesetzt wurde. Im Gegensatz dazu zeigten die Pferde der gesunden Kontrollgruppe keine asymmetrische, sich wiederholende Nickbewegung bei Belastung der rechten beziehungsweise linken Vordergliedmaße. Somit konnte gezeigt werden, dass durch die Betrachtung der graphischen Darstellung eine Unterscheidung zwischen einem lahmen und einem gesunden Pferd mittels Pose Estimation möglich ist. Als nächster Schritt ist eine Weiterentwicklung des Programms angestrebt, um die erfassten Parameter der Kopf- und Gliedmaßenbewegung in Bezug auf die Beinphase bestimmen zu können. Damit wird eine Berechnung der Messwerte und somit die Erfassung spezifischerer Daten möglich sein. Durch das Trainieren des Programmes mit weiteren Daten wird erwartet, dass eine stabile Anwendung von Referenzpunkten an der distalen Gliedmaße auch das Messen der Schrittlänge ermöglicht.

Der Einsatz von Pose Estimation für die Analyse von Hinterhand-Lahmheiten hat die Anwendung verschiedener anatomische Orientierungspunkte an den Hintergliedmaßen als optionale Referenzpunkte ermöglicht. Im Rahmen der Vorbereitung der Studie erwies sich zum einen das *Tuber coxae* als vielversprechend, da es bereits in verschiedenen Studien zur Lahmheitsuntersuchung verwendet wurde (May und Wyn-Jones 1987, Buchner et al. 1996, Church et al. 2009). Das Kniegelenk als Referenzpunkt ist beim Pferd bisher nicht mit transportablen Systemen untersucht worden, da die Fixierung eines Beschleunigungssensors an diesem Punkt nicht möglich ist. Nur in Studien mit Optical Motion Capture wurde er nach Kenntnis der Autorin als Referenzpunkt

verwendet (Kramer et al. 2000, Audigié et al. 2001). In beiden Studien wurden lahme und nicht lahme Pferde im Seitenvergleich gefilmt und dabei die Bewegung mehrerer Referenzpunkte in der Vertikalen und Gelenkwinkel im Verhältnis zur Zeit verglichen (Kramer et al. 2000, Audigié et al. 2001) sowie die Länge der Vorführphase und die Hufhöhe in der Bewegung (Kramer et al. 2000).

Nach Abschluss der Analyse konnten bei acht von neun Pferden eine Korrelation zwischen Lahmheitsgrad und der Differenz aus den berechneten Werten, die aus der Bewegung des linken und rechten Kniegelenks in der Horizontalen ermittelt worden waren, nachgewiesen werden. Bei dem Pferd mit fehlender Übereinstimmung der ermittelten Werte und des Lahmheitsgrades handelte es sich um ein dunkelbraunes Warmblut mit sehr gleichmäßig gefärbtem Fell, was wie oben erläutert, zum jetzigen Zeitpunkt noch zu Ungenauigkeiten in der Analyse der Software führen kann. Die Kontrollgruppe ohne Lahmheit enthielt ein Pferd, dessen Werte bei Vergleich des linken und rechten Zirkels eine Differenz aufwiesen, die bei den übrigen Pferden der Kontrollgruppe nicht vorhanden war. Diese Filmaufnahmen wurden abends auf einem Reitplatz bei untergehender Sonne angefertigt, wobei ein Teil des Platzes noch von Sonne beschienen wurde. Wie oben aufgeführt kann die Qualität des Videos durch Sonnenstrahlen beeinträchtigt werden, da diese eine weiß schimmernde Reflexion im Video verursachen. Das Erkennen der anatomischen Strukturen wird für den Algorithmus erschwert und führt zu Ungenauigkeiten in den Messergebnissen.

Das *Tuber coxae* als Referenzpunkt hingegen hat sich für die Bewegungsanalyse in Kreisvideos als nicht geeignet erwiesen. Der Vergleich der Medianwerte der als lahm erkannten, den nicht erkannten lahmen und den gesunden Pferden ermöglichte keine Klassifizierung in Bezug auf das Vorhandensein einer Lahmheit oder eine Lahmfreiheit. Kramer et al. (2000), Kramer et al. (2004) und Wyn-Jones (1988) haben gezeigt, dass ein zeitgleicher, paralleler Vergleich der Bewegung des linken und rechten *Tuber coxae* für das Erkennen einer Lahmheit stattfinden sollte, um eine Asymmetrie zu erkennen. Bei der Verwendung von Zirkelvideos kann nur eine Seite der Kruppe analysiert werden und ein direkter Vergleich mit der kontralateralen Seite ist nicht möglich.

Eine Sensitivität und Spezifität von fast 90 % bei der Verwendung des Kniegelenks als Referenzpunkt liefern vielversprechende Ergebnisse im Rahmen dieser ersten Untersuchungen. Das Trainieren des Programms mit mehr gelabelten Daten wird

zu einer stabileren Platzierung der Marker führen, um so auch bei ungünstigen Lichtverhältnissen und bei Pferden mit schlecht erkennbaren anatomischen Orientierungspunkten genaue Analyseergebnisse zu erhalten.

Bei der Verwendung einer Kraftmessplatte kann der direkte Gewichtsunterschied in der Belastung der einzelnen Beine gemessen werden (Byström et al. 2018). Sie liefert sehr genaue Messergebnisse und gilt immer noch als Goldstandard unter den Systemen zu Lahmheitsanalyse (Keegan et al. 2011, Keegan 2012, Serra Bragança et al. 2018). Unter Einbeziehung der Tatsachen, dass entweder nur einzelne Tritte aufgezeichnet werden können, was eine direkte Vergleichbarkeit mit der kontralateralen Gliedmaße nicht möglich macht oder dass die Kombination mit einem Laufband notwendig ist, um alle vier Beine gleichzeitig analysieren zu können (Serra Bragança et al. 2018) zeigt, dass dieses System für den alltäglichen Gebrauch ungeeignet ist (Keegan et al. 2011, Keegan 2012, Serra Bragança et al. 2018) und die Anwendung auf größere Kliniken und Universitäten beschränkt bleibt (Keegan 2007, Keegan 2012). Auch der Aspekt der fachlichen Expertise für die Auswertung und der zeitliche Aufwand, der durch das Gewöhnen der Pferde an das Laufband bedingt ist, sollten bedacht werden (Keegan 2007, Keegan et al. 2011, Back und Clayton 2013). Die Druckmessplatte ist mit weniger genauen Messdaten und der Tatsache, dass sie vor jeder Verwendung kalibriert werden muss, keine vollwertige Alternative (van Heel et al. 2004, Oosterlinck et al. 2010).

Mit der Anwendung von Optical Motion Capture ist grundsätzlich ein beliebiges Setzen der Marker möglich und es können große Mengen an Daten generiert werden (Serra Bragança et al. 2018). Dafür muss jedoch gewährleistet sein, dass ein Aufzeichnen der Marker aus verschiedenen Kamerawinkeln möglich ist (Guerra Filho 2005). Um eine ausreichend hohe Anzahl an auswertbaren Tritten zu erhalten, muss das Pferd entweder auf einem Laufband (Keegan 2012) oder in einem Raum mit fest installierten und geeichten Kameras gefilmt werden (Guerra Filho 2005). Dennoch ist es nicht immer möglich, ein Verdecken der Marker zu verhindern, was zu weniger verwertbaren Daten oder einem vollständigen Datenverlust führt (El-Gohary und McNames 2012, Kong et al. 2013). Aufgrund der Kosten, die durch die Ausstattung in Kombination mit Anpassen der Gegebenheiten vor Ort entstehen, ist auch dieses System vor allem für Forschungszentren und größere Kliniken geeignet (Keegan 2012, Bosch et al. 2018).

Bei der Anwendung von IMUs können je nach Lokalisation der Sensoren

verschiedene Parameter erhoben werden. Systeme, deren Beschleunigungssensoren ausschließlich am Kopf und Rumpf befestigt werden, können Minimal- und Maximalwerte der Vertikalbewegung im Seitenvergleich und in Relation zur Schrittfolge erfassen. Außerdem können die Auf- und Abwärtsamplituden der Messpunkte miteinander verglichen werden (Keegan et al. 2011, Leelamankong et al. 2020). Das Equimoves©-System bietet die Möglichkeit, mittels der Sensoren an den Beinen zusätzliche Parameter wie Schrittlänge oder bestimmte Winkel zu messen und so ein noch genaueres Gangbild des Pferdes zu erfassen (Bosch et al. 2018). In einer Studie von Weishaupt et al. (2001) mit zwei Equimetrix-Systemen, die an *Sternum* und *Sacrum* mit Bauchgurten befestigt waren, wurden gegenüber des zweiten Messsystems (Kraftmessplatte in Kombination mit Laufband) sowie drei erfahrenen Tierärztinnen und Tierärzten nur geringe Übereinstimmungen bei der Bewertung der Lahmheiten festgestellt. Die Autoren vermuten, dass die Kompensationsmechanismen des Pferdes, die zu einer Verschiebung der Bodenreaktionskraft innerhalb der vier Gliedmaßen führen und Auswirkungen auf die Beschleunigung einzelner Körperteile haben, sehr komplex sind und abhängig von der Lahmheitsursache variieren können, was eine Erfassung und Einordnung mit nur zwei Beschleunigungssensoren erschwert.

Die Studie von Shrestha et al. (2017) zeigt, dass die Anwendung von Radarsensoren grundsätzlich für die Lahmheitsanalyse bei Pferden geeignet ist, jedoch noch einige Probleme mit sich bringt. Zum einen sind bisher kaum Studien über die Bewegungsanalyse durch Radarsensoren bekannt, zum anderen ist die Analyse von aufrecht laufenden Menschen einfacher als von Tieren, die sich auf vier Beinen bewegen. Die geringe Anzahl an Studienteilnehmern (n=2) und die Tatsache, dass die Autoren und Autorinnen die schlechte Reproduzierbarkeit aufgrund nicht kontrollierbarer Bewegungen der Tiere als Limitation in der Studie angeführt haben (Shrestha et al. 2017), lassen auf weitere Untersuchungen zu diesem Forschungsansatz hoffen. Der Einsatz von Radarsensoren unabhängig von der Wettersituation und Distanz ist ein interessanter Aspekt für die Bewegungsanalyse bei Pferden (Shrestha et al. 2017).

Abhängig von der Wahl der Referenzpunkte konnte in der vorliegenden Studie gezeigt werden, dass anhand eines Systems, welches auf Basis künstlicher Intelligenz arbeitet, eine Lahmheits-Klassifizierung von Pferden anhand des Gangbildes möglich ist und eine hohe Übereinstimmung im Vergleich zur

klinischen Bewertung vorhanden ist. Diese Ergebnisse verdeutlichen das große Potenzial des Analyseansatzes. Durch die Anwendung von Pose Estimation können Einschränkungen, die bei anderen, modernen Systemen für Lahmheits- und Bewegungsanalyse vorhanden sind, zum Teil umgangen werden.

Zusammenfassend lässt sich sagen, dass noch einige Einschränkungen bei der Verwendung dieser Software auf einem Smartphone-Gerät vorhanden sind und daher das Programm einer weiteren Verbesserung bedarf. Vor allem die Beeinträchtigung der Videoqualität durch die bereits genannten Faktoren macht eine weitere Entwicklung der Software notwendig. Im Gegensatz dazu ermöglichten Markierungen am "Rand" des Pferdes sowie an gut sichtbaren anatomischen Strukturen wie Nüstern, Augen oder am Kronsaum eine exakte und reproduzierbare Erfassung von Daten zur Bewegung eines Pferdes.

Insgesamt zeigten die Probleme im Rahmen der Aufnahme und der nachfolgenden Analyse der Videos, dass die Vielfalt an gelabelten Bildern für das Training des Algorithmus noch erhöht werden muss. Die Tatsache, dass die Fehlerquote bei den Analysedaten im Vergleich zu den Trainingsdaten deutlich höher war, zeigt die noch vorhandene „Unsicherheit“ im Falle eines unbekanntes Videos. Bei einer Auflösung von 432 Pixeln in der vertikalen Achse kann dieser Fehler einen Unterschied von bis zu ~1,9% ausmachen. Folglich können die Referenzpunkte in einigen Sequenzen nicht richtig erkannt werden, was zu einem höheren Prozentsatz an ungenauen Platzierungen führt.

Im Verlauf des Trainings des Algorithmus wurde der Fehlerwert des Programmes immer wieder überprüft. Das Ausschließen der Referenzpunkte, die mit einer Sicherheit (engl. likelihood) $<60\%$ vom Algorithmus gesetzt wurden, führte in den Trainingsdaten nur zu einer Reduktion des Fehlerwertes von 2,6 Pixel auf 2,59 Pixel. Das zeigt, dass die Sicherheit bei den Trainingsdaten bereits sehr hoch ist und mit den bestehenden, gelabelten Daten auch durch weiteres Training nicht verbessert werden kann (Bengio et al. 1994, Glorot und Bengio 2010). Außerdem sollte zu diesem Zeitpunkt das Risiko der Überanpassung eines Netzes bedacht werden (Aghdam und Heravi 2017). Die hohe Unsicherheit des Netzes bei ungesehenen Daten kann durch eine größere Vielfalt im Datensatz der gelabelten Bilder verbessert werden. Mittels zusätzlicher Modifikation der Bilder, beispielsweise durch ein Erhöhen des Bildrauschens oder Änderungen von Farben, Helligkeit usw., könnte eine erhöhte Stabilität bei der Anwendung des Programmes

erreicht werden. Zusätzlich kann der Fehlerwert mit mehr gelabelten Daten und anderen Hyperparametern im weiteren Training des neuronalen Netzes zusätzlich reduziert werden.

Im Rahmen der Studie wurde die entwickelte Deep-Learning-Software zu Analyse von Lahmheiten anhand einer kleinen Studiengruppe untersucht. Aufgrund der geringen Stichprobengröße sind weitere Studien mit einem breiteren Patientenkollektiv erforderlich, um robuste Schätzungen für Sensitivität und Spezifität zu erhalten. Zusätzlich soll es mit weiterer Entwicklung des Programmes die Möglichkeit geben, den anatomischen Ursprung von Lahmheiten zu identifizieren. Hierfür ist jedoch eine ausreichend große Datengrundlage, die eine Einteilung der Videoaufnahmen nach gleicher Lahmheitsursache ermöglicht, notwendig.

Die in dieser Studie evaluierte Methodik, Pose Estimation für die Ganganalyse und die Lahmheitsdiagnostik in der Pferdemedizin einzusetzen, hat die Möglichkeit gezeigt, bei mittelgradig lahmen Pferden dem klinischen Bild entsprechende Analysedaten zu erhalten. Um Pferde mit geringgradiger Lahmheit sicher identifizieren zu können, ist weiteres Training der Software mit mehr Daten notwendig.

Das Software-Programm auf der Grundlage von Deep Learning bietet in der Veterinärmedizin vielfältige Möglichkeiten, den Arbeitsalltag von Tierärztinnen und Tierärzten zu erleichtern und die Diagnostik zu verbessern. Aus diesem Grund kommen diese neuen Techniken auch grundlegend dem Tierschutz zugute, da Untersuchungen standardisiert werden können und die Erkennung und Einordnung von Symptomen und Befunden entscheidend weiterentwickelt werden kann.

Erfahrene Tierärzte und Tierärztinnen können es als Unterstützung bei der täglichen klinischen Arbeit nutzen und die Dokumentation für eine retrospektive Bewertung der Behandlung nutzen. Mit der Auswertung von immer mehr Daten, die zum Trainieren des Programmes verwendet werden, können auch geringgradige Gangveränderungen erkannt und diagnostiziert werden. Ein weiterer möglicher Verwendungszweck kann die Ganganalyse bei Verdacht auf Ataxie oder auch die individuelle Bewertung des Trainings von Pferden sein.

V. ZUSAMMENFASSUNG

Lahmheiten bei Pferden sind ein seit langem bekanntes Problem, das sowohl das Wohlbefinden als auch die Nutzung von Pferden signifikant beeinflusst. Dennoch erfolgen die Erkennung und Klassifizierung von Lahmheiten hauptsächlich auf subjektiver Basis durch BesitzerInnen und den Tierarzt oder die Tierärztin. In der Pferdemedizin wurden in den letzten Jahren verschiedene Systeme zur Lahmheitsanalyse entwickelt. Optical Motion Capture (OMC) nutzt die Kombination aus Kamerasystemen und reflektierenden Markern, die am Pferd fixiert werden und eine Bewegungserfassung ermöglichen. Beschleunigungssensoren und Gyroskope sind Bestandteil von Inertial Measurement Unit (IMU)-Systemen, die am Pferd durch Klebeband oder Gurte befestigt werden und Bewegung direkt aufzeichnen können. Auf dem sich schnell entwickelnden Gebiet der künstlichen Intelligenz, des Deep Learning und der Smart-Device-Technologie wurde in der vorliegenden Studie eine Diagnosesoftware entwickelt, die helfen kann, zwischen lahmen und gesunden Pferden zu unterscheiden und die betroffene Gliedmaße zu lokalisieren. Deep Learning bildet einen Teilbereich des maschinellen Lernens, bei dem anhand von mehrschichtigen neuronalen Netzwerken ein Algorithmus trainiert wird. Dieser Algorithmus kann, nach entsprechendem Training, bestimmte Muster in hochdimensionalen Datensätzen erkennen und diese filtern. Die Technologie der Software basiert auf Pose Estimation, die bereits in vielen Bereichen der Human- und Veterinärmedizin eingesetzt wird, um die Bewegung von Gliedmaßen oder Körpern zu untersuchen, ohne dass Sensoren oder Marker am Objekt befestigt werden müssen. Anhand von Standbildern aus Videoaufnahmen, auf denen beliebig wählbare Referenzpunkte markiert wurden, wird ein Algorithmus auf die Erkennung dieser Punkte trainiert. Die Verwendung von 58 Referenzpunkten an gut erkennbaren anatomischen Orientierungspunkten bietet verschiedene Möglichkeiten der Ganganalyse in einer Messung. In dieser Studie wurde ein Software-Programm mit den Videos 65 verschiedener Pferde trainiert und anschließend dessen Analysefähigkeit anhand der erfassten Daten von 22 lahmen und acht nicht lahmen Pferden überprüft. Die erste Gruppe wurde für das Training des Netzes verwendet, anhand der zweiten und dritten Gruppe erfolgte die Evaluation lahmer sowie nicht lahmer Pferde. Die Ergebnisse zeigten, dass

Lahmheiten der Vordergliedmaßen durch die Visualisierung der Trajektorien der Referenzpunkte erkannt werden können. Bei Lahmheiten der Hintergliedmaßen lieferte das Kniegelenk als Referenzpunkt vielversprechende Ergebnisse, wohingegen sich der *T. coxae* als Referenzpunkt in dieser Konfiguration als ungeeignet erwies. Schwierige Lichteinstellungen wie direktes Sonnenlicht oder Dunkelheit, sowie eine dunkle, sehr gleichmäßige Fellfarbe (Rappen, Dunkelbraune) erschweren die genaue Platzierung von Referenzpunkten durch das Programm. Die Studie hat gezeigt, dass die Anwendung von Pose Estimation zur Erkennung von Lahmheiten möglich ist. Eine weitere Entwicklung mit einem größeren Datensatz ist unerlässlich, um die Präzision des Algorithmus weiter zu verbessern und vielfältige Analysen wie eine Klassifizierung nach Lahmheitsursache oder Gangbildanalysen zu ermöglichen.

VI. SUMMARY

Lameness in horses is a long-known problem that significantly affects both the well-being and the use of a horse. Yet, lameness detection and classification are mainly performed on a subjective basis by the owner and the veterinarian. Different systems for lameness analysis were developed in the last years in equine medicine. Optical Motion Capture (OMC) uses a combination of camera systems and reflective markers which are fixated on the horse and enable capturing the movement. Accelerometers and gyroscopes are part of Inertial Measurement Unit (IMU)-Systems, which are attached to the horse with tape or straps and directly record motion data. In the rapidly developing field of artificial intelligence, deep learning, and smart device technology, a diagnostic software was developed which can help distinguish between lame and sound horses and locate the affected limb. In the large field of machine learning, Deep Learning represents the training of an algorithm with a multi-layer neural network. After being trained, this algorithm is capable of detecting certain patterns in high-dimensional data. The technology is based on Pose Estimation, which is already used in human and veterinary medicine to study the movement of limbs or bodies without the need to attach devices or markers to the object. The use of 58 reference points at easily identifiable anatomical landmarks provides several options for gait analysis in one setup. Still frames of videos are extracted and variable points of interest labelled manually. Based on these still frames, an algorithm is trained on detecting these reference points afterwards. In this study, 22 horses with forelimb or hindlimb lameness and a control group of eight healthy sound horses were analysed with the program. The first group was used to train the network, followed by evaluation of the horses with and without lameness. The results showed that forelimb lameness could be detected by visualizing the trajectories of the reference points; for hindlimb lameness, the stifle joint showed promising results as a fiducial point, whereas the T. coxae proved unsuitable as a reference point in this setup. Difficult light settings such as direct sunlight or darkness or a dark, very uniform coat colour (black, dark brown) make accurate reference point placement difficult. In this study it was possible to demonstrate the use of pose estimation for lameness detection. Further development of the program with a larger data set is essential to improve the accuracy of the algorithm and to enable continuing studies on classification by the origin of

lameness or gait analyses.

VII. LITERATURVERZEICHNIS

(1998). USDA. Part I: Baseline Reference of 1998 Equine Health and Management. USDA:APHIS:VS, National Animal Health Monitoring System. Fort Collins (CO): : N280.898.

Aghdam, H. H. und Heravi, E. J. (2017). Convolutional Neural Networks. Guide to Convolutional Neural Networks: A Practical Application to Traffic-Sign Detection and Classification. Cham, Springer International Publishing: 85-130.

Aghdam, H. H. und Heravi, E. J. (2017). Pattern Classification. Guide to Convolutional Neural Networks: A Practical Application to Traffic-Sign Detection and Classification. Cham, Springer International Publishing: 15-83.

American Association of Equine Practitioners, AAEP. "Lameness Scale". <https://aaep.org/horsehealth/lameness-exams-evaluating-lame-horse>. zuletzt besucht: 26.September 2022

Andriluka, M. , Pishchulin, L. und Gehler, P. (2014). 2D Human Pose Estimation: New Benchmark and State of the Art Analysis, IEEE: 3686-3693

Arkell, M. , Archer, R. M. , Guitian, F. J. und May, S. A. (2006). "Evidence of bias affecting the interpretation of the results of local anaesthetic nerve blocks when assessing lameness in horses." Vet Rec 159(11): 346-349.

Aubreville, M. , Stathonikos, N. , Bertram, C. A. , Klopffleisch, R. , ter Hoeve, N. D. , Ciompi, F. , Wilm, F. , Marzahl, C. , Donovan, T. A. , Maier, A. K. , Breen, J. , Ravikumar, N. , Chung, Y. , Park, J. , Nateghi, R. , Pourakpour, F. , Fick, R. , Hadj, S. B. , Jahanifar, M. , Rajpoot, N. M. , Dexl, J. , Wittenberg, T. , Kondo, S. , Lafarge, M. W. , Koelzer, V. H. , Liang, J. , Wang, Y. , Long, X. , Liu, J. , Razavi, S. , Khademi, A. , Yang, S. , Wang, X. und Veta, M. B., K. (2022). "Mitosis domain generalization in histopathology images - The MIDOG challenge." ArXiv abs/2204.03742.:1-19

Audigié, F. , Pourcelot, P. , Degueurce, C. , Geiger, D. und Denoix, J. M. (2001). "Kinematic analysis of the symmetry of limb movements in lame trotting horses." *Equine Vet J Suppl*(33): 128-134.

Back, W. und Clayton, H. M. (2013). 1. History. *Equine Locomotion*. P. R. van Weeren. Edinburgh Scotland ; New York, Saunders Elsevier: 1-30.

Back, W. und Clayton, H. M. (2013). 2. Measurement Techniques for Gait Analysis. *Equine Locomotion*. H. M. Clayton and H. C. Schamhardt. Edinburgh Scotland ; New York USA, Saunders Elsevier: 31-60.

Back, W. und Clayton, H. M. (2013). 9. Gait Adaption in Lameness. *Equine Locomotion*. H. H. Buchner. Edinburgh Scotland ; New York, Saunders Elsevier: 175-197.

Barrey, E. (1999). "Methods, applications and limitations of gait analysis in horses." *Vet J* 157(1): 7-22.

Barstow, A. , Pfau, T. , Bolt, D. M. , Smith, R. K. und Weller, R. (2014). "Design and validation of a computer-aided learning program to enhance students' ability to recognize lameness in the horse." *J Vet Med Educ* 41(1): 1-8.

Bartmann, C. P. und Gehlen, H. (2017). 1 Klinische Untersuchung. *Handbuch Pferdepraxis*. W. Brehm, H. Gehlen, B. Ohnesorge and A. Wehrend. Stuttgart, Enke: 16-24.

Bassage, L. H. und Ross, M. W. (2011). Chapter 10 - Diagnostic Analgesia. *Diagnosis and Management of Lameness in the Horse (Second Edition)*. M. W. Ross and S. J. Dyson. Saint Louis, W.B. Saunders: 100-135.

Baxter, G. M. , Stashak, T. S. und Keegan, K. G. (2020). 2 Examination for Lameness. *Adams and Stashak's Lameness in Horses*. Hoboken, NJ: 67-188.

Bengio, Y. , Simard, P. und Frasconi, P. (1994). "Learning long-term dependencies with gradient descent is difficult." *IEEE Trans Neural Netw* 5(2): 157-166.

Bertram, C. A. , Aubreville, M. , Marzahl, C. , Maier, A. und Klopffleisch, R. (2019). "A large-scale dataset for mitotic figure assessment on whole slide images of canine cutaneous mast cell tumor." *Scientific Data* 6(1): 274.

Biercher, A. , Meller, S. , Wendt, J. , Caspari, N. , Schmidt-Mosig, J. , De Decker, S. und Volk, H. A. (2021). "Using Deep Learning to Detect Spinal Cord Diseases on Thoracolumbar Magnetic Resonance Images of Dogs." *Frontiers in Veterinary Science* 8: 721167.

Bosch, S. , Serra Bragança, F. , Marin-Perianu, M. , Marin-Perianu, R. , van der Zwaag, B. J. , Voskamp, J. , Back, W. , van Weeren, R. und Havinga, P. (2018). "EquiMoves: A Wireless Networked Inertial Measurement System for Objective Examination of Horse Gait." *Sensors (Basel)* 18(3): 850.

Brehm, W. , Burk, J. , Dellling, U. , Hagen, J. , Köhler M. , Litzke, L.-F. , Nowak, M. , Rijkenhuizen, A. , Schusser, G. F. , Tietje, S. und Troillet, A. (2017). *Handbuch Pferdepraxis. 21 - Krankheiten des Bewegungsapparates.* W. Brehm, H. Gehlen, B. Ohnesorge and A. Wehrend. Stuttgart, Enke: 849-1107.

Buchner, H. H. , Savelberg, H. H. , Schamhardt, H. C. und Barneveld, A. (1995). "Temporal stride patterns in horses with experimentally induced fore- or hindlimb lameness." *Equine Veterinary Journal* 27(S18): 161-165.

Buchner, H. H. , Savelberg, H. H. , Schamhardt, H. C. und Barneveld, A. (1996). "Head and trunk movement adaptations in horses with experimentally induced fore- or hindlimb lameness." *Equine Vet J* 28(1): 71-76.

Buchner, H. H. , Savelberg, H. H. , Schamhardt, H. C. , Merkens, H. W. und Barneveld, A. (1994). "Kinematics of treadmill versus overground locomotion in

horses." *Vet Q* 16 Suppl 2: S87-90.

Budras, K.-D. , Sack, W. O. und Röck, S. (2012). *Anatomy of the Horse*. Hannover, Schlütersche.: 112-184

Budsberg, S. , Verstraete, C. und Soutas-Little, R. W. (1987). "Force plate analysis of the walking gait in healthy dogs." *American journal of veterinary research* 48 6: 915-918.

Byström, A. , Egenvall, A. , Roepstorff, L. , Rhodin, M. , Bragança, F. S. , Hernlund, E. , van Weeren, R. , Weishaupt, M. A. und Clayton, H. M. (2018). "Biomechanical findings in horses showing asymmetrical vertical excursions of the withers at walk." *PLoS One* 13(9): e0204548.

Chen, V. C. , Tamoush, D. und Miceli, W. J. (2014). *Radar Micro-Doppler Signatures: Processing and Applications*, Institution of Engineering and Technology: 1-17.

Church, E. E. , Walker, A. M. , Wilson, A. M. und Pfau, T. (2009). "Evaluation of discriminant analysis based on dorsoventral symmetry indices to quantify hindlimb lameness during over ground locomotion in the horse." *Equine Vet J* 41(3): 304-308.

Dalla Costa, E. , Minero, M. , Lebelt, D. , Stucke, D. , Canali, E. und Leach, M. C. (2014). "Development of the Horse Grimace Scale (HGS) as a Pain Assessment Tool in Horses Undergoing Routine Castration." *PLOS ONE* 9(3): e92281.

Davidson, E. J. (2018). "Lameness Evaluation of the Athletic Horse." *Veterinary Clinics of North America: Equine Practice*: 181-191.

DeepLabCut. "DeepLabCut User Guide".
https://deeplabcut.github.io/DeepLabCut/docs/standardDeepLabCut_UserGuide.h

[tml](#). zuletzt besucht: 28.September 2022

Dyson, S. und Pollard, D. (2020). "Application of a Ridden Horse Pain Ethogram and Its Relationship with Gait in a Convenience Sample of 60 Riding Horses." *Animals (Basel)* 10(6): 1044.

Eichelberger, P. , Ferraro, M. , Minder, U. , Denton, T. , Blasimann, A. , Krause, F. und Baur, H. (2016). "Analysis of accuracy in optical motion capture - A protocol for laboratory setup evaluation." *J Biomech* 49(10): 2085-2088.

El-Gohary, M. und McNames, J. (2012). "Shoulder and Elbow Joint Angle Tracking With Inertial Sensors." *IEEE Transactions on Biomedical Engineering* 59(9): 2635-2641.

Esteva, A. , Kuprel, B. , Novoa, R. A. , Ko, J. , Swetter, S. M. , Blau, H. M. und Thrun, S. (2017). "Dermatologist-level classification of skin cancer with deep neural networks." *Nature* 542: 115.

Fuller, C. J. , Bladon, B. M. , Driver, A. J. und Barr, A. R. (2006). "The intra- and inter-assessor reliability of measurement of functional outcome by lameness scoring in horses." *Vet J* 171(2): 281-286.

Glorot, X. und Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. *AISTATS*: 249-256.

Guerra Filho, G. (2005). "Optical Motion Capture: Theory and Implementation." *RITA* 12: 61-90.

Gulshan, V. , Peng, L. und Coram, M. (2016). "Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs."

Haarmeier, T. und Thier, P. (2006). "Detection of speed changes during pursuit eye movements." *Exp Brain Res* 170(3): 345-357.

Hammarberg, M. , Egenvall, A. , Pfau, T. und Rhodin, M. (2016). "Rater agreement of visual lameness assessment in horses during lungeing." *Equine Veterinary Journal* 48(1): 78-82.

Hardeman, A. M. , Serra Bragança, F. M. , Swagemakers, J. H. , van Weeren, P. R. und Roepstorff, L. (2019). "Variation in gait parameters used for objective lameness assessment in sound horses at the trot on the straight line and the lunge." *Equine Vet J* 51(6): 831-839.

Hardeman, A. M. , Van Weeren, P. R. , Serra Bragança, F. M. , Warmerdam, H. und Bok, H. G. J. (2022). "A first exploration of perceived pros and cons of quantitative gait analysis in equine clinical practice." *Equine Veterinary Education* 34(10): e438-e444.

He, K. , Zhang, X. , Ren, S. und Sun, J. (2016). "Deep Residual Learning for Image Recognition." *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*: 770-778.

Insafutdinov, E. , Andriluka, M. und Pishchulin, L. (2017). "ArtTrack: Articulated Multi-Person Tracking in the Wild." 6457-6465.

Insafutdinov, E. , Pishchulin, L. , Andres, B. , Andriluka, M. und Schiele, B. (2016). "DeeperCut: A Deeper, Stronger, and Faster Multi-person Pose Estimation Model." *ECCV*: 34-50.

Jung, A. (2021). Untersuchung der Bewegungsasymmetrie von neun Pferden im Trab vergleichend mittels subjektiver Beurteilung von Videos und berechneter Asymmetrie-Indices der akzelerometrischen Messungen auf einer zirkulären Bahn einer Führanlage, Veterinärmedizinischen Universität Wien.

Keegan, K. G. (2007). "Evidence-based lameness detection and quantification." *Vet Clin North Am Equine Pract* 23(2): 403-423.

Keegan, K. G. (2012). Objective measures of lameness evaluation. *American College of Veterinary Surgeons Symposium*: 127-131.

Keegan, K. G. (2019). "Reliability of equine visual lameness classification." *Vet Rec* 12;184(2): 60-62.

Keegan, K. G. , Kramer, J. , Yonezawa, Y. , Maki, H. , Pai, P. F. , Dent, E. V. , Kellerman, T. E. , Wilson, D. A. und Reed, S. K. (2011). "Assessment of repeatability of a wireless, inertial sensor-based lameness evaluation system for horses." *Am J Vet Res* 72(9): 1156-1163.

Keegan, K. G. , Yonezawa, Y. , Pai, P. F. , Wilson, D. A. und Kramer, J. (2004). "Evaluation of a sensor-based system of motion analysis for detection and quantification of forelimb and hind limb lameness in horses." *Am J Vet Res* 65(5): 665-670.

Keg, P. R. , van Weeren, P. R. , Back, W. und Barneveld, A. (1997). "Influence of the force applied and its period of application on the outcome of the flexion test of the distal forelimb of the horse." *Vet Rec* 141(18): 463-466.

Kelmer, G. , Keegan, K. G. , Kramer, J. , Wilson, D. A. , Pai, F. P. und Singh, P. (2005). "Computer-assisted kinematic evaluation of induced compensatory movements resembling lameness in horses trotting on a treadmill." *Am J Vet Res* 66(4): 646-655.

Kil, N. , Ertelt, K. und Auer, U. (2020). "Development and Validation of an Automated Video Tracking Model for Stabled Horses." *Animals (Basel)* 10(12): 2258.

Kong, W. , Sessa, S. , Cosentino, S. , Zecca, M. , Saito, K. , Wang, C. , Imtiaz, U. , Lin, Z. , Bartolomeo, L. , Ishii, H. , Ikai, T. und Takanishi, A. (2013). "Development of a real-time IMU-based motion capture system for gait rehabilitation." 2013 IEEE International Conference on Robotics and Biomimetics (ROBIO): 2100-2105.

Kramer, J. , Keegan, K. G. , Kelmer, G. und Wilson, D. A. (2004). "Objective determination of pelvic movement during hind limb lameness by use of a signal decomposition method and pelvic height differences." Am J Vet Res 65(6): 741-747.

Kramer, J. , Keegan, K. G. , Wilson, D. A. , Smith, B. K. und Wilson, D. J. (2000). "Kinematics of the hind limb in trotting horses after induced lameness of the distal intertarsal and tarsometatarsal joints and intra-articular administration of anesthetic." Am J Vet Res 61(9): 1031-1036.

LeCun, Y. , Bengio, Y. und Hinton, G. (2015). "Deep learning." Nature 521(7553): 436-444.

Lecun, Y. , Kavukcuoglu, K. und Farabet, C. (2010). Convolutional Networks and Applications in Vision: 253-256.

Leelamankong, P. , Estrada, R. , Mählmann, K. , Rungsri, P. und Lischer, C. (2020). "Agreement among equine veterinarians and between equine veterinarians and inertial sensor system during clinical examination of hindlimb lameness in horses." Equine Vet J 52(2): 326-331.

López-Sanromán, F. J. , Holmbak-Petersen, R. , Santiago, I. , Gómez de Segura, I. A. und Barrey, E. (2012). "Gait analysis using 3D accelerometry in horses sedated with xylazine." Vet J 193(1): 212-216.

Ludloff, A. (1998). Einführung. Praxiswissen Radar und Radarsignalverarbeitung. Wiesbaden, Vieweg+Teubner Verlag: 1-10.

Marunova, E. , Dod, L. , Witte, S. und Pfau, T. (2021). "Smartphone-Based Pelvic Movement Asymmetry Measures for Clinical Decision Making in Equine Lameness Assessment." *Animals (Basel)* 11(6): 1665.

Mathis, A. , Mamidanna, P. , Cury, K. M. , Abe, T. , Murthy, V. N. , Mathis, M. W. und Bethge, M. (2018). "DeepLabCut: markerless pose estimation of user-defined body parts with deep learning." *Nat Neurosci* 21(9): 1281-1289.

May, A. , Gesell-May, S. , Muller, T. und Ertel, W. (2022). "Artificial intelligence as a tool to aid in the differentiation of equine ophthalmic diseases with an emphasis on equine uveitis." *Equine Vet J* 54(5): 847-855.

May, S. A. und Wyn-Jones, G. (1987). "Identification of hindleg lameness." *Equine Vet J* 19(3): 185-188.

Morris, E., Seeherman, H. (1987). Redistribution of ground reaction forces in experimentally induced equine carpal lameness. *Exercise Physiology* 2: 553-563.

Muller-Quirin, J. , Dittmann, M. T. , Roepstorff, C. , Arpagaus, S. , Latif, S. N. und Weishaupt, M. A. (2020). "Riding Soundness-Comparison of Subjective With Objective Lameness Assessments of Owner-Sound Horses at Trot on a Treadmill." *J Equine Vet Sci* 95: 103314.

Nath, T. , Mathis, A. , Chen, A. C. , Patel, A. , Bethge, M. und Mathis, M. W. (2019). "Using DeepLabCut for 3D markerless pose estimation across species and behaviors." *Nat Protoc* 14(7): 2152-2176.

Oosterlinck, M. , Pille, F. , Huppes, T. , Gasthuys, F. und Back, W. (2010). "Comparison of pressure plate and force plate gait kinetics in sound Warmbloods at walk and trot." *Vet J* 186(3): 347-351.

Parkes, R. S. , Weller, R. , Groth, A. M. , May, S. und Pfau, T. (2009). "Evidence

of the development of 'domain-restricted' expertise in the recognition of asymmetric motion characteristics of hindlimb lameness in the horse." *Equine Vet J* 41(2): 112-117.

Pfau, T. , Noordwijk, K. , Sepulveda Caviedes, M. F. , Persson-Sjodin, E. , Barstow, A. , Forbes, B. und Rhodin, M. (2018). "Head, withers and pelvic movement asymmetry and their relative timing in trot in racing Thoroughbreds in training." *Equine Vet J* 50(1): 117-124.

Quellec, G. , Charrière, K. , Boudi, Y. , Cochener, B. und Lamard, M. (2017). "Deep image mining for diabetic retinopathy screening." *Medical Image Analysis* 39: 178-193.

Rantanen, N. W. , Jorgensen, J. S. und Genovese, R. L. (2011). Chapter 16 - Ultrasonographic Evaluation of the Equine Limb: Technique. *Diagnosis and Management of Lameness in the Horse (Second Edition)*. M. W. Ross and S. J. Dyson. Saint Louis, W.B. Saunders: 182-205.

Rhodin, M. , Persson-Sjodin, E. , Egenvall, A. , Serra Bragança, F. M. , Pfau, T. , Roepstorff, L. , Weishaupt, M. A. , Thomsen, M. H. , van Weeren, P. R. und Hernlund, E. (2018). "Vertical movement symmetry of the withers in horses with induced forelimb and hindlimb lameness at trot." *Equine Vet J* 50(6): 818-824.

Rhodin, M. , Pfau, T. , Roepstorff, L. und Egenvall, A. (2013). "Effect of lungeing on head and pelvic movement asymmetry in horses with induced lameness." *Vet J* 198 Suppl 1: e39-45.

Roepstorff, L. , Wiestner, T. , Weishaupt, M. A. und Egenvall, E. (2013). "Comparison of microgyro-based measurements of equine metatarsal/metacarpal bone to a high speed video locomotion analysis system during treadmill locomotion." *Vet J* 198 Suppl 1: e157-160.

Ross, M. W. (2011). Chapter 2 - Lameness in Horses: Basic Facts Before Starting. *Diagnosis and Management of Lameness in the Horse (Second Edition)*. M. W. Ross and S. J. Dyson. Saint Louis, W.B. Saunders: 3-8.

Ross, M. W. (2011). Chapter 3 - Anamnesis (History). *Diagnosis and Management of Lameness in the Horse (Second Edition)*. M. W. Ross and S. J. Dyson. Saint Louis, W.B. Saunders: 8-15.

Ross, M. W. (2011). Chapter 5 - Observation: Symmetry and Posture. *Diagnosis and Management of Lameness in the Horse (Second Edition)*. M. W. Ross and S. J. Dyson. Saint Louis, W.B. Saunders: 32-43.

Ross, M. W. (2011). Chapter 6 - Palpation. *Diagnosis and Management of Lameness in the Horse (Second Edition)*. M. W. Ross and S. J. Dyson. Saint Louis, W.B. Saunders: 43-63.

Ross, M. W. (2011). Chapter 7 - Movement. *Diagnosis and Management of Lameness in the Horse (Second Edition)*. M. W. Ross and S. J. Dyson. Saint Louis, W.B. Saunders: 64-80.

Ross, M. W. (2011). Chapter 8 - Manipulation. *Diagnosis and Management of Lameness in the Horse (Second Edition)*. M. W. Ross and S. J. Dyson. Saint Louis, W.B. Saunders: 80-88.

Seitzinger, A. H. (2000). A comparison of the economic costs of equine lameness, colic, and equine protozoal myeloencephalitis (EPM). *Proceedings of the 9th International Symposium on Veterinary Epidemiology and Economics, 2000*: 1-4.

Serra Bragança, F. M. , Rhodin, M. und van Weeren, P. R. (2018). "On the brink of daily clinical application of objective gait analysis: What evidence do we have so far from studies using an induced lameness model?" *Vet J* 234: 11-23.

Serra Bragança, F. M. , Rhodin, M. , Wiestner, T. , Hernlund, E. , Pfau, T. , van

Weeren, P. R. und Weishaupt, M. A. (2018). "Quantification of the effect of instrumentation error in objective gait assessment in the horse on hindlimb symmetry parameters." *Equine Vet J* 50(3): 370-376.

Sewak, M. , Karim, M. R. und Pujari, P. (2018). 2 Introduction to Convolutional Neural Networks. *Practical convolutional neural networks: implement advanced deep learning models using Python*, Packt Publishing Ltd: 30-45.

Shrestha, A. , Le Kernec, J. , Fioranelli, F. , Marshall, J. F. und Voute, L. (2017). "Gait analysis of horses for lameness detection with radar sensors." 1-6.

Slater, J. (2016). *National Equine Health Survey (NEHS) 2016*, Blue Cross for Pets2016.:1-15.

Starke, S. D. und May, S. A. (2017). "Veterinary student competence in equine lameness recognition and assessment: a mixed methods study." *Vet Rec* 181(7): 168.

The MathWorks, I. "MathWorks® - Accelerating the pace of engineering and science". <https://de.mathworks.com/help/matlab/ref/gradient.html>. zuletzt besucht: 30.September 2022

Titterton, D. und Weston, J. (2004). 4 Gyroscope Technology 1. *Strapdown Inertial Navigation Technology*, Institution of Engineering and Technology: 59-112.

van Heel, M. C. V. , Barneveld, A. , van Weeren, P. R. und Back, W. (2004). "Dynamic pressure measurements for the detailed study of hoof balance: the effect of trimming." *Equine Veterinary Journal* 36(8): 778-782.

van Weeren, P. R. , Pfau, T. , Rhodin, M. , Roepstorff, L. , Serra Bragança, F. und Weishaupt, M. A. (2017). "Do we have to redefine lameness in the era of quantitative gait analysis?" *Equine Vet J* 49(5): 567-569.

van Weeren, P. R. , Pfau, T. , Rhodin, M. , Roepstorff, L. , Serra Bragança, F. und Weishaupt, M. A. (2018). "What is lameness and what (or who) is the gold standard to detect it?" *Equine Vet J* 50(5): 549-551.

Verschooten, F. und Verbeeck, J. (1997). "Flexion test of the metacarpophalangeal and interphalangeal joints and flexion angle of the metacarpophalangeal joint in sound horses." *Equine Vet J* 29(1): 50-54.

Wang, Y. , Li, J. , Zhang, Y. und Sinnott, R. (2021). Identifying lameness in horses through deep learning. *Proceedings of the 36th Annual ACM Symposium on Applied Computing*: 976-985.

Weishaupt, M. A. , Hogg, H. P. , Wiestner, T. , Denoth, J. , Stussi, E. und Auer, J. A. (2002). "Instrumented treadmill for measuring vertical ground reaction forces in horses." *Am J Vet Res* 63(4): 520-527.

Weishaupt, M. A. , Wiestner, T. , Hogg, H. P. , Jordan, P. , Auer, J. A. und Barrey, E. (2001). "Assessment of gait irregularities in the horse: eye vs. gait analysis." *Equine Vet J Suppl*(33): 135-140.

Wyn-Jones, G. (1988). Chapter 1 - The Diagnosis of the Causes of Lameness. *Equine Lameness*, Blackwell Scientific Publ.: 1-22.

Yala, A. , Lehman, C. , Schuster, T. , Portnoi, T. und Barzilay, R. (2019). "A Deep Learning Mammography-based Model for Improved Breast Cancer Risk Prediction." *Radiology* 292(1): 60-66.

Zeiler, M. , Krishnan, D. , Taylor, G. und Fergus, R. (2010). "Deconvolutional networks." 2528-2535.

VIII. ANHANG

Tabelle 2: Einteilung der Lahmheitsgrade (modifiziert nach (Wyn-Jones 1988, Baxter et al. 2020, American Association of Equine Practitioners 2022))

Einteilung nach AAEP		Einteilung UK	
Grad	Befunde	Grad	Befunde
0	Keine Lahmheit sichtbar.	0	Keine Lahmheit.
1	Die Lahmheit ist nur schwer erkennbar und unabhängig von den Gegebenheiten (z.B. unter dem Sattel, auf dem Zirkel, auf der Schrägen, auf hartem Boden) nicht kontinuierlich sichtbar.	1-2	Die Lahmheit ist im Schritt oder Trab nur schwer zu erkennen.
2	Die Lahmheit ist im Schritt und im Trab auf der Geraden schwer erkennbar, aber unter bestimmten Gegebenheiten (z.B. unter dem Sattel, auf dem Zirkel, auf der Schräge, auf hartem Boden) konstant sichtbar.	3-4	Die Lahmheit ist im Schritt nur schwer zu erkennen, im Trab gut sichtbar.
3	Die Lahmheit ist im Trab unter allen Gegebenheiten konstant sichtbar.	5-6	Die Lahmheit ist im Trab gut zu sehen.
4	Die Lahmheit ist im Schritt deutlich sichtbar.	7-8	Deutliche Lahmheit im Schritt, Traben ist kaum/nicht möglich.
5	Die Lahmheit führt zu einer minimalen Gewichtsbelastung in der Bewegung und/oder in Ruhe oder zu vollständiger Bewegungsunfähigkeit.	9-10	Das Pferd nimmt keine Last auf.

IX. DANKSAGUNG

Zuallererst gelten mein Dank meiner Doktormutter Prof. Dr. Anna May für die hervorragende Betreuung über die letzten zwei Jahre, die Unterstützung und Verwirklichung im Rahmen der Studie, bei der Veröffentlichung des Papers und der Dissertation durch unermüdliches Korrekturlesen zu diversen Tages- und Nachtzeiten, kritische Diskussionen und konstruktive Beiträge und dem Beantworten meiner zahlreichen Fragen. 1000 Dank für die Chance und die Hilfe, noch „im Herbst“ einreichen zu können!

...und Dr. Stefan Gesell-May für das Vertrauen, mir dieses Thema zu überlassen, die Unterstützung bei der Planung und Durchführung der Studie, für das kritische Erörtern immer neuer aufkommender Fragen und Diskussionspunkte, für die ständige Erreichbarkeit und die fachliche Anregung!

Tobias Müller möchte ich für das Beantworten all meiner Fragen zum Thema Künstliche Intelligenz, Deep Learning und Pose Estimation und das Korrekturlesen meiner Texte in Bezug auf fachliche Richtigkeit danken - und für die vielen hilfreichen Links!

Vielen Dank an Mathias Raths für Ideen und Hilfe bei der Erstellung der Excel-Tabellen zur Analyse der Daten.

Bei Prof. Dr. Marcus Doherr möchte ich mich sehr herzlich für die Anregungen und Unterstützung bei der Statistik und die vielen kritischen, sehr hilfreichen Anmerkungen zu meinen Texten bedanken - egal ob aus Kanada, Berlin oder Südafrika!

Ein großes Dankeschön gilt meinem Cousin James L. Zainaldin, dafür dass er über Nacht mein Paper korrekturgelesen und es mit fachlicher Expertise in Bezug auf Diktion und Kommasetzung im Englischen aufgewertet hat!

Für das Organisieren etlicher Pferde für die ersten Videoaufnahmen, viele Stunden in diversen Reitställen bei Schnee und Eis sowohl vor als auch hinter der Kamera geht ein großes Dankeschön an meine Freundin und ehemalige Kollegin Anna Mayet.

In Sachen Computer möchte ich Lena-Marie Neumann danken, die mir mit dem Formatieren einige schlaflose Stunden erspart hat und kurz vor Schluss dazu beigetragen hat, die Abgabe noch rechtzeitig zu schaffen.

Meinen Kollegen aus Parsdorf danke ich sehr für die Unterstützung bei der Aufnahme der Videos und am Ende für das Einspringen in der Klinik, damit ich die Dissertation abschließen konnte.

Meinen Eltern, ohne die ich es nie so weit geschafft hätte und die mich in jedem Abschnitt in den letzten zwei Jahren uneingeschränkt unterstützt haben, danke ich von ganzem Herzen.