

An den FH-Masterstudiengang Data Science und Engineering	
Name Studierende/r:	Christina Hager
PKZ:	2210595006

ANTRAG FÜR EIN THEMA FÜR EINE MASTERARBEIT

Ich ersuche um Genehmigung des folgenden Themas für meine Masterarbeit:

(Arbeits-)Titel:	
Landmarken-Erkennung anhand von Wundbildern einzelner Körperbereiche	
Das Exposé im Umfang von 1-2 Seiten ist diesem Antrag beigelegt.	
23.01.2024	<i>Hager Christina</i>
Datum	Unterschrift der/des Studierenden

VEREINBARUNG ÜBER DIE BETREUUNG UND BEGUTACHTUNG EINER MASTERARBEIT

Die Masterarbeit wird betreut und begutachtet von:

Titel Nachname Vorname:	FH-Prof. Dr. David Christian Schedl
-------------------------	-------------------------------------

Um eine zeitgerechte Begutachtung zu ermöglichen, ist die Masterarbeit bis zum 20. Juni (erster Masterprüfungstermin) bzw. bis zum 31. August (zweiter Masterprüfungstermin) abzugeben. Die/der Begutachter/in kann unabhängig davon einen früheren Termin für die Abgabe der Masterarbeit verlangen:

Datum	Unterschrift der FH-Betreuungsperson und Begutachter/in

KOOPERATIONSVEREINBARUNG MIT EINEM UNTERNEHMEN (wenn relevant)

Die Masterarbeit wird in Kooperation mit folgendem Unternehmen durchgeführt:

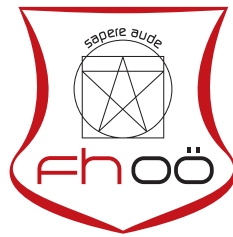
Firmenname:	RISC Software GmbH
Adresse:	Softwarepark 32a, 4232 Hagenberg im Mühlkreis

Seitens des Unternehmens wird die Arbeit betreut von:

Titel Nachname Vorname:	Sophie Kaltenleithner, MSc
Funktion im Unternehmen:	<i>Researcher & Developer</i>
Kontaktdaten:	<i>sophie.kaltenleithner@risc-software.at</i>
<i>23.01.2024</i>	<i>S. Kalt.</i>
Datum	Unterschrift der firmenseitigen Betreuungsperson

Landmarken-Erkennung anhand von Wundbildern einzelner Körperbereiche

Christina Hager, BSc



MASTERARBEIT (EXPOSÉ)

eingereicht am
Fachhochschul-Masterstudiengang

Data Science und Engineering

in Hagenberg

im Januar 2024

Betreuung (Vorschlag):
FH-Prof. Dr. David Christian Schedl

Inhaltsverzeichnis

Exposé	1
1 Kooperation mit RISC Software GmbH	1
2 Motivation und Ziel der Arbeit	1
2.1 Wissenschaftliche Fragestellung	2
3 Stand der Forschung	2
3.1 Monocular-Human-Pose-Estimation	2
3.2 Evaluierungsmethoden	3
3.3 Verfügbare Datensätze	4
Quellenverzeichnis	5
Literatur	5

Exposé

1 Kooperation mit RISC Software GmbH

Die Masterarbeit soll in Kooperation mit der RISC Software GmbH geschrieben werden. Dieses Unternehmen versteht sich als Forschungs- und Entwicklungsunternehmen und ist im Softwarepark Hagenberg ansässig. Entstanden ist die RISC Software GmbH aus dem RISC Institut der Johannes Kepler Universität Linz (JKU) und ist noch heute zu 80% im Besitz der JKU und zu 20% der Upper Austrian Research GmbH. [4]

Verankert ist diese Masterarbeit in der Medizin Informatik Abteilung des Unternehmens, die sich auf medizinische Modellierung und Simulation fokussiert. Dies umfasst auch den Forschungsbereich objektive Diagnosen und Dokumentation basierend auf dreidimensionalen, an Patient:innen angepassten, Modellen. [3]

2 Motivation und Ziel der Arbeit

Für das Therapieren von sowohl akuten als auch chronischen Wunden ist deren (Verlaufs-) Dokumentation von Bedeutung. Dabei werden beispielsweise Typ, Größe, Ursache oder auch optische Beschreibung der Wunde vermerkt. [13]

In der Vergangenheit wurde bei der RISC Software GmbH mit Burncase eine intuitive Hilfestellung zur Visualisierung und Dokumentation von Wunden und deren Heilungsverlauf geschaffen. Das Forschungsprojekt fokussierte sich dabei auf die Dokumentation und Flächenabschätzung von Brandwunden mithilfe eines anpassbaren dreidimensionalen Modells. Bei der Behandlung von Brandwunden ist das Verhältnis der Flächengrößen von verbrannter Hautoberflächen (TBSA-B) zur totalen Hautoberfläche (TBSA) für die weitere Behandlung bedeutend. Mit BurnCase konnte man zeigen, dass die Berechnung dieses Verhältnisses anhand eines 3D Körpermodells objektiver ist [5], als jene mit bekannten Schätzregeln bzw. -formeln wie Rule-of-Nines [14] oder Rule-of-Palms [11].

Auch diese Arbeit soll sich mit der Visualisierung und Dokumentation von Wunden anhand von virtuellen Modellen von Patient:innen beschäftigen. Aktuell werden Brandwunden entsprechend nach Verbrennungsgrad farbkodiert manuell auf das 3D Modell gezeichnet. Dies soll in Zukunft vereinfacht werden indem entsprechende Informationen zu einer Wunde aus einem RGB Bild extrahiert werden. Mit diesen Informationen soll dann ein semi-automatisches Mapping der Wunde auf ein Patient:innen Modell erfolgen können.

Dies würde jedoch den Umfang dieser Arbeit überschreiten. Deshalb soll vorerst eine umfassende Recherche zu gängigen Methoden im Bereich Pose-Estimation und Landmarken-Erkennung durchgeführt werden. Außerdem sind entsprechende Datensätze zu evaluieren, wobei diese gegebenenfalls zu editieren oder zu erweitern sind. Bilder von bestehenden Pose-Estimation Datensätzen unterscheiden sich von klassischen Wund-Bildern. Denn letztere zeigen oft nur Teilausschnitte von menschlichen Körpern wie Hände oder Füße und sind zum Teil auch stark deformiert. Ein mögliches Endergebnis dieser Arbeit könnte ein System sein, dass bereits automatisch von einem Bild ausgehend einen Wundbereich grob an einem deformierbaren 3D Körpermodell mit einer Bounding Box kennzeichnen kann.

2.1 Wissenschaftliche Fragestellung

Wie können bestehende Pose-Estimation Ansätze verbessert und evaluiert werden um sie im Bereich der Wunddokumentation einsetzbar zu machen? Dabei gilt es, aus RGB Fotos von größeren Körperbereichen wie Armen oder Beinen, Informationen zur Körperposition zu extrahieren. Diese sollen später als Anhaltspunkte dienen, um die im Bild gezeigten Bereiche auf einem deformierbaren, menschlichen 3D Modell automatisch zu markieren.

Es ist in der Arbeit sowohl die Eignung von bestehenden Ansätzen zu prüfen, als auch zu evaluieren, wie diese Ansätze dahingehend verbessert werden können, sodass auch von Bildern kleinerer Körperbereiche Informationen wie 2D Koordinaten von anatomischen Landmarken extrahiert werden können. Wird innerhalb der Arbeit ein gut funktionierendes System gefunden, so soll dies auch auf Körperregionen evaluiert werden, die Wunden bzw. deren Segmentierung zeigen.

3 Stand der Forschung

3.1 Monocular-Human-Pose-Estimation

Human Pose-Estimation ist ein fester Bestandteil von Computer Vision und versucht Informationen zur menschlichen Pose zu finden. Bei solchen Informationen kann es sich einerseits um die räumliche Position von anatomischen Landmarken wie Gelenken, oder aber auch die Form eines Körpers handeln. [2]

Dabei kann man grundlegend zwischen zwei Kategorien unterscheiden. Je nachdem ob 2D Koordinaten von anatomischen Landmarken allein oder dabei auch Tiefeninformationen vorhergesagt werden sollen, spricht man von 2D oder 3D Pose-Estimation. [2]

Körperposen können sehr vielfältig sein und weisen hohe Freiheitsgrade auf. Daher stößt die Pose-Estimation auf viele Problem wie zum Beispiel die Komplexität und Einzigartigkeit von Posen. Auch Kameraeinstellungen oder -perspektiven können die Qualität durch Bildunschärfe mitbestimmen. Solchen Problemen wollen aktuelle Ansätze mit Deep Learning Methoden begegnen. [9]

Architektur solcher Frameworks

Im Inneren weisen solche Netz-Architekturen Ähnlichkeiten auf und können in zwei Teile gegliedert werden. Einerseits wird ein Pose-Encoder benötigt, mit dem versucht wird, high-level Features zu extrahieren. Bei einem solchen Pose-Encoder handelt es sich meist um gängige Bildklassifikationsnetze, die bereits auf Datensätze wie Imagenet vortrainiert wurden. Andererseits folgt dann ein Pose-Decoder, der den Ziel-Output entweder durch Detektion oder Regression schätzt. Als Output bieten sich einerseits 2D bzw. 3D Koordinaten von anatomischen Landmarken oder ein 3D Mesh an, wobei bei Pose-Decodern basierend auf Detektion ersteres und bei denen basierend auf Regression eher letzteres gewählt wird.[9]

Repräsentation von Posen

Grob gibt es zwei Möglichkeiten menschliche Posen darzustellen. Einerseits kann man Körper und deren Posen mithilfe von Keypoints darstellen. Dies erfolgt durch die Verwendung von 2D- oder 3D-Koordinaten, was eine einfache und intuitive Repräsentation des Körperskeletts ermöglicht. Benachbarte Keypoints werden über Linien verbunden. Dies erinnert an eine Baumstruktur, die viele Keypoints eines Körpers enthält und diese über Kanten verbindet. [15]

Andererseits gibt es modellbasierte Darstellungen, die auf den inhärenten strukturellen Merkmalen des menschlichen Körpers aufbauen und mehr Informationen im Vergleich zur Keypoint-basierten Darstellung anbieten. Dabei wird zwischen Part-basierten volumetrischen Modellen und statistischen 3D-Modellen des menschlichen Körpers unterschieden. Part-basierte volumetrische Modelle repräsentieren Gliedmaßen als Zylinder oder Ellipsoide und lösen Herausforderungen wie die Darstellung verdeckter Teile. Statistische 3D-Modelle des menschlichen Körpers bieten eine detaillierte Beschreibung der Körperhaltung und -form. [9] Als Beispiel ist das Skinned Multi-Person Linear Model (SMPL) [10] zu nennen. SMPL ist ein skelettgesteuertes menschliches Modell, das dafür entwickelt wurde Körperformen und -posen darstellen zu können. Dabei ist SMPL als ein parametrisches 3D Modell zu verstehen, bei dem einerseits die Proportion und Größe des Körpers über Form-Parameter verändert und andererseits mithilfe von Posen-Parameter entsprechend deformiert werden kann. [15]

3.2 Evaluierungsmethoden

Folgende Metriken werden insbesondere bei der 2D Pose-Estimation häufig verwendet und könnten auch für diese Arbeit relevant sein.

Prozentsatz der korrekten Teile

Diese Metrik dient dazu, die Präzision von Körperteilvorhersagen zu bewerten. Eine solche Vorhersage gilt als genau, wenn die geschätzten Endpunkte des entsprechenden Körperteils innerhalb eines festgelegten Schwellenwerts im Vergleich zu den Ground-Truth-Endpunkten liegen. Allerdings gibt es Einschränkungen bei dieser Metrik, da Körperteile je nach Perspektive und Entfernung verkürzt werden und somit keine präzise Messung möglich ist. [9]

Prozentsatz der korrekten Keypoints

Hier soll die Genauigkeit von 2D-Keypoint-Vorhersagen bewertet werden. Ein Keypoint wird dann genau vorhergesagt, wenn der Abstand zwischen dem Ground-Truth-Keypoint und dem vorhergesagten Keypoint innerhalb eines bestimmten Schwellenwertes liegt. [9]

3.3 Verfügbare Datensätze

Mögliche Datensätze, die für 2D Pose-Estimation geeignet wären, sind beispielsweise Leeds Sports Pose (LSP) [6], Frames Labeled in Cinema (FLIC) [12] und der MPII Single Person [1] Datensatz. Diese enthalten Bilder von jeweils einer Person, deren Körperpose, je nach Datensatz, mit 10-16 Keypoints annotiert ist. [9] Ein größerer Datensatz wäre der Microsoft Common Objects in Context (MS COCO) Datensatz. [8] Dieser beinhaltet im Gegensatz zu den vorhin angesprochenen Daten, Bilder von mehreren Personen und dementsprechend Posen, wobei hier für die Annotationen 17 verschiedene Keypoints verwendet werden. [9]

Für das Evaluieren von Modellen anhand Bildern von Wunden, kann die RISC GmbH Datensätze zur Verfügung stellen. Dort gibt es einerseits einen internen Datensatz zu Brandwunden und andererseits den WoundsDB [7] Datensatz, der Daten zu chronischen Wunden beinhaltet.

Quellenverzeichnis

Literatur

- [1] Mykhaylo Andriluka u. a. „2D Human Pose Estimation: New Benchmark and State of the Art Analysis“. In: *2014 IEEE Conference on Computer Vision and Pattern Recognition*. 2014, S. 3686–3693. DOI: 10.1109/CVPR.2014.471 (siehe S. 4).
- [2] Miniar Ben Gamra und Moulay A. Akhloufi. „A review of deep learning techniques for 2D and 3D human pose estimation“. *Image and Vision Computing* 114 (2021), S. 104282. DOI: <https://doi.org/10.1016/j.imavis.2021.104282> (siehe S. 2).
- [3] RISC Software GmbH. *Forschungsprojekt BurnCase 3D*. <https://www.risc-software.at/referenzprojekte/burncase/> [Zugriff: 15.01.2024] (siehe S. 1).
- [4] RISC Software GmbH. *Unternehmen*. <https://www.risc-software.at/unternehmen/> [Zugriff: 15.01.2024]. 2024 (siehe S. 1).
- [5] H.L. Haller u. a. „“Understanding burns”: Research project BurnCase 3D—Overcome the limits of existing methods in burns documentation“. *Burns* 35.3 (2009), S. 311–317. DOI: <https://doi.org/10.1016/j.burns.2008.07.010> (siehe S. 1).
- [6] Sam Johnson und Mark Everingham. „Clustered Pose and Nonlinear Appearance Models for Human Pose Estimation“. In: *British Machine Vision Conference*. 2010. URL: <https://api.semanticscholar.org/CorpusID:7318714> (siehe S. 4).
- [7] Michał Kręcihwost u. a. „Chronic wounds multimodal image database“. *Computerized Medical Imaging and Graphics* 88 (2021), S. 101844. DOI: <https://doi.org/10.1016/j.compmedimag.2020.101844> (siehe S. 4).
- [8] Tsung-Yi Lin u. a. *Microsoft COCO: Common Objects in Context*. 2015. arXiv: 1405.0312 [cs.CV] (siehe S. 4).
- [9] Wu Liu u. a. *Recent Advances in Monocular 2D and 3D Human Pose Estimation: A Deep Learning Perspective*. 2021. arXiv: 2104.11536 [cs.CV] (siehe S. 2–4).
- [10] Matthew Loper u. a. „SMPL: A Skinned Multi-Person Linear Model“. *Seminal Graphics Papers: Pushing the Boundaries, Volume 2* (2015). URL: <https://api.semanticscholar.org/CorpusID:5328073> (siehe S. 3).
- [11] N.D. Rossiter, P. Chapman und I.A. Haywood. „How big is a hand?“ *Burns* 22.3 (1996), S. 230–231. DOI: [https://doi.org/10.1016/0305-4179\(95\)00118-2](https://doi.org/10.1016/0305-4179(95)00118-2) (siehe S. 1).

- [12] Benjamin Sapp und Ben Taskar. „MODEC: Multimodal Decomposable Models for Human Pose Estimation“. *2013 IEEE Conference on Computer Vision and Pattern Recognition* (2013), S. 3674–3681. URL: <https://api.semanticscholar.org/CorpusID:12576235> (siehe S. 4).
- [13] Cornelia Schneider u. a. *Die Versorgung chronischer Wunden durch das österreichische Gesundheitssystem – eine Übersicht*. Mai 2022. DOI: [10.5281/zenodo.6517149](https://doi.org/10.5281/zenodo.6517149) (siehe S. 1).
- [14] A.B. Wallace. „THE EXPOSURE TREATMENT OF BURNS“. *The Lancet* 257.6653 (1951). Originally published as Volume 1, Issue 6653, S. 501–504. DOI: [https://doi.org/10.1016/S0140-6736\(51\)91975-7](https://doi.org/10.1016/S0140-6736(51)91975-7) (siehe S. 1).
- [15] Jinbao Wang u. a. „Deep 3D human pose estimation: A review“. *Computer Vision and Image Understanding* 210 (2021), S. 103225. DOI: <https://doi.org/10.1016/j.cviu.2021.103225> (siehe S. 3).