



Technische Universität München  
TUM School of Medicine and Health

# Klassifikation von anatomischen Varianten der Wirbelsäule und Darstellung von deren Bedeutung für die Performance Deep Learning basierter Labelling- und Segmentierungsalgorithmen

Luca Firouz Franco Malagutti

Vollständiger Abdruck der von der TUM School of Medicine and Health der Technischen  
Universität München zur Erlangung eines

Doktors der Medizin (Dr. med.)

genehmigten Dissertation.

Vorsitz: Apl. Prof. Dr. Bernhard Haslinger

Prüfende der Dissertation:

1. Apl. Prof. Dr. Jan Stefan Kirschke
2. Apl. Prof. Dr. Maria Wostrack

Die Dissertation wurde am ~~07.11.2023~~ bei der Technischen Universität München eingereicht  
und durch die TUM School of Medicine and Health am ~~13.03.2024~~ angenommen.



# Inhaltsverzeichnis

Abbildungsverzeichnis .....	3
Tabellenverzeichnis .....	3
Abkürzungen.....	4
1. Vorwort .....	6
2. Einleitung .....	8
3. Material und Methoden.....	12
3.1 Systematik der Einordnung von Übergangsanomalien .....	12
3.1.1 Allgemeines .....	12
3.1.2 Klassifikation zervikothorakaler Übergangsanomalien - Halsrippen..	12
3.1.3 Allgemeine Definitionen des thorakolumbalen Übergangs .....	15
3.1.4 Differenzierung thorakolumbalen Übergangswirbel .....	17
3.1.5 Allgemeine Definitionen des lumbosakralen Übergangs.....	19
3.1.6 Differenzierung lumbosakralen Übergangswirbel.....	23
3.1.7 Praktische Aspekte bei der Erstellung des Goldstandards .....	23
3.2 Datenerhebung .....	24
3.2.1 Allgemeine Informationen zur Datenerhebung .....	24
3.2.2 Inklusionskriterien.....	27
3.2.3 Exklusionskriterien.....	27
3.2.4 Beschreibung des definitiven Gesamtdatensatzes .....	28
3.2.5 Das "Training – Dataset" .....	29
3.2.6 Das "Public – Test – Dataset" .....	29
3.2.7 Das "Hidden – Test – Dataset" .....	30
3.3 Teilnahme an der VerSe – 20 Challenge.....	32
3.3.1 Allgemeine Informationen zur VerSe Challenge (2019 und 2020).....	32
3.3.2 Das Anduin-Tool.....	32
3.3.3 Anforderungen an eine gute Segmentierung.....	34
3.3.4 Parameter zur Bestimmung der Performance .....	34
4. Ergebnisse .....	36
4.1 Primäre Datenanalyse .....	36
4.1.1 Allgemeine Informationen zur Datenanalyse .....	36
4.1.2 Grundprävalenzen div. anatomischer Varianten der Wirbelsäule .....	37
a. Prävalenz thorakolumbalen Übergangswirbel.....	37

b. Prävalenz lumbosakraler Übergangswirbel.....	38
c. Prävalenz numerischer Anomalien der LWS .....	38
4.1.3 Koprävalenzen der einzelnen anatomischen Varianten .....	39
4.2 Resultate der VerSe – 20 Challenge.....	49
4.2.1 Gegenüberstellung von VerSe – 19 und VerSe – 20.....	49
a. Resultate der VerSe - 19 Challenge .....	49
b. Resultate der VerSe - 20 Challenge .....	51
c. Vergleich der Resultate von VerSe - 19 und VerSe - 20 .....	53
4.2.2 Bedeutung anatomischer Varianten für die VerSe.....	54
5. Diskussion.....	57
5.1 Klinische Relevanz.....	57
5.2 Grundsätzliche Aspekte .....	58
5.3 Zervikothorakaler Übergang.....	61
5.4 Thorakolumbaler Übergang.....	62
5.5 Lumbosakraler Übergang .....	64
5.6 Prävalenzen anatomischer Varianten der Wirbelsäule .....	65
5.7 VerSe - 20 Challenge.....	67
6. Zusammenfassung .....	69
7. Literaturverzeichnis.....	70

## Abbildungsverzeichnis

<i>Abbildung 1: Halsrippe unilateral an C7 in axialer Ansicht.....</i>	<i>13</i>
<i>Abbildung 2: Halsrippe unilateral an C7 in coronarer Ansicht.....</i>	<i>13</i>
<i>Abbildung 3: Halsrippen bilateral an C7 in axialer Ansicht.....</i>	<i>14</i>
<i>Abbildung 4: Halsrippen bilateral an C7 in coronarer Ansicht.....</i>	<i>14</i>
<i>Abbildung 5: Thorakolumbaler Übergangswirbel vom thorakalen Typ .....</i>	<i>18</i>
<i>Abbildung 6: Thorakolumbaler Übergangswirbel vom lumbalen Typ.....</i>	<i>18</i>
<i>Abbildung 7: Lumbosakraler Übergangswirbel nach Castellvi Grad IIa.....</i>	<i>20</i>
<i>Abbildung 8: Lumbosakraler Übergangswirbel nach Castellvi Grad IIb.....</i>	<i>21</i>
<i>Abbildung 9: Lumbosakraler Übergangswirbel nach Castellvi Grad IIIa.....</i>	<i>21</i>
<i>Abbildung 10: Lumbosakraler Übergangswirbel nach Castellvi Grad IIIb.....</i>	<i>22</i>
<i>Abbildung 11: Lumbosakraler Übergangswirbel nach Castellvi Grad IV .....</i>	<i>22</i>

## Tabellenverzeichnis

<i>Tabelle 1: Teildatensätze des Projekts .....</i>	<i>26</i>
<i>Tabelle 2: Der definitive Gesamtdatensatz.....</i>	<i>31</i>
<i>Tabelle 3: Häufigkeit thorakolumb. ÜW bei vorliegendem lumbosakr. ÜW.....</i>	<i>41</i>
<i>Tabelle 4: Häufigkeit num. Anomalien der LWS bei vorliegendem lumbosakr. ÜW .....</i>	<i>42</i>
<i>Tabelle 5: Häufigkeit lumbosakr. ÜW bei vorliegendem thorakolumb. ÜW.....</i>	<i>44</i>
<i>Tabelle 6: Häufigkeit num. Anomalien der LWS bei vorliegendem thorakolumb. ÜW .....</i>	<i>45</i>
<i>Tabelle 7: Häufigkeit lumbosakr. ÜW bei vorliegender num. Anomalie der LWS .....</i>	<i>47</i>
<i>Tabelle 8: Häufigkeit thorakolumb. ÜW bei vorliegender num. Anomalie der LWS.....</i>	<i>48</i>
<i>Tabelle 9: Ergebnisse der VerSe – 19 Challenge.....</i>	<i>51</i>
<i>Tabelle 10: Ergebnisse der VerSe – 20 Challenge.....</i>	<i>52</i>
<i>Tabelle 11: Vergleich der Performance von VerSe – 19 und VerSe – 20 .....</i>	<i>54</i>
<i>Tabelle 12: Vergleich der Performance an „Normal-Scans“ und „Varianten-Scans“ .....</i>	<i>56</i>

## Abkürzungen

AB	Aortenbifurkation
Abb.	Abbildung
BWK	Brustwirbelkörper
BWS	Brustwirbelsäule
C7	Halswirbel Nummer 7
CAD	Computer Aided Diagnosis
CM	Conus Medullaris
CNN	Convolutional Neural Network
CT	Computer Tomographie
d-mean	Localisation Distance
DICE-Score	Dice Coefficient
div.	diverse
DXA	Dual Energy X-Ray Absorptiometry
FoV	Field of View
HWS	Halswirbelsäule
ID-rate	Identification Rate
ILL	Ligamentum iliolumbale
L1	Lendenwirbel Nummer 1
L4	Lendenwirbel Nummer 4
L5	Lendenwirbel Nummer 5
L6	Lendenwirbel Nummer 6

LSIVDA	Lumbosacral Intervertebral Disc Angle
Lumbosakr.	Lumbosakral
LWS	Lendenwirbelsäule
m / w	männlich / weiblich
M.	Musculus
Max.	Maximal
MICCAI	Medical Image Computing and Computer-Assisted Intervention
MRT	Magnetresonanz Tomographie
n	Anzahl (in Tabellen)
NifTI	Neuroimaging Informatics Technology Initiative
Num.	Numerisch
PACS	Picture Archiving and Communication System
Proc. / Procc.	Processus (Singular / Plural)
RRA	Arteria renalis dextra
S1	Sakralwirbel Nummer 1
SD	Standardabweichung
Th11	Brustwirbel Nummer 11
Th12	Brustwirbel Nummer 12
Th13	Brustwirbel Nummer 13
Thorakolumb.	Thorakolumbal
ÜW	Übergangswirbel
VerSe	Vertebral Segmentation Challenge

# 1. Vorwort

Die im Folgenden dargelegte Dissertation basiert zu großen Anteilen auf meiner Mitarbeit in der Arbeitsgruppe "Deep Spine", bestehend aus Mitarbeitern des Instituts für Diagnostische und Interventionelle Neuroradiologie des Klinikums Rechts der Isar in München, sowie des Departments für Informatik der Technischen Universität München, welche im Rahmen der Teilnahme an der "Large Scale Vertebral Segmentation Challenge 2020" (VerSe 20) erfolgte. Die VerSe 20 wurde im Rahmen der MICCAI – Konferenz 2020 (International Conference on Medical Image Computing and Computer-assisted Intervention) vom 04. bis 08. Oktober 2020 in Lima, Peru, abgehalten. Dabei habe ich zunächst, indem ich im Sinne der Datenerhebung auf das institutsinterne "Picture Archiving and Communication System" (PACS) sowie auf externe Quellen (s.u.) zugegriffen habe, die aus dem Vorgängerprojekt (VerSe 19, bestehend aus Daten von insgesamt 141 Patienten) übernommenen 105 Scans von 86 Patienten um zusätzliche 214 Scans von 214 Patienten erweitert, sodass der finale Datensatz von 319 Scans von 300 Patienten entstanden ist. Die Akquise der Neudaten erfolgte hierbei durch mich. Bei allen erhobenen Daten, d. h. Neudaten sowie präexistente Scans aus dem VerSe 19 Projekt, handelt es sich um CT-Scans, welche unterschiedliche Abschnitte der Wirbelsäule abbilden (Inklusionskriterien s.u.). Die neuen, im Zuge des VerSe 20 Projektes erhobenen Scans wurden dabei insbesondere nach zwei Gesichtspunkten ausgewählt, nämlich einerseits einer möglichst großen Diversität verschiedener CT-Scanner, sowie andererseits nach dem möglichst reichhaltigen Vorliegen anatomischer Varianten der Wirbelsäule. Die Entwicklung und Weiterentwicklung des automatischen Segmentierungsalgorithmus "Anduin" erfolgte durch die informatischen Mitarbeiter der Arbeitsgruppe, während die händische Korrektur der durch diesen generierten Wirbelsäulensegmentierungen durch mich, allerdings teilweise in Zusammenarbeit mit zwei weiteren studentischen Kolleginnen, E. Niederreiter und K. Grau, vorgenommen wurde. Die Durchsicht und Kontrolle dieser händischen Korrekturen wiederum erfolgte durch zwei Neuroradiologen der Arbeitsgruppe, Prof. Dr. J. Kirschke und Dr. H. Liebl. Ich wirkte ferner bei der Erstellung des Goldstandards mit, sowie als Koautor bei der Verfassung der im Hinblick auf die Wettbewerbsteilnahme angefertigten wissenschaftlichen Arbeit "A



Computed Tomography Vertebral Segmentation Dataset with Anatomical Variations and Multi-Vendor Scanner Data" von H. Liebl, welche am 28.10.2021 in der Zeitschrift "Nature Scientific Data" publiziert wurde (Liebl et al., 2021). Des Weiteren basiert die auf die informatischen Inhalte des Projektes sich konzentrierende, ebenfalls von Mitarbeitern unserer Arbeitsgruppe verfasste, wissenschaftliche Arbeit "A Vertebrae Labelling and Segmentation Benchmark for Multi-detector CT Images" von A. Sekuboyina auf den Ergebnissen der Wettbewerbsteilnahme (Sekuboyina et al., 2021). Die im Anschluss für den Schreibeil dieser Dissertation erstellten Auswertungen und Berechnungen im Hinblick auf die Prävalenzen verschiedener anatomischer Varianten der Wirbelsäule, der Koprävalenzen derselben untereinander, d.h. die Häufigkeit des gleichzeitigen Vorliegens verschiedener anatomischer Varianten der Wirbelsäule bei einem einzelnen Individuum, sowie der Vergleich der verschiedenen Performance-Indices der Segmentierungsalgorithmen, welche als solche der Arbeit von A. Sekuboyina entnommen worden sind (Sekuboyina et al., 2021), erfolgten in Eigenarbeit durch mich.

## 2. Einleitung

Die Wirbelsäule lässt sich beim erwachsenen Menschen untergliedern in einen zervikalen, einen thorakalen und einen lumbalen Abschnitt, sowie das Sakrum (Kreuzbein) und den Coccyx (Steißbein). Geht man von der anatomischen Norm aus, so setzt sich der zervikale Anteil aus sieben Wirbeln zusammen, der thorakale Anteil aus zwölf Wirbeln und der lumbale Anteil sowie das Sakrum und das Steißbein aus jeweils fünf Wirbeln. Anatomische Varianten wie numerische Aberrationen, d.h. ein Abweichen der zuvor erwähnten Normverteilung der Wirbel auf die verschiedenen Abschnitte der Wirbelsäule, aber auch nichtnumerische Anomalien wie etwa Übergangswirbel, auf die im Folgenden genauer eingegangen werden soll, sind in der unselektierten Normalbevölkerung häufig anzufinden (Thawait et al., 2012; Tins & Balain, 2016) und können jeden Abschnitt der Wirbelsäule betreffen. Entsprechend der klinischen Relevanz diesbezüglich wird in dieser Arbeit das Hauptaugenmerk auf den thorakolumbalen und den lumbosakralen Übergang gelegt werden (s.u.).

Während die Halswirbelsäule im Grunde konstant aus sieben Wirbeln zu bestehen scheint, sind numerische Aberrationen der Brust- und Lendenwirbelsäule häufiger (Wigh, 1980). Auch übersteigt die Häufigkeit von Übergangswirbeln des thorakolumbalen und des lumbosakralen Übergangs diejenige von Anomalien des zervikothorakalen Übergangs wie bspw. Halsrippen (Thawait et al., 2012). Insgesamt sind laut Tins et al. in der Normalbevölkerung numerische Aberrationen häufiger zu beobachten als Übergangswirbel (Tins & Balain, 2016).

Fehlen einem Wirbel am Übergang zu einem angrenzenden Wirbelsäulenabschnitt die spezifischen Charakteristika des Abschnittes, dem der Wirbel nach normbasierter Zählung in kraniokaudaler Richtung zuzuordnen ist (Bsp.: Th12 mit Stummelrippe uni- oder bilateral oder ohne jegliche artikulierenden Rippen), oder weist der Wirbel Charakteristika beider angrenzender Abschnitte gleichzeitig auf (Bsp.: L5 mit Proc. transversus, der uni- oder bilateral mit dem Sakrum verschmolzen ist oder artikuliert), so handelt es sich um einen Übergangswirbel (Wigh, 1980). Es soll an dieser Stelle erwähnt werden, dass die Identifikation eines Wirbels als Übergangswirbel auch stets bedeutet, dass in Abhängigkeit verschiedener morphologischer Charakteristika (s.u.) die zunächst vermutete Zugehörigkeit des Wirbels zu einem bestimmten Wirbelsäulenabschnitt infrage gestellt werden muss (Bsp.: Vermeintlicher L1 mit

Stummelrippen könnte sich auch als Th13 herausstellen.). Die Prävalenz von Übergangswirbeln in der Normalbevölkerung wird von Park et al. auf etwa 12,6 % geschätzt (Park et al., 2016).

Russell Wigh hat gezeigt, dass ein zuverlässiges und einheitliches Klassifikationssystem für Wirbelsäulenanomalien von großer Bedeutung für die klinische Praxis ist (Wigh, 1980). Während es in einem klinischen Kontext insbesondere die zwischen den einzelnen Wirbeln austretenden Nervenwurzeln, beziehungsweise die jeweils dazugehörigen Kennmuskeln und Dermatome sind, die eine wertvolle Auskunft darüber liefern können, wie ein zunächst nicht klar zu beziffernder Wirbel letztendlich einzuordnen ist (vgl. Kapitel 5.2) (Thawait et al., 2012), so ist doch im Rahmen eines solchen Projektes wie der VerSe 20, in dessen Rahmen lediglich radiologische Aspekte betrachtet werden können, vorrangig eine einheitliche, reproduzierbare Nomenklatur vonnöten, welche bei der bildgeberischen Diagnostik, und ebenso bei der Nutzung automatisierter Segmentierungsalgorithmen zur Anwendung kommen kann. Auch Russell Wigh macht auf die Notwendigkeit einer solchen umfassenden Standardisierung der Nomenklatur und Klassifikation von Wirbelsäulenanomalien aufmerksam, und er begründet dies in einem klinischen Zusammenhang dadurch, dass es innerhalb einer betrachteten Stichprobe von Patienten, welche jeweils eine Laminektomie bei lumbaler Diskushernie erhalten haben, ausschließlich bei denjenigen zu nomenklatorischen Fehlern und letztlich operativen Eingriffen am falschen Wirbelsäulensegment kam, bei denen eine anatomische Variante der Wirbelsäule vorlag (Wigh, 1980). Für eine einheitliche Systematik in diesem Sinne liegen in der wissenschaftlichen Literatur bereits diverse Vorschläge vor (Castellvi et al., 1984; Thawait et al., 2012; Wigh, 1980), jedoch sollte zunächst erwähnt werden, dass prinzipiell zwei Möglichkeiten bestehen, nämlich einerseits die Nummerierung der einzelnen Wirbelkörper in der konventionellen, kraniokaudalen Zählweise, unter Vernachlässigung morphologischer Aspekte, und andererseits die Zuordnung der einzelnen Wirbel anhand morphologischer Charakteristika, wie in Kapitel 5 näher erörtert werden wird.

Die folgende Arbeit wird sich primär auf ein von uns im Rahmen der auf der MICCAI Conference 2020 in Lima, Peru abgehaltenen Vertebral Segmentation Challenge 2020 vorgestelltes Klassifikationssystem berufen, welches mitunter auf den Erkenntnissen der Fachliteratur basiert und dessen Reproduzierbarkeit regelmäßig mithilfe automatisierter Segmentierungsalgorithmen überprüft worden ist, wie später

in dieser Arbeit näher erläutert werden wird. Dieses Klassifikationssystem wurde in einer im Oktober 2021 von Liebl et al. in der Fachzeitschrift "Nature Scientific Data" veröffentlichten Arbeit bereits vorgestellt (Liebl et al., 2021). Es soll jedoch in dieser Arbeit nochmals in seinen Einzelheiten und seinem Zustandekommen nachvollziehbar erklärt werden (vgl. Kapitel 3).

Das Prinzip der sog. "computer aided diagnosis" (CAD) ist insbesondere für die Neuroradiologie von großem Interesse, und mögliche Anwendungsgebiete wurden in der Literatur diskutiert (Kaka et al., 2021). Im Bereich der Wirbelsäulendiagnostik ist bereits gezeigt worden, dass die Erkennung osteoporotischer Veränderungen durch CAD deutlich verbessert werden kann (Hussain & Han, 2019). Auch Löffler et al. sehen das Prinzip der CAD als eine Chance zur Vermeidung von Diagnosefehlern und zur verbesserten Detektion osteoporotischer Frakturen in CT-Bilddaten an (Löffler et al., 2020). Bei der Identifikation osteoporotischer Wirbel haben sich diverse, von der Verwendung automatisierter Segmentierungsalgorithmen in CT-Daten abhängige Methoden bereits als der DXA-Knochendichtemessung überlegen erwiesen (Löffler et al., 2021). Es wurde gezeigt, dass mithilfe dieser Methoden ein sinnvolles Osteoporose Screening der Wirbelsäule möglich ist (Löffler et al., 2021). Analog dazu wäre also denkbar, dass die Verwendung von CAD langfristig auch zur Reduktion des bereits erwähnten Problems der operativen Eingriffe am falschen Wirbelsäulensegment, beitragen kann. Vor dem Hintergrund der beschriebenen Problematik ist zu diesem Zweck eine einheitliche Systematik zur Klassifikation anatomischer Varianten der Wirbelsäule prinzipiell unerlässlich (Wigh, 1980). Die Segmentierungsalgorithmen, die bei der Erstellung des o.g. Klassifikationssystems eine wesentliche Rolle gespielt haben, arbeiten mit einem sog. "Deep Learning" (Sekuboyina et al., 2021). Das "Deep Learning" bezeichnet hierbei einen Subtypen des sog. "Machine Learning", also einer mit einer gewissen Anpassungsfähigkeit funktionierenden künstlichen Intelligenz, die anhand eines Dateninputs zu lernen imstande ist (Choi et al., 2020). Besonders für das "Deep Learning" innerhalb des Spektrums des "Machine Learning" ist die Verwendung sog. "Tiefer Neuronaler Netzwerke", eines Informationsverarbeitungssystems also, das seiner Funktionsweise nach dem menschlichen Gehirn entlehnt ist, und dabei mehrere Verarbeitungsebenen – "Input Layer", mehrere "Hidden Layer", "Output Layer" – beinhaltet (Choi et al., 2020). Der Effekt des "Deep Learning" ist grundsätzlich umso größer, je größer der verwendete Dateninput ist (Choi et al., 2020). Da sein

Einsatzgebiet insbesondere auch die Domäne der Bilderkennung umfasst (Choi et al., 2020), eignet es sich als Funktionsweise für ein Programm, welches Fertigkeiten in Mustererkennung, dem Labelling und der Segmentierung von Abschnitten der Wirbelsäule erwerben soll (Sekuboyina et al., 2021). Das bereits erwähnte Klassifikationssystem wurde im Rahmen dieser Arbeit entwickelt und hat den Anspruch, alle existenten Wirbelsäulenanomalien einheitlich zuzuordnen und der Erkennung durch automatisierte Segmentierungsalgorithmen zugänglich zu machen. Die verbesserte Erkennung sowie Prozessierung anatomischer Varianten der Wirbelsäule durch automatisierte Segmentierungsalgorithmen sind schließlich für die Steigerung der allgemeinen Performance derselben von erheblicher Bedeutung (Sekuboyina et al., 2021).

## 3. Material und Methoden

### 3.1 Systematik der Einordnung von Übergangsanomalien

#### 3.1.1 Allgemeines

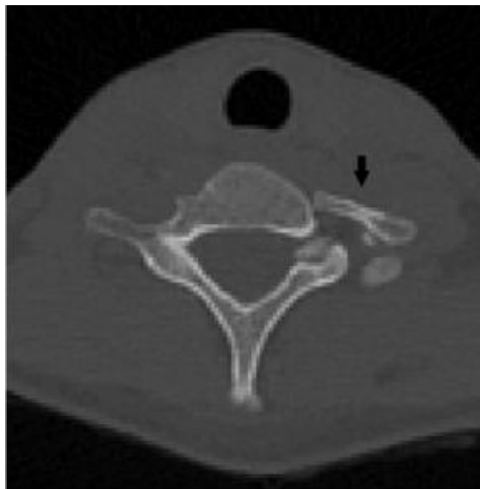
Da eine korrekte und reproduzierbare Segmentierung der Wirbelsäule durch automatisierte Algorithmen vorrangig morphologischer Anhaltspunkte bedarf, orientiert sich die von uns im Rahmen dieses Forschungsprojektes ausgearbeitete Systematik der Klassifikation von Übergangswirbeln jeglicher Art in erster Linie an solchen Kriterien, die zuverlässig in der CT-Bildgebung zu ermitteln sind (Liebl et al., 2021). Schließlich beschränkt sich der in diesem Projekt bearbeitete Bilddatensatz auf diese Modalität (Liebl et al., 2021). Im Entstehungsprozess dieser einheitlichen Klassifikation, die im Vorfeld der Wettbewerbsteilnahme bei der MICCAI-Challenge 2020 von uns erstellt wurde, spielten auch die Erkenntnisse vorbestehender Literatur über anatomische Varianten der Wirbelsäule eine tragende Rolle (Brewin et al., 2009; Castellvi et al., 1984; Thawait et al., 2012; Wigh, 1980). Inwiefern diesbezüglich die Vorschläge diverser Autoren in die von uns vorgestellte Systematik integriert wurden oder aber im Hinblick auf unsere Arbeit verworfen worden sind, und an welcher Stelle neue theoretische Ansätze im Rahmen unseres Projektes entwickelt worden sind, ist in Kapitel 5 ausführlich dargestellt. Im Folgenden sollen die theoretischen Grundlagen des Klassifikationssystems erläutert werden, das ich in Zusammenarbeit mit den anderen Mitgliedern der Arbeitsgruppe erstellt habe, und das letztendlich auch in die Arbeit von Liebl et al. Eingang gefunden hat (Liebl et al., 2021).

#### 3.1.2 Klassifikation zervikothorakaler Übergangsanomalien - Halsrippen

Wie bereits Thawait et al. haben auch wir uns dazu entschieden, die Halswirbelsäule als fixiert 7-gliedrig anzuerkennen, und gleichzeitig den siebten Halswirbel (C7) per Definition als den einzigen Wirbel festzulegen, der mit einer Halsrippe artikulieren kann (Liebl et al., 2021; Thawait et al., 2012). Bezüglich der Kriterien für das Vorhandensein einer solchen Halsrippe beziehen wir uns auf die von Brewin et al.

erarbeiteten Kriterien (Brewin et al., 2009), die im Folgenden genannt sind. Zum besseren Verständnis siehe auch die Abb. 1 – 4.

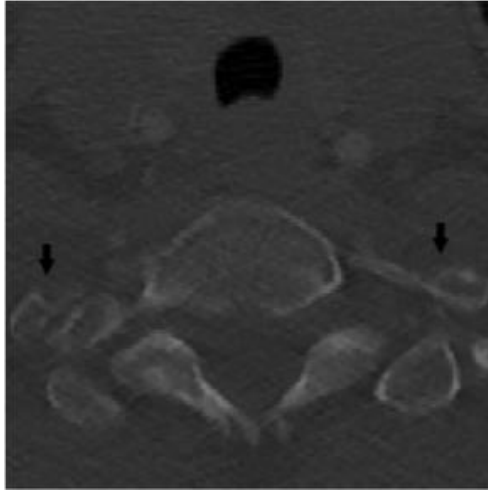
- i) Verbindung der Rippe mit dem Proc. transversus von C7
- ii) Ausschluss einer direkten Verbindung der Rippe zum Manubrium sterni
- iii) Die Verbindung der Rippe mit dem Proc. transversus von C7 ist von artikulierender Natur und keine knöcherner Verschmelzung.



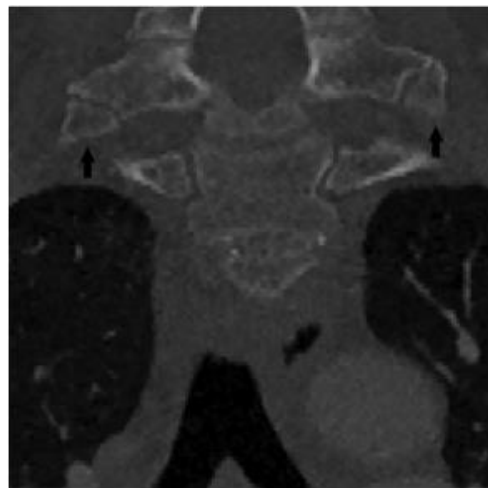
*Abbildung 1:  
Halsrippe unilateral (Pfeil) an C7 in axialer Ansicht*



*Abbildung 2:  
Halsrippe unilateral (Pfeil) an C7 in coronarer Ansicht*



*Abbildung 3:  
Halsrippen bilateral (Pfeile) an C7 in axialer Ansicht*



*Abbildung 4:  
Halsrippen bilateral (Pfeile) an C7 in coronarer Ansicht*



### 3.1.3 Allgemeine Definitionen des thorakolumbalen Übergangs

Prinzipiell kann die Ursache jeder numerischen Aberration der Wirbelsäule (d.h. die Verteilung der freien Wirbelkörper auf die Abschnitte der Wirbelsäule weicht von dem Schema 7/12/5 ab), sowohl durch isolierte Veränderungen am thorakolumbalen Übergang als auch durch isolierte Veränderungen am lumbosakralen Übergang, oder durch Veränderungen an den beiden Übergängen gleichzeitig, bedingt sein. In jedem Fall also interessiert die genaue Anzahl vorliegender Brustwirbel sowie Lendenwirbel. Bei der genaueren Differenzierung rein numerischer Aberrationen kann also die Anzahl an Brustwirbeln verändert sein (11 oder 13 anstatt 12) oder die Anzahl an Lendenwirbeln (4 oder 6 anstatt 5) oder eben beide in verschiedenen denkbaren Kompositionen.

Es war für unser Projekt zunächst essenziell, über exakte Kriterien für das Vorliegen eines regelhaften thorakalen bzw. lumbalen Wirbels in Abgrenzung zum Vorliegen eines Übergangswirbels zu verfügen. Hierbei haben wir uns im Wesentlichen der Kriterien nach Russell Wigh bedient (Liebl et al., 2021; Wigh, 1980). Demnach ist ein regelhafter thorakaler Wirbel ein solcher, der beidseits mit Rippen normaler Länge (> 38 mm) artikuliert (Wigh, 1980). Ein regelhafter lumbaler Wirbel ist ein solcher, der beidseits knöchern verschmolzene Procc. transversi aufweist (Wigh, 1980). Folglich wird jeder Wirbel am thorakolumbalen Übergang, der nach diesem Vorgehen weder als regulär thorakal noch als regulär lumbal kategorisiert werden kann, als thorakolumbalen Übergangswirbel angesehen (Wigh, 1980).

Die Tatsache, dass die im Rahmen des Projektes erarbeiteten Kriterien primär dem Ziel dienen, eine reproduzierbare maschinelle Segmentierbarkeit insbesondere auch eben solcher Übergangswirbel zu ermöglichen (Sekuboyina et al., 2021), verlangt der Klassifikation eine genauere Differenzierung dieser thorakolumbalen Übergangswirbel ab. So muss ein solcher Wirbel in dieser Arbeit als "dem thorakalen Bereich zugehörig" oder "dem lumbalen Bereich zugehörig" identifizierbar sein, damit es keine Verwirrungen geben kann bei der Zuordnung des Wirbels zu einem Label durch den Algorithmus oder die den Goldstandard erstellenden Mitarbeiter (siehe auch Kapitel 3.1.7). Grundsätzlich schlägt Russell Wigh die Einführung von vier klar voneinander abzugrenzenden Kategorien vor, welche im Folgenden genannt sind (Wigh, 1980), welche jedoch für die Zwecke dieser Arbeit nicht scharf genug

zwischen dem thorakalen und dem lumbalen Typ eines thorakolumbalen Übergangswirbels zu unterscheiden vermögen, sodass sie um die in Kap. 3.1.4 aufgeführten Differenzierungskriterien erweitert wurden (Liebl et al., 2021).

□ Normaler thorakaler Wirbel:

Ein Wirbel am thorakolumbalen Übergang, der bilateral mit jeweils einer Rippe artikuliert, deren Länge mehr als 38 mm beträgt, kann laut Russell Wigh als normaler thorakaler Wirbel bezeichnet werden (Wigh, 1980).

□ Normaler lumbaler Wirbel:

Enthält ein Wirbel am thorakolumbalen Übergang bilateral je einen Processus transversus, der mit dem Wirbelkörper knöchern verschmolzen ist, so handelt es sich laut Russell Wigh um einen normalen lumbalen Wirbel (Wigh, 1980).

□ Thorakolumbalen Übergangswirbel vom thorakalen Typ:

Artikuliert ein Wirbel am thorakolumbalen Übergang uni- oder bilateral mit einer Stummelrippe (Länge = 38 mm oder kleiner), und weist/weisen die Rippe(n) eine charakteristische Rippenmorphologie auf, so handelt es sich laut Russell Wigh um einen thorakolumbalen Übergangswirbel vom thorakalen Typ (d.h. dem thorakalen Abschnitt der Wirbelsäule zuzurechnen) (Wigh, 1980).

□ Thorakolumbalen Übergangswirbel vom lumbalen Typ:

Analog zum vorigen Punkt liefert laut Russell Wigh bei einem Wirbel am thorakolumbalen Übergang das Vorhandensein eines nicht knöchern mit dem Wirbelkörper verschmolzenen Processus transversus uni- oder bilateral, in Ermangelung charakteristischer Rippenmorphologie (z.B. Bei Vorliegen eines akzessorischen Ossifikationszentrums mit bulbärer Form), einen Hinweis darauf, dass es sich um einen thorakolumbalen Übergangswirbel vom lumbalen Typ handelt (d.h. dem lumbalen Abschnitt der Wirbelsäule zuzurechnen.) (Wigh, 1980).

### 3.1.4 Differenzierung thorakolumbaler Übergangswirbel

Bei der genaueren Differenzierung thorakolumbaler Übergangswirbel nach thorakalem bzw. lumbalem Typ erwies sich im Rahmen dieser Arbeit eine Kombination aus "harten" Grundkriterien und "weichen" Zusatzkriterien als sinnvoll (Liebl et al., 2021), die von studentischen und ärztlichen Mitarbeitern der Forschungsgruppe zwecks der Erstellung des Goldstandards definiert worden sind und im Folgenden genannt werden. Zum besseren Verständnis siehe auch die Abb. 5 und 6.

#### i) Harte Kriterien

- Ein thorakolumbaler Übergangswirbel ist dann der Brustwirbelsäule zuzurechnen, wenn er mindestens eine artikulierende Rippe der Länge > 38 mm aufweist.
- Ein thorakolumbaler Übergangswirbel, der zwei Stummelrippen (= 38 mm oder < 38 mm) enthält bzw. andere Strukturen (z.B. akzessorisches Ossifikationszentren), die jedenfalls keine Rippen der Länge > 38 mm sind, kann anhand der weichen Kriterien näher subklassifiziert werden.

#### ii) Weiche Kriterien, die eine Zugehörigkeit des ÜW zur Lendenwirbelsäule suggerieren:

- Horizontaler Verlauf der Procc. transversi bzw. der Stummelrippen
- Facies articulares der Procc. articulares in sagittaler Stellung
- Geradliniger, nach dorsal gerichteter Verlauf der Procc. spinosi

#### iii) Weiche Kriterien, die eine Zugehörigkeit des ÜW zur Brustwirbelsäule suggerieren:

- Deszendierender Verlauf der Procc. transversi bzw. der Stummelrippen
- Facies articulares der Procc. articulares in coronarer Stellung
- Schräger, nach dorsokaudal gerichteter Verlauf der Procc. spinosi

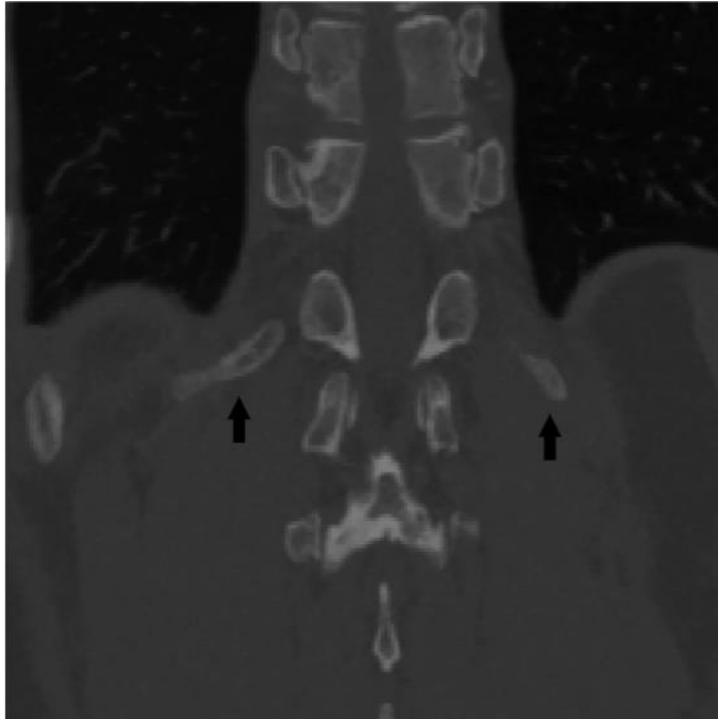


Abbildung 5:  
Thorakolumbaler Übergangswirbel vom thorakalen Typ mit Stummelrippen (Pfeile) an Th12



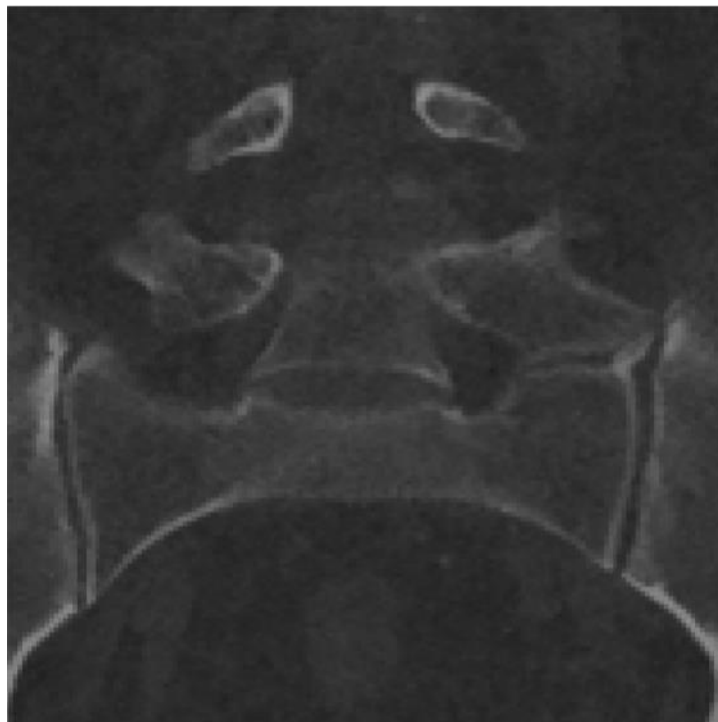
Abbildung 6:  
Thorakolumbaler Übergangswirbel vom lumbalen Typ mit Stummelrippen (Pfeile) an L1

### 3.1.5 Allgemeine Definitionen des lumbosakralen Übergangs

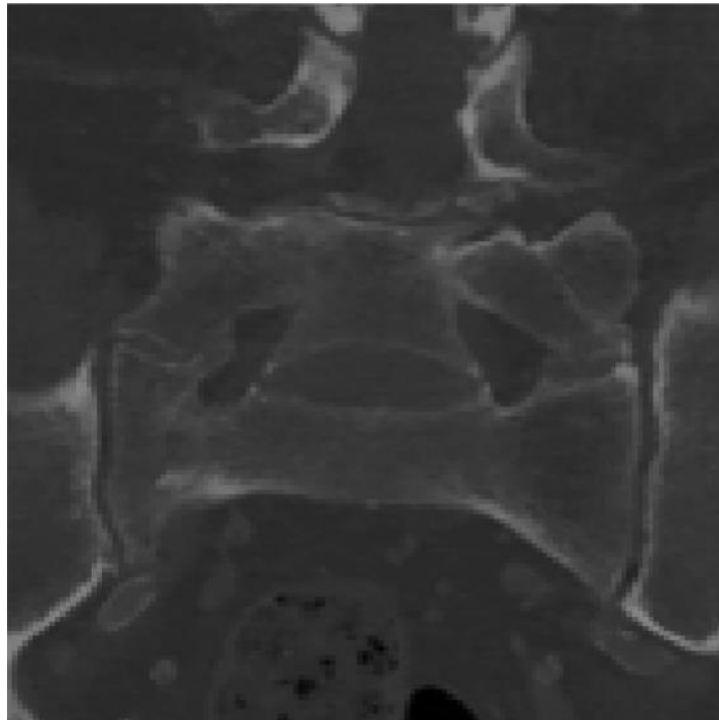
Auch am lumbosakralen Übergang können rein numerische Aberrationen von Anomalien mit Übergangswirbeln unterschieden werden. Jeder Wirbel am lumbosakralen Übergang, der ein freier Wirbel ist und dessen Procc. transversi nicht knöchern mit dem Sakrum verschmolzen sind (uni- oder bilateral) oder mit dem Sakrum artikulieren (uni- oder bilateral) kann als regelhafter lumbaler Wirbel angesehen werden (Wigh, 1980). Liegt ein Übergangswirbel vor, der sich in eine der Kategorien der Castellvi-Klassifikation einordnen lässt (Castellvi et al., 1984), so kann dessen Zuordnung zum lumbalen bzw. sakralen Teil der Wirbelsäule ähnlich wie in Kapitel 3.1.4 beschrieben mithilfe weiterer "weicher" Kriterien vorgenommen werden. Die Klassifikation nach Castellvi liefert im Rahmen unseres Projektes keine harten Kriterien für die Zuordnung eines Übergangswirbels zu einem bestimmten Wirbelsäulenabschnitt, sondern erfüllt hauptsächlich methodische Zwecke im Hinblick auf die korrekte Segmentierung (siehe Kapitel 3.1.7). Sicher steht jedoch zumindest, dass jeder Wirbel, der sich anhand der Castellvi-Klassifikation als Castellvi IIa/b, IIIa/b, oder IV kategorisieren lässt, in unserem Datensatz als lumbosakraler Übergangswirbel angesehen wurde. Auf die exakte Bestimmung des Castellvi Grades I in unseren Daten wurde aufgrund der erschwerten Abgrenzbarkeit und eingeschränkten Aussagekraft verzichtet (vgl. Kapitel 4.1.2). Die Kriterien der Castellvi – Klassifikation sind im Folgenden aufgelistet. Zum besseren Verständnis siehe auch die Abb. 7 – 11.

- Castellvi I:  
Der Übergangswirbel weist zum Sakrum hin uni- (Castellvi Ia) oder bilateral (Castellvi Ib) einen vergrößerten, dysplastischen Processus transversus mit kraniokaudalem Durchmesser = 19 mm oder > 19 mm auf (Castellvi et al., 1984).
- Castellvi II:  
Der Übergangswirbel weist zum Sakrum hin uni- (Castellvi IIa) oder bilateral (Castellvi IIb) einen vergrößerten, dysplastischen Processus transversus mit Pseudarthrose zur Ala ossis sacri auf (Castellvi et al., 1984).

- Castellvi III:  
Der Übergangswirbel weist uni- (Castellvi IIIa) oder bilateral (Castellvi IIIb) einen knöchern mit der Ala ossis sacri verschmolzenen Processus transversus auf (Castellvi et al., 1984).
- Castellvi IV:  
Hier liegt eine Mischvariante vor, bei der der Übergangswirbel auf einer Seite das Bild eines Castellvi II, kontralateral das Bild eines Castellvi III zeigt (Castellvi et al., 1984).



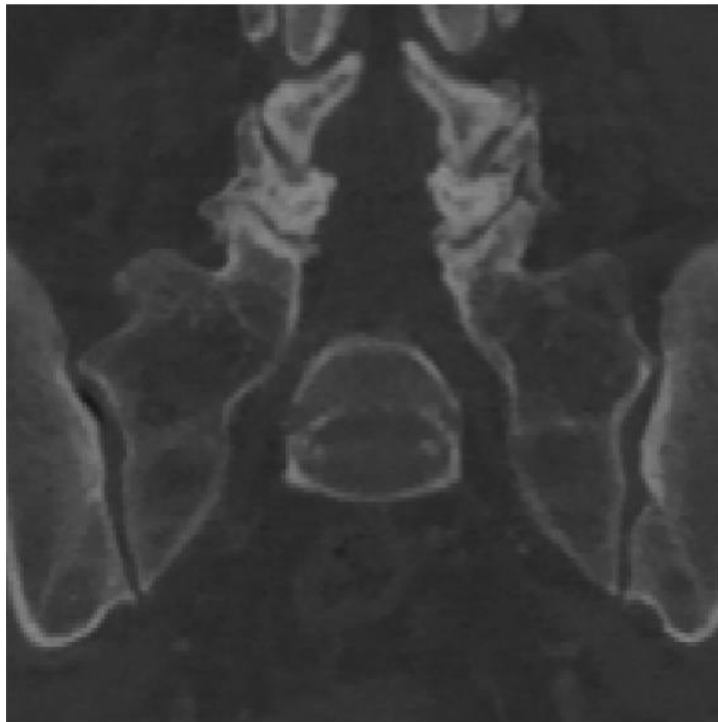
*Abbildung 7:  
Lumbosakraler Übergangswirbel nach Castellvi Grad IIa*



*Abbildung 8:  
Lumbosakraler Übergangswirbel nach Castellvi Grad IIb*



*Abbildung 9:  
Lumbosakraler Übergangswirbel nach Castellvi Grad IIIa*



*Abbildung 10:  
Lumbosakraler Übergangswirbel nach Castellvi Grad IIIb*



*Abbildung 11:  
Lumbosakraler Übergangswirbel nach Castellvi Grad IV*



### 3.1.6 Differenzierung lumbosakraler Übergangswirbel

Neben der Castellvi-Klassifikation, die in dieser Arbeit zunächst der Identifikation lumbosakraler Übergangswirbel als solcher dient, haben wir uns auch hier einer Reihe "weicher" Kriterien bedient, um die Zugehörigkeit eines solchen Übergangswirbels zu einem der beiden angrenzenden Wirbelsäulenabschnitte zu bestimmen. Es muss erwähnt werden, dass die hier vorgenommene Unterscheidung ein Konstrukt ist, das primär der Vollständigkeit unserer theoretischen Annahmen dient, dass jedoch faktisch in dem von uns erhobenen Datensatz kein einziger Fall identifiziert wurde, in dem ein lumbosakraler Übergangswirbel als dem Sakrum zugehörig kategorisiert wurde (Liebl et al., 2021). Es fanden sich also ausschließlich solche lumbosakralen Übergangswirbel, die sakralisierte Lumbalwirbel waren (Liebl et al., 2021). Dennoch sollen diese rein theoretischen Überlegungen im Folgenden dargestellt werden.

- i) Weiche Kriterien, die eine Zugehörigkeit des ÜW zur Lendenwirbelsäule suggerieren:
  - Vorliegen jeglicher Kategorien nach Castellvi mit Ausnahme der Kategorie 3b
  - Kubische Form des Wirbelkörpers in sagittaler Ansicht
  - Verbindung des Wirbels zum ILL (sofern erkennbar)
  
- ii) Weiche Kriterien, die eine Zugehörigkeit des ÜW zum Sakrum suggerieren:
  - Vorliegen der Kategorie nach Castellvi 3b
  - Konische Form des Wirbelkörpers in sagittaler Ansicht

### 3.1.7 Praktische Aspekte bei der Erstellung des Goldstandards

Von größerer Bedeutung als die oben genannte theoretischen Einteilung der lumbosakralen Übergangswirbel nach lumbal oder sakral war in unserem Projekt eine klare Vorstellung davon, welche Wirbel überhaupt durch den Algorithmus bzw. im Rahmen der Erstellung des Goldstandards gelabelt werden sollen, und wie genau dieses Labelling beschaffen sein sollte. Hierbei orientierten wir uns wiederum an der

Castellvi-Klassifikation (Castellvi et al., 1984). Wir beschlossen, dass die Sequenz der Nummerierung insofern stereotyp vorgenommen werden sollte, als das Label L5 (sofern vorhanden!) entweder das Ende der Segmentierung darstellt oder dass darauf das Label L6 folgt (Liebl et al., 2021). Auf ein Label S1 wurde verzichtet (Liebl et al., 2021). Ebenso wurde festgelegt, dass mit dem Sakrum knöchern verschmolzene Strukturen nicht segmentiert werden sollen (Liebl et al., 2021). Daraus geht hervor, dass alle Übergangswirbel der Kategorien Castellvi 3a/b oder 4 nicht segmentiert wurden (Liebl et al., 2021).

Vereinfachend, jedoch morphologisch durchaus begründbar, wurde angenommen, dass es sich bei allen segmentierten Wirbeln am lumbosakralen Übergang um solche Wirbel handelt (Normalwirbel oder Übergangswirbel), die der lumbalen Wirbelsäule zuzurechnen sind. Diese Annahme ist vereinbar mit der Tatsache, dass keine sakralen, sondern ausschließlich lumbale Label verwendet wurden (Sekuboyina et al., 2021). Allerdings wurde die Möglichkeit, dass ein nichtsegmentierter Wirbel formal als Teil der Lendenwirbelsäule anzusehen ist, nicht ausgeschlossen (z.B. ein lumbosakraler Übergangswirbel der Kategorie Castellvi 3a mit typischer Morphologie eines Lendenwirbels).

## 3.2 Datenerhebung

### 3.2.1 Allgemeine Informationen zur Datenerhebung

Für die Erstellung des Gesamtdatensatzes im Rahmen der Teilnahme an der MICCAI-Challenge 2020 wurden 319 CT-Scans von insgesamt 300 Patienten akquiriert (Liebl et al., 2021). Dabei setzt sich dieses Kollektiv aus zwei nach definierten Gesichtspunkten erstellten Untergruppen zusammen, nämlich einer aus 150 Patienten bestehenden Untergruppe, die möglichst vielfältig die verschiedenen anatomischen Varianten der Wirbelsäule abbildet, und einer weiteren, ebenfalls aus 150 Patienten bestehenden Untergruppe, die Scans von verschiedenen CT-Scanner-Herstellern aufweist. Zusätzlich zu den im Projekt erneut verwendeten CT-Daten von 86 Patienten aus dem Vorgängerprojekt VerSe 19 (Sekuboyina et al., 2021), die entsprechend den vorliegenden anatomischen Verhältnissen bzw. des Scanner-Typs entweder der ersten oder der zweiten Untergruppe zugeteilt wurden, wurden für den Gesamtdatensatz CT-Daten von 214 weiteren Patienten erhoben (Liebl et al., 2021).

Für die Erhebung dieser Neudaten habe ich aus dem hausinternen PACS eine Anzahl von zunächst 308 CT-Scans, und aus einem externen, von Glocker et al. (Glocker et al., 2013) veröffentlichten Datensatz 30 weitere CT-Scans extrahiert (Liebl et al., 2021). Nachdem eine bestimmte Anzahl der erhobenen Scans willkürlich ausgewählt wurde, ergab sich in Addition zu den 86 Fällen aus VerSe 19 der resultierende Gesamtdatensatz, bestehend aus 319 CT-Scans von 300 Patienten, das anschließend in drei Untergruppen aufgeteilt wurde ("Training-Dataset", "Private-Test-Dataset", "Hidden-Test-Dataset"; siehe Kapitel 3.2.5, 3.2.6, 3.2.7). Einzelheiten zum Datensatz können Tabelle 1 entnommen werden (s.u.), die den von der Arbeitsgruppe publizierten Daten entlehnt ist (Liebl et al., 2021). Die Datenakquise und -verarbeitung wurde vorab von der Ethikkommission des Klinikums Rechts der Isar genehmigt, während auf eine schriftliche Einwilligung der betreffenden Patienten aufgrund umfassender Anonymisierungsmaßnahmen, welche im Sinne des Datenschutzes durchgeführt wurden, verzichtet wurde (Liebl et al., 2021). Im Folgenden sollen nun die exakten Ein- und Ausschlusskriterien sowie die Einteilung der Daten in die Untergruppen näher erläutert werden.

*Tabelle 1: Teildatensätze des Projekts (Liebl et al., 2021)*

	Hidden Test Dataset	Public Test Dataset	Training Dataset	Gesamt-dataset
Anzahl Patienten (n)	100	100	100	300
Geschlecht w/m (n)	45/55	49/51	50/50	144/156
Mittl. Alter (n +/- SD)	57,9 +/- 17,6	54,5 +/- 17,2	56,3 +/- 18	56,2 +/- 17,6
Anzahl Scans (n)	103	103	113	319
Anzahl gelabelter Wirbel (n)	1348	1366	1428	4142
Davon cervicale Wirbel (n)	193	164	224	581
Davon thorakale Wirbel (n)	728	770	757	2255
Davon lumbale Wirbel (n)	427	432	447	1306
Nach Scanner-Hersteller:				
Philips intern (n)	40	49	57	146
Siemens intern (n)	23	14	36	73
GE extern (n)	10	10	0	20
Siemens extern (n)	10	10	10	30
Toshiba extern (n)	10	10	0	20
Unbekannt extern (n)	10	10	10	30

### 3.2.2 Inklusionskriterien

Auf diejenigen Daten, die aus dem VerSe 19 Projekt übernommen worden sind, treffen die unten genannten Inklusionskriterien womöglich nicht exakt zu. Hierfür muss auf die entsprechende Literatur verwiesen werden (Sekuboyina et al., 2021).

Für alle neu erhobenen Daten gelten die im Zuge dieses Projektes ausgearbeiteten Inklusionskriterien, die im Folgenden aufgelistet sind und auch von Liebl et al. als solche angeführt werden (Liebl et al., 2021).

- Modalität der Bildgebung: CT
- Auflösung mit max. Schichtdicke 1,5 mm axial, 1 mm coronar, 3 mm sagittal
- Untersuchungszeitraum zwischen 05.02.2016 und 01.03.2020
- Minimale Anzahl vollständig abgebildeter freier Wirbel: 7
- Patientenalter > 18 Jahre

Wie zuvor erwähnt, ist die Datensuche im hausinternen PACS insbesondere nach zwei gezielten Gesichtspunkten erfolgt (Liebl et al., 2021). Erstens wurde im Sinne des Trainingseffekts für den im Projekt verwendeten Algorithmus ein ausreichend großes Patientenkollektiv mit ärztlich befundeter oder im Prozess festgestellter Übergangsanomalie der Wirbelsäule angestrebt (Liebl et al., 2021). Zudem wurde versucht, bei den inkludierten Daten eine möglichst große Vielfalt hinsichtlich verschiedener CT-Scanner-Hersteller (Siemens, Toshiba, GE) zu erreichen, mit dem Hintergrund, eine gewisse Generalisierbarkeit der Ergebnisse abzubilden (Liebl et al., 2021).

### 3.2.3 Exklusionskriterien

Die ebenfalls von Liebl et al. angeführten Exklusionskriterien wurden, ebenso wie die Inklusionskriterien, eigens für die neu erhobenen Daten im Rahmen des Projekts erarbeitet (Liebl et al., 2021). Im Folgenden sind diese Kriterien aufgelistet.

- Vorhandensein destruktiver Wirbelkörpermetastasen im abgebildeten Abschnitt
- Vorhandensein von Frakturen und/oder starken degenerativen Veränderungen

der Wirbelsäule, wenn dadurch die Integrität des betroffenen Wirbels beeinträchtigt war.

An dieser Stelle sollte erwähnt werden, dass das Vorliegen von Hämangiomen, Schmorlschen Knoten, Fremdmaterialien (z.B. bei Spondylodese / Kyphoplastie) sowie leichter oder moderater degenerativer Veränderungen nicht als Ausschlusskriterium betrachtet wurde (Liebl et al., 2021).

### 3.2.4 Beschreibung des definitiven Gesamtdatensatzes

Alle im Datensatz eingeschlossenen Scans wurden im Rahmen des Datenexports in das NIfTI-Format ("Neuroimaging Informatics Technology Initiative") umgewandelt, was unter anderem eine Anonymisierung der Daten ermöglichte (Liebl et al., 2021). Die von Glocker et al. hinzugezogenen Daten lagen bereits anonymisiert im NIfTI-Format vor (Glocker et al., 2013) (Vgl. Link). Der finale Datensatz enthält 319 CT-Scans von insgesamt 300 verschiedenen, anonymisierten Patienten, wobei ein ausgewogenes Geschlechterverhältnis zu verzeichnen ist (144 weibliche Patientinnen, 156 männliche Patienten.) (Liebl et al., 2021). Das Patientenalter ist mit einem Mittelwert von 56,2 Jahren und einer Standardabweichung von 17,6 Jahren beschrieben (Liebl et al., 2021). Die Gesamtanzahl vollständig abgebildeter Wirbel beträgt 4142 und lässt sich in 581 zervikale, 2255 thorakale und 1306 lumbale Wirbel untergliedern (Liebl et al., 2021). Ferner lässt sich der Datensatz nach den unterschiedlichen Scanner-Herstellern gliedern in 146 Scans vom hauseigenen Philips-Scanner, 73 Scans vom hauseigenen Siemens-Scanner, 20 externe Scans vom Hersteller GE Medical Systems, 30 externe Scans vom Hersteller Siemens, 20 externe Scans vom Hersteller Toshiba sowie 30 externe Scans aus dem Glocker-Datensatz, bei denen der Gerätehersteller unbekannt ist (Liebl et al., 2021). Für alle Angaben zum Patientenkollektiv siehe auch Tabelle 1. Im Sinne der Teilnahme an der VerSe – 20 Challenge wurden die Daten gleichmäßig auf drei Teildatensätze verteilt, wie im Kapitel 3.2.1 bereits erwähnt wurde. Der Datensatz ist öffentlich zugänglich und über den folgenden Link: (<https://osf.io/t98fz/>) einzusehen. Im Folgenden werden nun die resultierenden Teildatensätze genauer beschrieben werden. Die Verteilung der einzelnen identifizierten Subtypen anatomischer Varianten der Wirbelsäule auf die unterschiedlichen Teildatensätze ist in Tabelle 2 dargestellt.

### 3.2.5 Das "Training – Dataset"

Das Training-Dataset setzt sich aus 113 Scans von insgesamt 100 Patienten (50 weiblich, 50 männlich) mit einem durchschnittlichen (Mittelwert) Alter von 56,3 Jahren (SD = 18,0 Jahre) zusammen (vgl. Tabelle 1) (Liebl et al., 2021). Der gerätespezifisch neuselektierte Anteil beschränkt sich in diesem Teildatensatz auf 10 externe Scans vom Hersteller Siemens sowie 10 Scans aus dem Glocker-Datensatz von unbekanntem Geräte-Hersteller (vgl. Tabelle 1) (Liebl et al., 2021). Insgesamt beinhaltet sind 1428 vollständig abgebildete Wirbel, die sich unterteilen lassen in 224 zervikale, 757 thorakale und 447 lumbale Wirbel (vgl. Tabelle 1) (Liebl et al., 2021). Das Training-Dataset diente im Hinblick auf die VerSe – 20 Challenge zunächst nur der Präsentation einer gewissen Anzahl von Scans gegenüber den jeweiligen Algorithmen der Teilnehmer (vgl. Kapitel 3.3.1), um bei diesen, wie der Name vermuten lässt, einen Trainingseffekt zu erzielen (Sekuboyina et al., 2021). Die Ergebnisse dieses ersten Segmentierungsprozesses wurden nicht in den Auswertungsvorgang und die Ermittlung der Performance einbezogen (Sekuboyina et al., 2021).

### 3.2.6 Das "Public – Test – Dataset"

Demgegenüber besteht das Public-Test-Dataset aus 103 Scans von insgesamt 100 Patienten (49 weiblich, 51 männlich) mit einem durchschnittlichen (Mittelwert) Alter von 54,5 Jahren (SD = 17,2 Jahre) (vgl. Tabelle 1) (Liebl et al., 2021). Die Aufteilung des gerätespezifisch neuselektierten Anteils auf die verschiedenen Hersteller kann wie folgt angegeben werden: 10 externe Scans vom Hersteller Siemens, 10 externe Scans vom Hersteller Toshiba, 10 externe Scans vom Hersteller GE Medical Systems sowie 10 Scans aus dem Glocker-Datensatz von unbekanntem Geräte-Hersteller (vgl. Tabelle 1) (Liebl et al., 2021). Der Teildatensatz enthält eine Gesamtzahl von 1366 vollständig abgebildeten Wirbeln, davon 164 zervikale, 770 thorakale und 432 lumbale Wirbel (vgl. Tabelle 1) (Liebl et al., 2021). Das Public-Test-Dataset wurde in Form unsegmentierter Scans allen Teams zur Verfügung gestellt (vgl. Kapitel 3.3.1), woraufhin die verschiedenen Teams, die durch ihren jeweiligen Algorithmus erstellten Segmentierungen zum Zweck der Auswertung und Punktevergabe wiederum den anderen Teilnehmern zur Verfügung stellten

(Sekuboyina et al., 2021).

### 3.2.7 Das "Hidden – Test – Dataset"

Zuletzt ist das Hidden-Test-Dataset zu nennen, das 103 Scans von insgesamt 100 Patienten (45 weiblich, 55 männlich) mit einem durchschnittlichen (Mittelwert) Alter von 57,9 Jahren (SD = 17,6 Jahre) enthält (vgl. Tabelle 1) (Liebl et al., 2021). Der gerätespezifisch neuselektierte Anteil lässt sich untergliedern in 10 externe Scans vom Hersteller Siemens, 10 externe Scans vom Hersteller Toshiba, 10 externe Scans vom Hersteller GE Medical Systems sowie 10 Scans aus dem Glocker-Datensatz von unbekanntem Geräte-Hersteller (vgl. Tabelle 1) (Liebl et al., 2021). Die Gesamtzahl vollständig dargestellter Wirbel in diesem Teildatensatz beträgt 1348, davon 193 zervikale, 728 thorakale und 427 lumbale Wirbel (vgl. Tabelle 1) (Liebl et al., 2021).

Das Hidden-Test-Dataset spielte aufgrund seiner hohen Aussagekraft eine große Rolle bei der Auswertung und bei der Berechnung der Performance-Werte der einzelnen Algorithmen (Sekuboyina et al., 2021). In diesem letzten Durchgang stellten die verschiedenen Teams ihre Algorithmen einem sog. Docker Container zur Verfügung, sodass nach der Darreichung der unsegmentierten Scans aus dem Hidden-Test-Dataset abermals Segmentierungen prozessiert wurden, die allerdings keine nachträgliche Korrektur mehr erhalten konnten (Sekuboyina et al., 2021).



Tabelle 2: Der definitive Gesamtdatensatz (Liebl et al., 2021)

	Hidden Test Dataset	Public Test Dataset	Training Dataset	Gesamtdatensatz
<b>Halsrippen</b>				
1 Halsrippe (n)	3	5	4	12
2 Halsrippen (n)	2	2	2	6
<b>Thorakolumbale Varianten</b>				
12-gliedrige BWS (n)	44	44	43	131
11-gliedrige BWS (n)	1	4	3	8
13-gliedrige BWS (n)	2	2	2	6
<b>Lumbosakrale Varianten</b>				
5-gliedrige BWS (n)	52	54	59	165
4-gliedrige BWS (n)	1	2	0	3
6-gliedrige BWS (n)	29	38	28	85
<b>Stummelrippen</b>				
Stummelrippe bei Th12 (n)	8	12	17	37
Stummelrippe bei Th13 (n)	2	2	1	5
Stummelrippe bei L1 (n)	18	14	16	48
<b>Lumbosakraler Übergang nach Castellvi</b>				
Castellvi 0/1 (n)	46	52	48	146
Castellvi 2a (n)	10	8	17	35
Castellvi 2b (n)	10	8	7	25
Castellvi 3a (n)	5	7	2	14
Castellvi 3b (n)	8	8	12	28
Castellvi 4 (n)	3	2	2	7

## 3.3 Teilnahme an der VerSe – 20 Challenge

### 3.3.1 Allgemeine Informationen zur VerSe Challenge (2019 und 2020)

Die Large Scale Vertebral Segmentation Challenge 2020 (VerSe – 20) wurde, ebenso wie im Jahr zuvor die Large Scale Vertebral Segmentation Challenge 2019 (VerSe – 19) im Rahmen der MICCAI Konferenz (International Conference on Medical Image Computing and Computer Assisted Intervention) abgehalten (Liebl et al., 2021). Dabei nahmen in beiden Jahren mehrere internationale Teams an der Challenge teil, indem jeweils ein Segmentierungsalgorithmus pro Team vorgestellt wurde (Sekuboyina et al., 2021). Die Performance der einzelnen Algorithmen an den zur Verfügung gestellten Scans wurde anschließend quantifiziert (Sekuboyina et al., 2021). Wie in den Kapiteln 3.2.5 bis 3.2.7 bereits beschrieben, wurden im Zuge der Challenge jeweils eine Trainings- und zwei Testphasen durchlaufen (s.o.) (Sekuboyina et al., 2021). Für die Bestimmung der Performance eines Algorithmus gilt hierbei generell, dass die Teilperformance an den Scans des sog. Hidden-Test-Dataset die höchste Aussagekraft besitzt, da in diesem Teildatensatz keine nachträglichen Korrekturen an der erstellten Segmentierung möglich waren, und somit die Qualität der automatisierten Segmentierung am besten abgebildet wird (Sekuboyina et al., 2021). Eine Beschreibung der im Rahmen der Challenge verwendeten Parameter zur Quantifizierung der jeweiligen Performance ist dem Kapitel 3.3.4 zu entnehmen. Weitere Informationen zu den Teilnehmern und den verschiedenen Algorithmen beider VerSe Challenges (VerSe – 19 und VerSe – 20) können unter den folgenden Links eingesehen werden: VerSe – 19: <https://verse2019.grand-challenge.org/> und VerSe – 20: <https://verse2020.grand-challenge.org/>

### 3.3.2 Das Anduin-Tool

Die automatisierten Schritte in Labelling und Segmentierung von Wirbelsäulendarstellungen in den vorhandenen CT-Daten im Rahmen der Teilnahme an der Vertebral Segmentation Challenge 2020 wurden im Wesentlichen durch das von uns verwendete Framework Anduin geleistet, das unter dem folgenden Link:

[www.anduin.bonescreen.de](http://www.anduin.bonescreen.de) öffentlich zugänglich ist (Sekuboyina et al., 2021). Es wurden hierfür ausschließlich CT-Daten mit einer isotropischen Auflösung von 1 mm oder einer sagittalen Auflösung von max. 3 mm in den Datensatz eingeschlossen (Liebl et al., 2021). Die Segmentierung durch Anduin läuft in mehreren automatisierten Einzelschritten ab (Liebl et al., 2021; Sekuboyina et al., 2021). Mithilfe eines sog. CNN (Convolutional neural Network) wird zunächst eine sog. Heatmap erzeugt, d.h. eine niedrigaufgelöste visuelle Codierung aller interessierenden Strukturen (in diesem Fall knöcherne, zur Wirbelsäule gehörige Strukturen), die primär der Erstellung einer sog. Bounding Box dient, also eines abgegrenzten Teilbereichs, der alle o.g. Strukturen enthält (Liebl et al., 2021; Sekuboyina et al., 2021). In einem nächsten Schritt ermöglicht ein sog. Butterfly Net ein vorläufiges Labelling, indem auf dreidimensionalem Niveau die sog. Centroids, die als Mittel- und Ausgangspunkt des eigentlichen Segmentationsvorganges verstanden werden können, in der Mitte der jeweiligen Wirbelkörper platziert werden (Liebl et al., 2021; Sekuboyina et al., 2021). An dieser Stelle besteht die Möglichkeit einer manuellen Korrektur der genauen Lage der Centroids durch den Anwender des Anduin Tools (Sekuboyina et al., 2021). In einem letzten Schritt erfolgt mittels eines U-Net die hochaufgelöste, fein differenzierte Segmentierung der einzelnen Wirbel inklusive der sog. Subregions (Wirbelbogen, Wirbelfortsätze) für jeden Wirbel gesondert, sowie im Anschluss das Zusammenfügen der auf diese Weise erstellten Patches, sodass am Ende des Prozesses eine Segmentation Mask, d.h. eine farblich codierte, hochaufgelöste Segmentierung aller vollständig im bearbeiteten CT dargestellten Wirbel, steht (Liebl et al., 2021; Sekuboyina et al., 2021). Die anschließende manuelle Segmentierungskorrektur erfolgte mithilfe des Tools ITK-Snap, welches unter dem folgenden Link erreicht werden kann: (<http://www.itksnap.org/pmwiki/pmwiki.php>), durch drei eigens für diese Aufgabe trainierte Medizinstudenten, darunter ich selbst. Die Kontrolle der Segmentierungskorrekturen unterlag zwei Neuroradiologen. Hinsichtlich der Funktionsweisen der von den anderen Teams in der VerSe – 20 Challenge vorgestellten Segmentierungsalgorithmen, die sich in einigen Punkten grundlegend von der Funktionsweise von Anduin unterscheiden, muss auf die entsprechende Literatur verwiesen werden (Sekuboyina et al., 2021).

### 3.3.3 Anforderungen an eine gute Segmentierung

Sekuboyina et al. beschreiben die komplexen Herausforderungen, die sich für eine automatisierte Segmentierung der Wirbelsäule in CT-Daten ergeben (Sekuboyina et al., 2021). Schließlich zeigt nicht jedes CT genau denselben Abschnitt der Wirbelsäule, sodass alle Wirbel im dargestellten Abschnitt (= Field of view bzw. FoV) anhand morphologischer Merkmale korrekt identifiziert werden müssen. Ferner liegt auf der Hand, dass sowohl anatomische Varianten wie etwa Übergangsanomalien als auch diverse Pathologien der Wirbelsäule wie z.B. ossäre Metastasen, Wirbelkörperfrakturen, Schmorlsche Knoten, Hämangiome etc., aber auch Fremdmaterialien wie Metall o.ä. eine exakte Segmentierung erschweren können (Liebl et al., 2021). Ebenfalls zu erwähnen im Zusammenhang mit möglichen Herausforderungen in dieser Hinsicht sind die unterschiedlichen Gegebenheiten der Bildqualität je nach Hersteller des CT-Scanners (Liebl et al., 2021).

### 3.3.4 Parameter zur Bestimmung der Performance

Für die Bestimmung bzw. Quantifizierung der Korrektheit und Vollständigkeit eines Labelling- und Segmentierungsergebnisses standen in diesem Zusammenhang vier denkbare Parameter zur Verfügung, von denen zwei, nämlich die sog. ID-rate und die sog. Localisation distance (auch: d-mean) der Beurteilung des Labelling und zwei weitere, nämlich der sog. DICE-Score und die sog. Hausdorff distance zur Beurteilung der Segmentierung verwendet werden konnten (Sekuboyina et al., 2021). Nachdem bei der VerSe – 19 Challenge alle vier Parameter zum Einsatz gekommen waren, beschränkte man sich zum Zweck der Unverfälschtheit der Ergebnisse bei der VerSe – 20 Challenge lediglich auf die Verwendung der ID-rate (Labelling) und des DICE-Scores (Segmentierung) (Sekuboyina et al., 2021). Die beiden in der VerSe -20 nicht mehr verwendeten Parameter (d-mean, Hausdorff distance) sind nämlich ihrer Natur nach wenig sensitiv für das komplette Fehlen eines Labels (Sekuboyina et al., 2021). Dementsprechend handelt es sich bei ihnen um solche Parameter, die auch bei unvollständigem Labeling eine Einschätzung über die Restperformance hinsichtlich der richtigen Lage der restlichen, platzierten Label bzw. der Korrektheit der Segmentierung derer als solche erlaubt, während bei der ID-rate und dem DICE-Score ein solches Fehlen eines Wirbels stärker ins Gewicht fällt

(Sekuboyina et al., 2021). Aufgrund der Tatsache, dass im Datensatz der VerSe – 20 Challenge bewusst eine überproportional hohe Anzahl an Scans integriert worden war, die anatomische Varianten abbilden, lag es schließlich nahe, dass ein vollständiges, lückenloses Labelling als ein Kriterium von hoher Wichtigkeit für die Beurteilung der Performance angesehen werden musste, und dass daher die ID-rate und der DICE-Score die sinnvolleren Parameter darstellten (Sekuboyina et al., 2021).

#### i) ID-rate

Die ID-rate stellt ein Maß für die Korrektheit des Labelling dar, indem sie die Anzahl korrekt platzierter Labels bzw. Centroids mit der Gesamtanzahl der auf dem Scan abgebildeten Wirbel ins Verhältnis setzt (Sekuboyina et al., 2021). Dabei gilt ein Wirbel dann als korrekt identifiziert, wenn der Centroid erstens demjenigen Wirbel am nächsten ist, den er bezeichnet (also keinem anderen), und zweitens keinen größeren Abstand als 20 mm zur wahren Position aufweist, an der er sich befinden müsste (Sekuboyina et al., 2021). Für das bessere Verständnis der ID-rate siehe auch die untenstehende Formel:

$$ID - rate = \frac{\text{Anzahl korrekt identifizierter Wirbel}}{\text{Anzahl komplett abgebildeter Wirbel}}$$

#### ii) DICE-Score

Der DICE-Score dient, wie zuvor erwähnt, der Beurteilung der Segmentierung, und unterscheidet zwischen den beiden Termini der sog. Prediction (P), also dem Vorschlag des Algorithmus bezüglich der einem Wirbel zugehörig zu verstehenden Flächen in den einzelnen Schichten der Aufnahme, sowie der sog. Ground Truth (T), die als die wahre Gesamtheit der in allen Schichten enthaltenen Flächen betrachtet werden kann, die einem bestimmten Wirbel zuzuordnen sind (Sekuboyina et al., 2021). Vereinfachend lässt sich sagen, dass der DICE-Score ein Maß für die Überschneidung der beiden Mengen von P und T darstellt (Sekuboyina et al., 2021). Für das bessere Verständnis des DICE-Scores siehe auch die untenstehende Formel:

$$DICE - Score(P; T) = \frac{1}{N} \sum_{i=1}^N \left( \frac{2|P_i \cap T_i|}{|P_i| + |T_i|} \right)$$

## 4. Ergebnisse

### 4.1 Primäre Datenanalyse

#### 4.1.1 Allgemeine Informationen zur Datenanalyse

Nachdem der Gesamtdatensatz und zusätzlich auch alle weiteren im Vorgängerprojekt (VerSe19) enthaltenen Daten, teils nach Identifikation weiterer, akzidenteller Anomalien, schließlich mithilfe der erarbeiteten Klassifikationskriterien (vgl. Kapitel 3.1) auf den Aspekt anatomischer Varianten der Wirbelsäule hin klassifiziert worden waren, konnten im Folgenden einige für diese Arbeit essentielle Berechnungen angestellt werden, die nun in ihren Einzelheiten dargestellt werden sollen. Es sollte an dieser Stelle angemerkt werden, dass die hierfür verwendeten Daten aufgrund des Zeitpunktes der Durchführung der Berechnungen, sowie der Tatsache, dass im Anschluss daran noch geringfügige Anpassungen im Datensatz des VerSe20 Projektes vorgenommen worden sind (etwa Ein- oder Ausschluss einzelner Individuen), geringfügig von dem von Liebl et al. veröffentlichten Datensatz abweichen können. Die Abweichungen sind jedoch derartig minimal, dass sie für die Aussagekraft der Berechnungen als unerheblich angesehen werden können. Zudem beruhen die Berechnungen auf einer archivierten und bei Bedarf einsehbaren Datentabelle, sodass sie reproduzierbar bleiben. Die Einbeziehung derjenigen Daten des VerSe19 Datensatzes, die nicht Eingang in das VerSe20 Projekt erhalten hatten, wurde unabhängig von der Teilnahme an der diesjährigen VerSe Challenge vorgenommen, und erfüllt hauptsächlich methodische Zwecke für die zuvor erwähnten Berechnungen. Dabei wurden zunächst aus den vorhandenen Daten, die nicht variantenspezifisch erhoben wurden, Werte für die Grundprävalenz verschiedener anatomischer Varianten berechnet, und dann anschließend mithilfe des variantenspezifisch selektierten Anteils der Versuch unternommen, Zusammenhänge hinsichtlich der Koprävalenz der einzelnen Varianten untereinander aufzudecken. Um die Plausibilität der berechneten Größen einschätzen zu können, wurden sie mit den Prävalenzangaben anderer Autoren in Beziehung gesetzt, wie dem Kapitel 5 dieser Arbeit zu entnehmen ist. Für eine bessere Übersicht der Ergebnisse siehe auch die Tabellen 3 – 8.

## 4.1.2 Grundprävalenzen diverser anatomischer Varianten der Wirbelsäule

Die näherungsweise Ermittlung einer Grundprävalenz, welche die Verteilung der zu untersuchenden anatomischen Phänomene in der Allgemeinbevölkerung widerspiegeln soll, erfordert für die Durchführung der Berechnungen eine diesbezüglich gänzlich unselektierte Stichprobe, bei deren Erhebung das Vorliegen eben dieser Anomalien keine Rolle gespielt hat. Da dies auf den gerätespezifisch selektierten Anteil des VerSe20 Datensatzes sowie auf die Gesamtheit des VerSe19 Datensatzes zutrifft, wurden diese beiden Gruppen für die folgenden Berechnungen herangezogen. Da es jedoch für jede spezifische Fragestellung weitere Voraussetzungen zu erfüllen gilt, variiert die Anzahl betrachteter Scans je nach Fragestellung. So ist es beispielsweise naheliegend, dass bei der Ermittlung der Prävalenz thorakolumbalen Übergangswirbel nur solche Scans in die Berechnungen einbezogen werden können, die den thorakolumbalen Übergang auch tatsächlich abbilden, da ansonsten das Ergebnis verfälscht würde. Die jeweiligen Bedingungen für das betrachtete Kollektiv sind nachfolgend im Zusammenhang mit der entsprechenden Fragestellung genannt. Im Folgenden werden nun die Ergebnisse hinsichtlich der errechneten Prävalenzen für das Vorliegen thorakolumbalen und lumbosakralen Übergangswirbel, sowie für das Vorliegen einer numerischen Anomalie der LWS dargestellt. Auf die Berechnung der Prävalenzen für das Vorliegen einer numerischen Anomalie der BWS sowie das Vorliegen von Halsrippen wurde im Rahmen dieser Arbeit wegen der zu geringen Fallzahl verzichtet.

### a. Prävalenz thorakolumbalen Übergangswirbel

Zunächst soll an dieser Stelle genannt sein, dass das Vorliegen eines thorakolumbalen Übergangswirbels in unserem Datensatz übereinstimmend ist mit dem Vorliegen von Stummelrippen, d.h. Rippen einer Länge = 38 mm oder < 38 mm entweder uni- oder bilateral. Auch die Einteilung des Übergangswirbels nach thorakalem bzw. lumbalem Typ ist deckungsgleich mit dem Vorhandensein entsprechend thorakaler bzw. lumbaler Stummelrippen. Um nun eine Prävalenzberechnung vornehmen zu können, wurden als Stichprobe diejenigen Scans aus dem unselektierten Datenanteil (s.o.) verwendet, bei denen der

thorakolumbale Übergang vollständig abgebildet war, d.h. mindestens ein thorakaler und ein lumbaler Wirbel mussten im FoV enthalten sein. Entsprechend der genannten Kriterien ergab sich schließlich eine Anzahl von 165 Scans, die nun auf das Vorhandensein thorakolumbaler Übergangswirbel bzw. Stummelrippen überprüft wurden. Diese lagen vor bei insgesamt 36 Scans, sodass in diesem Fall die Prävalenz auf 21,82 % geschätzt werden kann. Von den 36 Fällen mit Übergangswirbeln ließen sich 20 (d.h. 55,56 %) dem lumbalen und 16 (d.h. 44,44 %) dem thorakalen Typ zuordnen.

#### b. Prävalenz lumbosakraler Übergangswirbel

Grundbedingung für die Stichprobe bei der Berechnung der Prävalenz lumbosakraler Übergangswirbel war, dass der lumbosakrale Übergang vollständig abgebildet ist, also zumindest die kaudalen Anteile der LWS sowie ansatzweise das Sakrum. Dies traf auf 185 Scans aus dem unselektierten Datenanteil (s.o.) zu. Für die nachfolgenden Berechnungen wurde aufgrund der schlechten Objektivierbarkeit für das Vorliegen eines Übergangswirbels vom Grad I nach Castellvi beschlossen, ausschließlich solche Wirbel als lumbosakrale Übergangswirbel zu werten, die nach der Castellvi-Klassifikation als Grad IIa, IIb, IIIa, IIIb oder IV kategorisierbar waren (vgl. Kapitel 3.1.5). Aus dem Grunddatensatz von 185 Scans wurden insgesamt 23 Fälle mit vorliegendem lumbosakralem Übergangswirbel identifiziert (davon nach Castellvi-Grad gegliedert 30,43 % Typ IIa, 39,13 % Typ IIb, 4,35 % Typ IIIa, 21,74 % Typ IIIb, 4,35 % Typ IV). Die Prävalenz lumbosakraler Übergangswirbel lässt sich demzufolge nach unseren Definitionen und Berechnungen auf 12,43 % schätzen.

#### c. Prävalenz numerischer Anomalien der LWS

Der Begriff numerischer Anomalien der LWS umfasst theoretisch betrachtet alle Fälle, bei denen die Anzahl an Wirbeln, die als lumbal kategorisiert werden, sei es regulär oder in Form eines thorakolumbalen oder lumbosakralen Übergangswirbels, ungleich fünf ist. Da in unserem Datensatz kein einziger Fall mit einer Anzahl lumbaler Wirbel  $< 4$  oder  $> 6$  beschrieben ist (Liebl et al., 2021), kann zumindest im Rahmen dieser Arbeit angenommen werden, dass sich die numerischen Anomalien der LWS in zwei Kategorien gliedern lassen, nämlich erstens die 4-Gliedrigkeit und



zweitens die 6-Gliedrigkeit der LWS. Um die Gliedrigkeit der LWS zweifelsfrei bestimmen zu können, mussten nun aus dem unselektierten Datenanteil (s.o.) diejenigen Fälle als Stichprobe extrahiert werden, welche die LWS komplett abbilden, sowie als abgrenzende Elemente Ansätze des Sakrum und mindestens ansatzweise der kranial an die LWS angrenzende letzte Brustwirbel. Die Bedingungen trafen auf 176 Scans zu, von denen 18 eine Abweichung von der 5-Gliedrigkeit der LWS zeigten (Dabei lag in 100 % der Fälle eine 6-Gliedrigkeit und in 0 % der Fälle eine 4-Gliedrigkeit der LWS vor). Die Prävalenz numerischer Anomalien der LWS beträgt demzufolge schätzungsweise 10,23 %.

#### 4.1.3 Koprävalenzen der einzelnen anatomischen Varianten

Nachdem nun Schätzungen für die Prävalenzen der einzelnen anatomischen Varianten der Wirbelsäule ermittelt worden waren, wurde in einem nächsten Schritt versucht, etwaige Assoziationen zwischen ihrem jeweiligen Auftreten zu untersuchen. Die zentrale Frage lautet nun also, ob die Wahrscheinlichkeit für das Vorliegen einer bestimmten Wirbelsäulen-anomalie im Vergleich zur in der nichtselektierten Bevölkerung vorherrschenden Grundwahrscheinlichkeit erhöht ist, wenn bereits eine andere Wirbelsäulen-anomalie bei dem betreffenden Individuum bekannt ist. Als jeweils die unselektierte Bevölkerung repräsentierende Referenzgrößen wurden die zuvor errechneten Schätzungen für die Grundprävalenzen der einzelnen anatomischen Varianten verwendet. Um nun derartige Assoziationen präzise untersuchen zu können, war es notwendig, anders als bei der Berechnung der Grundprävalenzen, Stichproben mit möglichst hohen Prävalenzwerten zu betrachten. Zu diesem Zweck wurde der gesamte VerSe20 Datensatz inklusive des speziell nach anatomischen Varianten selektierten Datenanteils für die nachfolgenden Berechnungen herangezogen. Ähnlich wie in Kapitel 4.1.2 beschrieben, gelten auch hier wieder für die einzelnen Fragestellung jeweils spezifische Anforderungen an die in den jeweiligen Stichproben verwendeten Scans.

i) 1. Fragestellung:

Häufigkeit thorakolumbalen ÜW bei vorliegendem lumbosakralem ÜW

Um zu untersuchen, wie hoch die Wahrscheinlichkeit eines Individuums ist, bei vorhandenem lumbosakralem Übergangswirbel zusätzlich über einen thorakolumbalen Übergangswirbel zu verfügen, musste zunächst aus dem Hochprävalenzkollektiv (s.o.) eine Stichprobe ausgewählt werden, deren beinhalten Scans die folgenden Strukturen abbilden:

- Den thorakolumbalen Übergang
- Die komplette LWS mit vollständiger Abbildung aller Lendenwirbel
- Den lumbosakralen Übergang
- Einen lumbosakralen Übergangswirbel vom Grad IIa, IIb, IIIa, IIIb oder IV nach der Castellvi-Klassifikation

Nach den o.g. Kriterien wurde eine insgesamt 91 Scans umfassende Stichprobe erstellt, die nachfolgend auf das Vorliegen thorakolumbalen Übergangswirbel, d.h. Stummelrippen am thorakolumbalen Übergang, untersucht wurde. Diese wurden in insgesamt 40 Scans identifiziert, wobei ein relativ ausgewogenes Verhältnis zwischen thorakalem und lumbalem Typ des thorakolumbalen Übergangswirbels verzeichnet werden kann (65,0 % lumbaler Typ, 35,0 % thorakaler Typ). Eine mögliche Tendenz für das Vorherrschen des lumbalen Typs bei fortschreitendem Sakralisierungsgrad des lumbosakralen Übergangswirbels kann aufgrund der Ergebnisse diskutiert werden. Für Einzelheiten siehe Tabelle 3. Abschließend kann unseren Berechnungen zufolge geschätzt werden, dass für ein Individuum mit vorhandenem lumbosakralem Übergangswirbel eine Wahrscheinlichkeit von 43,96 % besteht, zusätzlich einen thorakolumbalen Übergangswirbel aufzuweisen. Diese Wahrscheinlichkeit ist im Vergleich zur in Kapitel 4.1.2 genannten Grundwahrscheinlichkeit für das Vorliegen eines thorakolumbalen Übergangswirbels um mehr als den Faktor 2,0 erhöht.

Tabelle 3: Häufigkeit thorakolumbalen ÜW bei vorliegendem lumbosakralem ÜW

Castellvi-Grad	2a	2b	3a	3b	4	Alle
Anzahl Scans gesamt (n)	32	25	10	18	6	91
Anzahl Scans mit vorhandenem thorakolumb. ÜW, lumbaler Typ (n)	2 (6,3 %)	8 (32,0 %)	6 (60,0 %)	7 (38,9 %)	3 (50,0 %)	26 (28,6 %)
Anzahl Scans mit vorhandenem thorakolumb. ÜW, thorakaler Typ (n)	9 (28,1 %)	4 (16,0 %)	0 (0,0 %)	1 (5,6 %)	0 (0,0 %)	14 (15,4 %)
Anzahl Scans mit vorhandenem thorakolumb. ÜW, beide Typen (n)	11 (34,4 %)	12 (48,0 %)	6 (60,0 %)	8 (44,4 %)	3 (50,0 %)	40 (44,0 %)

ii) 2. Fragestellung:

Häufigkeit num. Anomalien der LWS bei vorliegendem lumbosakralem ÜW

Analog zu Fragestellung 1 soll als nächstes gezeigt werden, wie hoch die Wahrscheinlichkeit für ein Individuum mit vorhandenem lumbosakralem Übergangswirbel ist, zusätzlich eine numerische Anomalie der Wirbelsäule aufzuweisen. Die Scans in der für diese Fragestellung verwendeten Stichprobe müssen die folgenden Strukturen abbilden:

- Die komplette LWS mit vollständiger Abbildung aller Lendenwirbel
- Den lumbosakralen Übergang
- Einen lumbosakralen Übergangswirbel vom Grad IIa, IIb, IIIa, IIIb

oder IV nach der Castellvi-Klassifikation

Es ergab sich eine Stichprobe mit einer Anzahl von 91 Scans, von denen insgesamt 63 eine Abweichung von der 5-Gliedrigkeit der LWS aufwiesen. Interessant ist, dass in all diesen Fällen ausschließlich eine 6-Gliedrigkeit, und in keinem einzigen Fall eine 4-Gliedrigkeit der LWS festgestellt werden konnte. Einzelheiten sind der Tabelle 4 zu entnehmen. Aufgrund der Ergebnisse kann geschätzt werden, dass ein Individuum mit vorhandenem lumbosakralem Übergangswirbel mit einer Wahrscheinlichkeit von 69,23 % zusätzlich über eine numerische Anomalie der LWS verfügt. Diese Wahrscheinlichkeit ist im Vergleich zur in Kapitel 4.1.2 genannten Grundwahrscheinlichkeit für das Vorliegen einer numerischen Aberration der LWS um mehr als den Faktor 6,7 erhöht.

Tabelle 4: Häufigkeit num. Anomalien der LWS bei vorliegendem lumbosakralem ÜW

Castellvi-Grad	2a	2b	3a	3b	4	Alle
Anzahl Scans gesamt (n)	32	25	10	18	6	91
Anzahl Scans mit 4-gliedriger LWS (n)	0	0	0	0	0	0
Anzahl Scans mit 6-gliedriger LWS (n)	15 (46,9 %)	15 (60,0 %)	10 (100 %)	17 (94,4 %)	6 (100 %)	63 (69,2 %)
Anzahl Scans mit numerischer Aberration der LWS allgemein (n)	15 (46,9 %)	15 (60,0 %)	10 (100 %)	17 (94,4 %)	6 (100 %)	63 (69,2 %)

iii) 3. Fragestellung:

Häufigkeit lumbosakraler ÜW bei vorliegendem thorakolumbalem ÜW

Die 3. Fragestellung ähnelt der 1. Fragestellung inhaltlich, jedoch soll nun gezeigt werden, mit welcher Wahrscheinlichkeit bei vorliegendem thorakolumbalem Übergangswirbel zusätzlich ein lumbosakraler Übergangswirbel vorliegt. Es wurde

abermals eine Stichprobe aus dem Hochprävalenzkollektiv (s.o.) erstellt. Die darin enthaltenen Scans müssen die folgenden Strukturen abbilden:

- Den thorakolumbalen Übergang
- Die komplette LWS mit vollständiger Abbildung aller Lendenwirbel
- Den lumbosakralen Übergang
- Einen thorakolumbalen ÜW mit Stummelrippen

Von den insgesamt 73 Scans, welche gemäß den oben erwähnten Bedingungen die Stichprobe bildeten, wiesen 40 Scans einen lumbosakralen Übergangswirbel auf, der sich nach der Castellvi-Klassifikation als Grad IIa, IIb, IIIa, IIIb oder IV kategorisieren ließ. Dabei zeigte sich, dass lumbosakrale ÜW generell etwas häufiger bei thorakolumbalen ÜW vom lumbalen Typ, und etwas seltener bei solchen vom thorakalen Typ vorlagen. Einzelheiten hierzu können der Tabelle 5 entnommen werden. Es ergab sich nach unseren Berechnungen eine geschätzte Wahrscheinlichkeit von 54,79 %, dass ein Individuum mit vorhandenem thorakolumbalem Übergangswirbel zusätzlich über einen lumbosakralen Übergangswirbel verfügt. Diese Wahrscheinlichkeit ist im Vergleich zur in Kapitel 4.1.2 genannten Grundwahrscheinlichkeit für das Vorliegen eines lumbosakralen Übergangswirbels um mehr als den Faktor 4,4 erhöht.

Tabelle 5: Häufigkeit lumbosakraler ÜW bei vorliegendem thorakolumbalem ÜW

Thorakolumbaler ÜW	Lumbaler Typ	Thorakaler Typ	Alle
Anzahl Scans gesamt (n)	43	30	73
Anzahl Scans mit zusätzlich Castellvi 2a (n)	2 (4,7 %)	9 (30,0 %)	11 (15,1 %)
Anzahl Scans mit zusätzlich Castellvi 2b (n)	8 (18,6 %)	4 (13,3 %)	12 (16,4 %)
Anzahl Scans mit zusätzlich Castellvi 3a (n)	6 (14,0 %)	0 (0,0 %)	6 (8,2 %)
Anzahl Scans mit zusätzlich Castellvi 3b (n)	7 (16,3 %)	1 (3,3 %)	8 (11,0 %)
Anzahl Scans mit zusätzlich Castellvi 4 (n)	3 (7,0 %)	0 (0,0 %)	3 (4,1 %)
Anzahl Scans mit zusätzlichem lumbosakralem ÜW allgemein (n)	26 (60,5 %)	14 (46,7 %)	40 (54,8 %)

iv) 4. Fragestellung:

Häufigkeit num. Anomalien der LWS bei vorliegendem thorakolumbalem ÜW

Wie bei der 3. Fragestellung wird auch in der 4. Fragestellung von einem hypothetischen Individuum mit vorliegendem thorakolumbalem Übergangswirbel ausgegangen. Nun soll dessen Wahrscheinlichkeit für das gleichzeitige Vorhandensein einer numerischen Anomalie der LWS berechnet werden. Als Grundlage für die Berechnung dient erneut eine Stichprobe aus dem Hochprävalenzkollektiv (s.o.), welche die Anforderung erfüllen muss, dass die darin enthaltenen Scans die folgenden Strukturen abbilden:

- Den thorakolumbalen Übergang
- Die komplette LWS mit vollständiger Abbildung aller Lendenwirbel
- Den lumbosakralen Übergang
- Einen thorakolumbalen ÜW mit Stummelrippen

Es wurde auf diese Weise analog zur vorherigen Fragestellung eine 73 Scans beinhaltende Stichprobe erhoben, die 35 Fälle mit vorhandener numerischer Anomalie der LWS aufwies. In diesem Fall wurden unter den betreffenden Scans fast ausschließlich 6-gliedrige, und lediglich eine einzige 4-gliedrige LWS verzeichnet. Die allgemeine Wahrscheinlichkeit, bei vorhandenem thorakolumbalem Übergangswirbel zusätzlich über eine numerische Anomalie der LWS zu verfügen, lässt sich mithilfe der erhobenen Zahlen nun auf 47,95 % schätzen, obgleich die separaten Berechnungen für die einzelnen Subtypen der thorakolumbalen Übergangswirbel interessanterweise eine geschätzte Wahrscheinlichkeit von nur 3,33 % für das Auftreten einer numerischen Anomalie der LWS bei Vorliegen des thorakalen Subtyps, und demgegenüber eine Wahrscheinlichkeit von 79,07 % bei Vorliegen des lumbalen Subtyps ergaben. Einzelheiten können in der Tabelle 6 eingesehen werden. Die zuvor erwähnte, allgemeine Wahrscheinlichkeit ist im Vergleich zur in Kapitel 4.1.2 genannten Grundwahrscheinlichkeit für das Vorliegen einer numerischen Aberration der LWS um mehr als den Faktor 4,6 erhöht.

*Tabelle 6: Häufigkeit num. Anomalien der LWS bei vorliegendem thorakolumbalem ÜW*

Thorakolumbaler ÜW	Lumbaler Typ	Thorakaler Typ	Alle
Anzahl Scans gesamt (n)	43	30	73
Anzahl Scans mit 4-gliedriger LWS (n)	1 (2,3 %)	0 (0,0 %)	1 (1,4 %)
Anzahl Scans mit 6-gliedriger LWS (n)	33 (76,7 %)	1 (3,3 %)	34 (46,6 %)
Anzahl Scans mit numerischer Aberration der LWS allgemein (n)	34 (79,1 %)	1 (3,3 %)	35 (47,9 %)

v) 5. Fragestellung:

Häufigkeit lumbosakraler ÜW bei vorliegender num. Anomalie der LWS

Die letzten beiden Fragestellungen beziehen sich auf die Situation, dass die Wahrscheinlichkeit für ein hypothetisches Individuum mit bestehender numerischer Anomalie der LWS hinsichtlich des Vorliegens weiterer anatomischer Varianten der Wirbelsäule untersucht werden soll. Zunächst soll hierbei die Frage geklärt werden, mit welcher Wahrscheinlichkeit zusätzlich ein lumbosakraler Übergangswirbel vorliegt. Die aus dem Hochprävalenzkollektiv (s.o.) abgeleitete Stichprobe enthält ausschließlich Scans, die im Sinne der nachfolgenden Berechnungen die folgenden anatomischen Strukturen abbilden:

- Die komplette, in diesem Fall entweder 4- oder 6-gliedrige LWS
- Den lumbosakralen Übergang

Es soll an dieser Stelle vermerkt sein, dass das für die vorherigen Berechnungen angewandte Grundkriterium der vollständigen Abbildung sämtlicher Lendenwirbel für die letzten beiden Berechnungen insofern aufgeweicht wurde, als ein Fall auch dann in die Ausgangszahl  $n$  aufgenommen werden konnte, wenn bei ansonsten zweifelsfrei bestimmbarer Gliedrigkeit der LWS einzelne Lendenwirbel nur teilweise abgebildet waren. Dies diente der Erhöhung der ansonsten sehr geringen Fallzahl, und lässt sich dadurch rechtfertigen, dass das besagte Grundkriterium hier von geringerer Bedeutung war als bei den vorangegangenen Fragestellungen. Die gebildete Stichprobe setzt sich aus insgesamt 88 Scans zusammen (davon 85 Fälle mit 6-gliedriger LWS und lediglich 3 Fälle mit 4-gliedriger LWS), von denen 66 Scans zusätzlich einen lumbosakralen Übergangswirbel der Kategorie IIa, IIb, IIIa, IIIb oder IV nach der Castellvi-Klassifikation zeigten. Bei all diesen Fällen handelte es sich um Scans mit vorhandener 6-Gliedrigkeit der LWS. Die wenigen Fälle mit 4-Gliedrigkeit der LWS zeigten jeweils keinen lumbosakralen Übergangswirbel. Einzelheiten sind der Tabelle 7 zu entnehmen. Die geschätzte Wahrscheinlichkeit für das Vorliegen eines lumbosakralen Übergangswirbels bei vorhandener numerischer Anomalie der LWS kann mit 75,0 % angegeben werden. Diese Wahrscheinlichkeit ist im Vergleich zur in Kapitel 4.1.2 genannten Grundwahrscheinlichkeit für das Vorliegen eines lumbosakralen Übergangswirbels um mehr als den Faktor 6,0 erhöht.



Tabelle 7: Häufigkeit lumbosakraler ÜW bei vorliegender num. Anomalie der LWS

Num. Aberration der LWS	6-Gliedrigkeit	4-Gliedrigkeit	Alle
Anzahl Scans gesamt (n)	85	3	88
Anzahl Scans mit zusätzlich Castellvi 2a (n)	16 (18,8 %)	0 (0,0 %)	16 (18,2 %)
Anzahl Scans mit zusätzlich Castellvi 2b (n)	15 (17,6 %)	0 (0,0 %)	15 (17,0 %)
Anzahl Scans mit zusätzlich Castellvi 3a (n)	11 (12,9 %)	0 (0,0 %)	11 (12,5 %)
Anzahl Scans mit zusätzlich Castellvi 3b (n)	18 (21,2 %)	0 (0,0 %)	18 (20,5 %)
Anzahl Scans mit zusätzlich Castellvi 4 (n)	6 (7,1 %)	0 (0,0 %)	6 (6,8 %)
Anzahl Scans mit zusätzlichem lumbosakralem ÜW allgemein (n)	66 (77,6 %)	0 (0,0 %)	66 (75,0 %)

vi) 6. Fragestellung:

Häufigkeit thorakolumbalen ÜW bei vorliegender num. Anomalie der LWS

Zuletzt soll nun die Wahrscheinlichkeit für das Vorliegen eines thorakolumbalen Übergangswirbels bei vorhandener numerischer Anomalie der LWS geschätzt werden. Die für die Berechnungen verwendete Stichprobe soll daher ausschließlich Scans enthalten, welche die folgenden Strukturen abbilden:

- Den thorakolumbalen Übergang
- Die komplette, in diesem Fall entweder 4- oder 6-gliedrige LWS
- Den lumbosakralen Übergang

Die auf diese Weise aus dem Hochprävalenzkollektiv (s.o.) abgeleitete Stichprobe

enthielt insgesamt 84 Scans (darunter 81 Fälle mit 6-gliedriger LWS und 3 Fälle mit 4-gliedriger LWS), von denen bei 35 Scans Stummelrippen, d.h. thorakolumbale Übergangswirbel, identifizierbar waren. Es zeigte sich, dass beinahe ausschließlich die Fälle mit 6-gliedriger LWS, und lediglich ein einziger Fall mit 4-gliedriger LWS, einen Übergangswirbel aufwiesen. Somit ergibt sich bei Vorliegen einer numerischen Anomalie der LWS eine geschätzte Wahrscheinlichkeit von 41,67 %, dass zusätzlich ein thorakolumbaler Übergangswirbel vorliegt. Diese Wahrscheinlichkeit ist im Vergleich zur in Kapitel 4.1.2 genannten Grundwahrscheinlichkeit für das Vorliegen eines thorakolumbalen Übergangswirbels um mehr als den Faktor 1,9 erhöht. Für Einzelheiten siehe auch Tabelle 8.

*Tabelle 8: Häufigkeit thorakolumbaler ÜW bei vorliegender numerischer Anomalie der LWS*

Num. Aberration der LWS	6-Gliedrigkeit	4-Gliedrigkeit	Alle
Anzahl Scans gesamt (n)	81	3	84
Anzahl Scans mit vorhandenem thorakolumb. ÜW, lumbaler Typ (n)	33 (40,7 %)	1 (33,3 %)	34 (40,5 %)
Anzahl Scans mit vorhandenem thorakolumb. ÜW, thorakaler Typ (n)	1 (1,2 %)	0 (0,0 %)	1 (1,2 %)
Anzahl Scans mit vorhandenem thorakolumb. ÜW, beide Typen (n)	34 (42,0 %)	1 (33,3 %)	35 (41,7 %)

## 4.2 Resultate der Verse – 20 Challenge

### 4.2.1 Gegenüberstellung von VerSe – 19 und VerSe – 20

Wie in Kapitel 3.3.4 beschrieben, wurden die durch die verschiedenen Algorithmen erstellten Segmentierungen letztlich mithilfe geeigneter Parameter ausgewertet (Sekuboyina et al., 2021). In der VerSe – 20 Challenge kamen hierbei ausschließlich die sog. ID-rate sowie der sog. DICE-Score zum Einsatz (Sekuboyina et al., 2021). Aufgrund der besseren Vergleichbarkeit möchte ich mich im Folgenden auch bei der Darstellung der Performance der Algorithmen im Rahmen der VerSe – 19 Challenge auf diese beiden Parameter beschränken. Nachdem jeder einzelne Scan für beide Parameter jeweils eine Art Punktzahl erhalten hat, die sich gemäß der in Kapitel 3.3.4 gezeigten Formeln errechnet, konnte durch Mittelwertbildung aus den Ergebnissen der einzelnen Scans wiederum eine durchschnittliche Performance für den entsprechenden Algorithmus ermittelt werden (Sekuboyina et al., 2021). Diese wird im Folgenden für das Public-Test-Dataset und das Hidden-Test-Dataset separat angegeben werden. Die ID-rate stellt dabei den für die Bewertung des Labelling entscheidenden Wert dar, während der DICE-Score im Wesentlichen die Beurteilung der Segmentierung gewährleisten soll (Sekuboyina et al., 2021). Die folgenden Angaben zur Platzierung der einzelnen Teams in der Rangliste mag auf den ersten Blick wenig intuitiv erscheinen, wenn die die Performance beschreibenden Einzelwerte gleichberechtigt betrachtet werden. Wie zuvor bereits erwähnt, ist es daher für das Verständnis des Zustandekommens der Platzierungen wichtig, dass die auf das Hidden-Test-Dataset bezogenen Werte von höherem Gewicht sind als diejenigen, die sich auf das Public-Test-Dataset beziehen (Sekuboyina et al., 2021). Für die exakten Regelungen hinsichtlich der Errechnung der Punktzahl, die lediglich der Erstellung einer Rangliste dient, muss auf die Originalliteratur verwiesen werden, da diese keinen Mehrwert für die spezifischen Fragestellungen dieser Arbeit beinhalten, und ferner in ihrer Komplexität hier nicht dargestellt werden können (Sekuboyina et al., 2021).

#### a. Resultate der VerSe – 19 Challenge

Die drei Teams mit der besten Bewertung in Bezug auf die durchschnittliche

Performance bei der VerSe – 19 Challenge waren Payer C. et al. an erster Stelle, darauffolgend Lessmann N. et al., und auf dem dritten Platz Chen M. et al. (Lessmann et al., 2019; Payer et al., 2019; Sekuboyina et al., 2021). Payer C. et al konnten mit ihrem Algorithmus eine ID-rate von 95,65 (Public-Test-Dataset) und 94,25 (Hidden-Test-Dataset), sowie einen DICE-Score von 90,90 (Public-Test-Dataset) und 89,80 (Hidden-Test-Dataset) erzielen, während Lessmann N. et al eine ID-rate von 89,86 (Public-Test-Dataset) und 90,42 (Hidden-Test-Dataset) und einen DICE-Score von 85,08 (Public-Test-Dataset) und 85,76 (Hidden-Test-Dataset) erreichten (Lessmann et al., 2019; Payer et al., 2019; Sekuboyina et al., 2021). Chen M. et al reihten sich anschließend mit einer ID-rate von 96,94 (Public-Test-Dataset) und 86,73 (Hidden-Test-Dataset), sowie einem DICE-Score von 93,01 (Public-Test-Dataset) und 82,56 (Hidden-Test-Dataset) ein (Sekuboyina et al., 2021). An dieser Stelle soll auch das Team von Sekuboyina et al., die mit dem im Rahmen dieses Projektes von uns entwickelten und trainierten Algorithmus Anduin gearbeitet haben, vorgestellt werden (Sekuboyina et al., 2021). Sekuboyina et al. erreichten bei der VerSe – 19 Challenge eine ID-rate von 89,97 (Public-Test-Dataset) und 87,66 (Hidden-Test-Dataset), und einen DICE-Score von 83,06 (Public-Test-Dataset) und 83,18 (Hidden-Test-Dataset) (Sekuboyina et al., 2021). Um die Ergebnisse der VerSe – 19 Challenge abschließend zusammenfassen und den Ergebnissen der VerSe – 20 Challenge gegenüber stellen zu können, eignet es sich, zusätzlich auch das durchschnittliche Abschneiden der fünf besten Teams (Top-5), sowie die durchschnittliche Performance aller zusammengefassten Teams (Alle) zu betrachten. Die Top-5 wiesen demnach in der VerSe – 19 Challenge eine durchschnittliche ID-rate von 83,3 (Public-Test-Dataset) und 82,4 (Hidden-Test-Dataset) und einen durchschnittlichen DICE-Score von 82,5 (Public-Test-Dataset) und 78,9 (Hidden-Test-Dataset) auf (Sekuboyina et al., 2021). Die durchschnittliche Performance aller Teilnehmer lässt sich angeben mit einer ID-rate von 61,4 (Public-Test-Dataset) und 61,6 (Hidden-Test-Dataset), und einem DICE-Score von 71,2 (Public-Test-Dataset) und 71,3 (Hidden-Test-Dataset) (Sekuboyina et al., 2021). Eine übersichtliche Darstellung aller o.g. Resultate kann in Tabelle 9 eingesehen werden.

Tabelle 9: Ergebnisse der VerSe – 19 Challenge (Sekuboyina et al., 2021)

Team	ID-rate		DICE-Score	
	Public Test Dataset	Hidden Test Dataset	Public Test Dataset	Hidden Test Dataset
Payer C. et al (1. Platz)	95,65	94,25	90,9	89,8
Lessmann N. et al (2. Platz)	89,86	90,42	85,08	85,76
Chen M. et al (3. Platz)	96,94	86,73	93,01	82,56
Sekuboyina A. et al (Anduin)	89,97	87,66	83,06	83,18
Top – 5	83,3	82,4	82,5	78,9
Alle	61,4	61,6	71,2	71,3

#### b. Resultate der VerSe – 20 Challenge

Analog zu den Angaben zur VerSe – 19 Challenge werden nun auch im Bezug auf die VerSe – 20 Challenge zunächst die drei Teams mit den besten Werten hinsichtlich ihrer Performance vorgestellt werden. Den ersten Platz belegten dabei Chen D. et al., mit einer ID-rate von 95,61 (Public-Test-Dataset) und 96,58 (Hidden-Test-Dataset), sowie einem DICE-Score von 91,72 (Public-Test-Dataset) und 91,23 (Hidden-Test-Dataset) (Sekuboyina et al., 2021). An zweiter Stelle in der Rangliste befinden sich Payer C. et al., die in diesem Jahr eine ID-rate von 95,06 (Public-Test-Dataset) und 92,82 (Hidden-Test-Dataset) und einen DICE-Score von 91,65 (Public-Test-Dataset) und 89,71 (Hidden-Test-Dataset) erzielen konnten (Sekuboyina et al., 2021). An dritter Stelle sind Zhang A. et al zu nennen, mit einer ID-rate von 94,93

(Public-Test-Dataset) und 96,22 (Hidden-Test-Dataset), und einem DICE-Score von 88,82 (Public-Test-Dataset) und 89,36 (Hidden-Test-Dataset) (Sekuboyina et al., 2021). Sekuboyina et al., die unsere Arbeitsgruppe repräsentieren, erreichten eine ID-rate von 82,68 (Public-Test-Dataset) und 86,06 (Hidden-Test-Dataset), und einen DICE-Score von 78,05 (Public-Test-Dataset) und 79,52 (Hidden-Test-Dataset) (Sekuboyina et al., 2021). Als nächstes werden nun erneut zum Zweck der besseren Vergleichbarkeit der beiden Challenges die durchschnittlichen, jeweils zusammengefassten Resultate der fünf besten Teams (Top-5) und aller Teams (Alle) dargestellt werden. Die Top-5 weisen eine durchschnittliche ID-rate von 93,9 (Public-Test-Dataset) und 94,4 (Hidden-Test-Dataset) auf, und einen DICE-Score von 89,3 (Public-Test-Dataset) und 88,8 (Hidden-Test-Dataset), während die durchschnittliche ID-rate aller Teams 77,7 (Public-Test-Dataset) und 72,8 (Hidden-Test-Dataset) beträgt, und der durchschnittliche DICE-Score aller Teams sich mit 80,3 (Public-Test-Dataset) und 74,2 (Hidden-Test-Dataset) angeben lässt (Sekuboyina et al., 2021). Alle genannten Ergebnisse sind in Tabelle 10 im Sinne der besseren Übersicht dargestellt.

Tabelle 10: Ergebnisse der VerSe – 20 Challenge (Sekuboyina et al., 2021)

Team	ID-rate		DICE-Score	
	Public Test Dataset	Hidden Test Dataset	Public Test Dataset	Hidden Test Dataset
Chen D. et al (1. Platz)	95,61	96,58	91,72	91,23
Payer C. et al (2. Platz)	95,06	92,82	91,65	89,71
Zhang A. et al (3. Platz)	94,93	96,22	88,82	89,36
Sekuboyina A. et al (Anduin)	82,68	86,06	78,05	79,52
Top – 5	93,9	94,4	89,3	88,8
Alle	77,7	72,8	80,3	74,2

### c. Vergleich der Resultate von VerSe – 19 und VerSe – 20

Eine Gegenüberstellung der von den verschiedenen Teams erzielten Performance-Werte in den beiden Challenges dient primär der Beantwortung der Frage, ob das allgemeine Leistungsniveau automatisierter Segmentierungsalgorithmen zum Zeitpunkt der zweiten Challenge sichtbar angestiegen ist im Vergleich zum Vorjahr. Dieser Vergleich kann am besten gezogen werden, wenn man die zuvor erwähnten durchschnittlichen Werte der Top-5 sowie die durchschnittlichen Werte aller teilnehmenden Teams, jeweils bezogen auf die ID-rate und den DICE-Score, für VerSe – 19 und VerSe – 20 betrachtet und einander gegenüberstellt. Die interessierenden Werte wurden zuvor bereits genannt und sind in Tabelle 11 nochmals übersichtlich dargestellt. In dem direkten Vergleich ist für die VerSe – 20 Challenge eine deutliche Verbesserung der durchschnittlichen Performance sowohl aller Teams als auch der Top-5 im Vergleich zur VerSe – 19 Challenge zu verzeichnen. Während die durchschnittlichen Werte der ID-rate für beide Gruppen sowohl im Public-Test-Dataset als auch im Hidden-Test-Dataset ausnahmslos um mindestens 10 Punkte angestiegen sind, lässt sich, wenn auch in etwas geringerer Ausprägung, eine ähnliche Tendenz für den DICE-Score feststellen. Kein einziger der betrachteten Werte hat sich im Vergleich zur VerSe – 19 Challenge verschlechtert. Sekuboyina et al. begründen diesen allgemeinen Anstieg der Performance mit der im Rahmen der VerSe – 20 Challenge vorgenommenen Adressierung anatomischer Varianten der Wirbelsäule, deren Labelling und Segmentierung mithilfe des überrepräsentativ großen Anteils an entsprechenden Scans entscheidend verbessert werden konnten (Sekuboyina et al., 2021).

Tabelle 11: Vergleich der Performance von VerSe – 19 und VerSe – 20 (Sekuboyina et al., 2021)

		Public Test Dataset		Hidden Test Dataset	
		Alle	Top – 5	Alle	Top – 5
ID – rate	VerSe – 19	61,4	83,3	61,6	82,4
	VerSe – 20	77,7	93,9	72,8	94,4
DICE – Score	VerSe – 19	71,2	82,5	71,3	78,9
	VerSe – 20	80,3	89,3	74,2	88,8

#### 4.2.2 Bedeutung anatomischer Varianten für die VerSe

Wie in Kapitel 4.2.1 bereits erwähnt, ist laut Sekuboyina et al. vor allem die bessere Erkennung anatomischer Varianten der Wirbelsäule durch die Segmentierungsalgorithmen für den starken Leistungsanstieg Letzterer seit der VerSe – 19 Challenge verantwortlich (Sekuboyina et al., 2021). Dennoch – oder gerade deswegen – ist nun die Frage danach, wie zuverlässig die besagten anatomischen Varianten denn aktuell durch die Algorithmen erkannt bzw. korrekt gelabelt und segmentiert werden, von großem Interesse. Damit unweigerlich verbunden ist auch die Frage, ob diesbezüglich noch weiteres Steigerungspotential besteht, das ggf. zum Inhalt etwaiger Folgeprojekte werden könnte. Um sich dieser Frage anzunähern, wird nun im Folgenden ein Augenmerk auf die separate Performance der verschiedenen Algorithmen an denjenigen Scans gelegt werden, die anatomische Varianten der Wirbelsäule beinhalten. Der Datensatz der VerSe – 20 Challenge wird also im Sinne dieser Betrachtungen aufgeteilt in einen ersten Anteil, der ausschließlich normale Scans enthält (ohne vorhandene anatomische Varianten) und einen zweiten Anteil, der ausschließlich aus Scans mit vorhandenen anatomischen Varianten besteht. Für beide Untergruppen kann nun separat ein eigener DICE-Score ermittelt werden, sodass bei Betrachtung der beiden Werte ein direkter Vergleich hinsichtlich der Performance gezogen werden kann. Beispielhaft



werden nun, analog zu Kapitel 4.2.1, zunächst die DICE-Scores der drei Teams mit der besten Performance beschrieben werden, sowie anschließend die Werte von Sekuboyina et al., die Durchschnittswerte der Top-5, und die Durchschnittswerte aller Teilnehmer. Eine Übersicht über die DICE-Scores der einzelnen Teams, sowie der Durchschnittswerte für die Top-5 und die gesamte Teilnehmerschaft, bietet Tabelle 12. Chen D. et al., die bei der VerSe – 20 Challenge das beste Ergebnis erzielt haben, weisen hier für die normalen Scans einen DICE-Score von 93,42 (Public-Test-Dataset) und 92,19 (Hidden-Test-Dataset) auf, und für die Varianten-Scans einen DICE-Score von 84,21 (Public-Test-Dataset) und 87,02 (Hidden-Test-Dataset) (Sekuboyina et al., 2021). Bei Payer C. et al., dem Team mit der insgesamt zweitbesten Performance, betragen die ermittelten Werte für den DICE-Score für die normalen Scans 92,94 (Public-Test-Dataset) und 89,74 (Hidden-Test-Dataset), und für die Varianten-Scans 85,96 (Public-Test-Dataset) und 89,59 (Hidden-Test-Dataset) (Sekuboyina et al., 2021). Das Team mit der drittbesten Performance, Zhang A. et al., weist einen DICE-Score von 89,20 (Public-Test-Dataset) und 89,81 (Hidden-Test-Dataset) für die normalen Scans, sowie einen DICE-Score von 87,15 (Public-Test-Dataset) und 87,35 (Hidden-Test-Dataset) für die Varianten-Scans auf (Sekuboyina et al., 2021). Betrachtet man nun das Ergebnis von Sekuboyina et al. bzw. dem in Kapitel 3.3.2 beschriebenen Segmentierungsalgorithmus Anduin, so lassen sich hier für die normalen Scans ein DICE-Score von 81,14 (Public-Test-Dataset) und 82,38 (Hidden-Test-Dataset), und für die Varianten-Scans die Werte 64,39 (Public-Test-Dataset) und 66,89 (Hidden-Test-Dataset) verzeichnen (Sekuboyina et al., 2021). Von besonders großem Interesse sind nun die durchschnittlichen DICE-Scores aller Teilnehmer. Sie betragen für die normalen Scans 76,4 (Public-Test-Dataset) und 72,63 (Hidden-Test-Dataset), und für die Varianten-Scans 63,89 (Public-Test-Dataset) und 64,49 (Hidden-Test-Dataset) (Sekuboyina et al., 2021). Bei den Top-5 ließen sich Durchschnittswerte für den DICE-Score von 90,7 (Public-Test-Dataset) und 89,59 (Hidden-Test-Dataset) für die normalen Scans, und 81,1 (Public-Test-Dataset) und 84,53 (Hidden-Test-Dataset) für die Varianten-Scans ermitteln (Sekuboyina et al., 2021). Für eine bessere Übersicht sei an dieser Stelle nochmals auf Tabelle 12 verwiesen. Insbesondere bei Betrachtung der durchschnittlichen DICE-Scores der Top-5 Performer und der gesamten Teilnehmerschaft wird ersichtlich, dass diejenigen Scans, die anatomische Varianten der Wirbelsäule zeigen, deutlich schlechtere Werte aufweisen als die normalen Scans. Aufgrund

dieser Tatsache ist der Gedanke naheliegend, dass anatomische Varianten weiterhin als für den Labelling- und Segmentierungsprozess problematisch angesehen werden sollten, und dass ihre Erkennung durch Deep Learning vermittelte Algorithmen weiterhin optimiert werden könnte (Sekuboyina et al., 2021). Es zeigt sich bei genauerer Betrachtung jedoch auch, dass die Diskrepanz zwischen den DICE-Scores der normalen Scans und der Varianten-Scans stark vom jeweiligen Algorithmus abhängig ist, und somit bei einigen Teilnehmern deutlich geringer ausgeprägt ist als bei anderen (Sekuboyina et al., 2021).

*Tabelle 12: Vergleich der Performance an „Normal-Scans“ und „Varianten-Scans“ (Sekuboyina et al., 2021)*

Team	DICE-Score (Public)		DICE-Score (Hidden)	
	„Normale Scans“	„Varianten Scans“	„Normale Scans“	„Varianten Scans“
Chen D. et al (1. Platz)	93,42	84,21	92,19	87,02
Payer C. et al (2. Platz)	92,94	85,96	89,74	89,59
Zhang A. et al (3. Platz)	89,2	87,15	89,81	87,35
Sekuboyina A. et al (Anduin)	81,14	64,39	82,38	66,89
Top – 5	90,7	81,1	89,59	84,53
Alle	76,4	63,89	72,63	64,49

## 5. Diskussion

Im Rahmen dieser Arbeit wurden anatomische Varianten der Wirbelsäule hinsichtlich verschiedener Gesichtspunkte untersucht. Die hierbei ermittelten Werte bezüglich der Prävalenz der verschiedenen Varianten unterschieden sich teilweise deutlich von den präexistenten Angaben anderer Autoren (Apazidis et al., 2011; Thawait et al., 2012; Tins & Balain, 2016). Die vielfach in der Fachliteratur vorhandenen Vorschläge einer jeweiligen Klassifikation der unterschiedlichen anatomischen Varianten wurden im Hinblick auf die in unserem Projekt erarbeiteten Klassifikationskriterien an einigen Stellen in ihrer Gänze, oder aber teilweise oder in angepasster Fassung übernommen, an anderer Stelle aber auch gänzlich verworfen (Brewin et al., 2009; Carrino et al., 2011; Castellvi et al., 1984; Kim et al., 2008; McCulloch & Waddell, 1980; Park et al., 2016; Thawait et al., 2012; Tins & Balain, 2016; Tureli et al., 2014; Wigh, 1980; Yun et al., 2018). Im Folgenden sollen die diesbezüglichen Aussagen der bestehenden Literatur im Genaueren genannt und mit dieser Arbeit in Beziehung gesetzt werden. Es zeigte sich darüber hinaus, dass anatomische Varianten der Wirbelsäule eine große Herausforderung für automatisierte Segmentierungsalgorithmen darstellen und deren Performance zum aktuellen Zeitpunkt erheblich limitieren, sodass das gezielte Training der Letzteren im Hinblick auf die Erkennung und Verarbeitung solcher Varianten deren Funktionsweise bedeutend verbessern kann (Sekuboyina et al., 2021).

### 5.1 Klinische Relevanz

Russell Wigh beschreibt die Tatsache, dass das Vorliegen einer Wirbelsäulen-anomalie ein bedeutsamer prädisponierender Faktor für die Addressierung eines falschen Wirbelsäulensegments bei chirurgischen Eingriffen an der Wirbelsäule ist (Wigh, 1980). So kam es laut Wigh et al. innerhalb eines Kollektivs von 100 Patienten, die einer Laminektomie bei Diskusherniation zugeführt wurden, ausschließlich bei solchen Patienten zu einer Addressierung des falschen Wirbelsäulensegments, bei denen auch eine Wirbelsäulen-anomalie vorlag (Wigh, 1980). Es wird schließlich behauptet, dass ein besseres Verständnis der Anatomie der Wirbelsäule sowie auch deren Varianten, maßgeblich zur Reduktion derartiger Fehleingriffe beitragen könne, wobei insbesondere das Vorhandensein von

Übergangswirbeln noch zu Fehlern führe (Wigh, 1980). Auch Park et al. nennen die Verbesserung der klinischen Praxis bzw. die Vermeidung von Fehlern in Diagnostik und operativer Therapie als Hauptgrund für die Notwendigkeit einer einheitlichen Klassifikation von anatomischen Varianten der Wirbelsäule (Park et al., 2016). Die Bedeutung einer solchen Klassifikation insbesondere im Hinblick auf chirurgische Eingriffe an der Wirbelsäule wird zudem vom Thawait et al. thematisiert (Thawait et al., 2012). Konin et al. formulieren diese Notwendigkeit speziell für lumbosakrale Übergangswirbel, wobei ein klinischer Bezug zu Schmerzen im Bereich der Lendenwirbelsäule beschrieben wird (Konin & Walz, 2010). Überdies stellen auch Vinha et al., welche in ihrer Arbeit eine südeuropäische Referenzpopulation betrachten, und Hanhivaara et al., im Bezug auf eine Kohorte finnischer Studienteilnehmer, eine Korrelation zwischen dem Vorhandensein einer lumbosakralen Übergangsanomalie und der Entwicklung einer Lumbago fest (Hanhivaara et al., 2022; Vinha et al., 2022). In der jüngeren Literatur wurden außerdem lumbosakrale Übergangswirbel als unabhängiger Risikofaktor für das Vorliegen einer Spinalkanalstenose diskutiert (Harris et al., 2019; Rabau et al., 2021), sowie für das Vorhandensein einer Hüftdysplasie bzw. Hüftgelenksfehlstellung (Becker et al., 2022; Sun et al., 2021), und nicht zuletzt für das Auftreten erworbener, degenerativer Wirbelsäulenveränderungen (Griffith et al., 2022). Zusätzlich zu diesen an mehrfacher Stelle in der Literatur beschriebenen klinischen Aspekten, soll im Zusammenhang mit der im Rahmen dieser Arbeit thematisierten VerSe – 20 Challenge insbesondere die Bedeutung einer Klassifikation anatomischer Varianten der Wirbelsäule für das Prinzip der CAD (Computer-Aided-Diagnosis) betont werden. Schließlich steckt hierin laut Sekuboyina et al. ein bedeutsames Potential für die Verbesserung der Performance automatisierter Segmentierungsalgorithmen (Sekuboyina et al., 2021).

## 5.2 Grundsätzliche Aspekte

Bei der Erstellung des Goldstandards für das im Rahmen dieser Arbeit erhobene Datenmaterial orientierten wir uns klar an den Vorschlägen bestehender Literatur, wobei sich nicht selten Widersprüche zwischen den Arbeiten verschiedener Autoren offenbarten, und auch teilweise Kriterien in das System integriert wurden, die von uns als Arbeitsgruppe selbst, unabhängig von der präexistenten Literatur, erarbeitet

worden sind. Die Einzelheiten des Klassifikationssystems sind im Kapitel 3 dieser Arbeit, sowie in einer im Rahmen der VerSe – 20 Challenge veröffentlichten Arbeit von Liebl et al. nachzulesen (Liebl et al., 2021). Es wurde zunächst entschieden, den noch häufig gängigen Ansatz, den 20. freien Wirbel der Wirbelsäule bei kraniokaudaler Zählweise pauschal als ersten Lendenwirbel zu bezeichnen, zu verwerfen (Liebl et al., 2021). So kommen bereits Yun et al. zu dem Schluss, dass diese Art der Nummerierung gerade unter Berücksichtigung der vielfältigen anatomischen Varianten der Wirbelsäulen-anatomie veraltet und für sich genommen nicht ausreichend ist, eine gute klinische Praxis zu gewährleisten (Yun et al., 2018). Vielmehr sollte laut Yun et al. der Wirbelmorphologie bei der Klassifikation und Enumeration der Wirbelsäule mehr Beachtung geschenkt werden (Yun et al., 2018). Dieser Sichtweise schlossen wir uns als Arbeitsgruppe insbesondere aus dem Grund an, dass die Charakterisierung der einzelnen Wirbel anhand morphologischer Merkmale, im Unterschied zur Charakterisierung nach kraniokaudaler Zählweise, auch dann durchführbar ist, wenn, bspw. auf einem CT-Scan, nicht die gesamte Wirbelsäule abgebildet ist (Liebl et al., 2021). Tureli et al. haben die Verlässlichkeit diverser potentieller "Landmarks" überprüft, also anatomischer Strukturen, deren Lagebeziehung zur Wirbelsäule als Hinweis dienen können bei der Frage nach der korrekten Zuordnung einzelner Wirbel (Tureli et al., 2014). Zu den von Tureli et al. untersuchten Strukturen zählen das Ligamentum iliolumbale (ILL), die 12. Rippe, die Aortenbifurkation (AB), die Arteria renalis dextra (RRA) sowie der Conus medullaris (CM) (Tureli et al., 2014). Es zeigte sich, dass bis auf das ILL alle o.g. Strukturen als "Landmarks" ungeeignet sind, da sie in ihrer Lagebeziehung zur Wirbelsäule zu unbeständig sind (Tureli et al., 2014). Das ILL entspringt zwar relativ verlässlich von demjenigen Wirbel, der als kaudalster Lumbalwirbel zu bezeichnen ist, doch kamen Tureli et al. zu dem Schluss, dass es in der Bildgebung nicht zuverlässig genug zu identifizieren sei, als dass man es regelhaft als "Landmark" verwenden könne (Tureli et al., 2014). Carrino et al. haben unter anderem die Lagebeziehung des ILL zur Wirbelsäule untersucht, und sind dabei zu dem Schluss gekommen, dass das ILL zuverlässig vom kaudalsten Lumbalwirbel entspringt, welcher nicht zwangsläufig L5 sein muss (Carrino et al., 2011). Auf dem Boden dieser Erkenntnisse wurde im Rahmen unseres Projektes beschlossen, das ILL in unserer Klassifikation lediglich als ein weiches Kriterium bei der Differenzierung lumbosakraler Übergangswirbel einfließen zu lassen (siehe Kapitel 3.1.6) (Liebl et al., 2021). Alle weiteren zuvor

erwähnten anatomischen "Landmarks" wurden aus den oben genannten Gründen als für dieses Klassifikationssystem bedeutungslos angesehen (Liebl et al., 2021). Ebenfalls wurde durch Carrino et al. eine mögliche Assoziation zwischen dem vorliegenden Grad der Klassifikation nach O'Driscoll und dem Vorhandensein eines lumbosakralen Übergangswirbels untersucht, außerdem eine mögliche Assoziation zwischen einem von der Norm abweichenden LSIVDA ("lumbosacral intervertebral disc angle") und dem Vorhandensein eines lumbosakralen Übergangswirbels, sowie eine mögliche Assoziation zwischen dem Vorhandensein eines Übergangswirbels (thorakolumbal oder lumbosakral) und einer von der Norm abweichenden Anzahl präsakraler Wirbel (Carrino et al., 2011). Eine signifikante Assoziation zum Vorliegen eines lumbosakralen Übergangswirbels ließ sich für das Vorhandensein eines Typ-4-Diskus nach O'Driscoll et al. nachweisen, jedoch nicht für einen von der Norm abweichenden LSIVDA (Carrino et al., 2011). Ferner zeigten Carrino et al., dass eine von der Norm abweichende Anzahl präsakraler Wirbel nicht signifikant mit dem Vorliegen eines thorakolumbalen Übergangswirbels, wohl aber mit dem Vorliegen eines lumbosakralen Übergangswirbels, assoziiert ist (Carrino et al., 2011). Da die Klassifikation nach O'Driscoll die Beurteilung der Bandscheibenmorphologie am lumbosakralen Übergang in MRT-Bildern vorsieht (O'Driscoll et al., 1996), jedoch in unserem Projekt ausschließlich CT-Bilder verwendet wurden, hat diese Klassifikation keine tragende Rolle im Hinblick auf diese Arbeit gespielt (Liebl et al., 2021). Auch der zuvor erwähnte LSIVDA findet bei uns keine weitere Verwendung, was mit seiner unzulänglichen Aussagekraft begründbar ist (Carrino et al., 2011). An dieser Stelle sei erwähnt, dass eine gewisse Assoziation einer numerischen Anomalie der Wirbelsäule (wenngleich in dieser Arbeit lediglich numerische Varianten der LWS betrachtet worden sind), nach unseren Berechnungen nicht nur mit dem Vorhandensein eines thorakolumbalen Übergangswirbels besteht (vgl. Carrino et al.), sondern auch mit dem Vorhandensein eines lumbosakralen Übergangswirbels (vgl. Kapitel 4.1). Von klinischer Bedeutung erscheint im Falle des Vorliegens einer anatomischen Variante der Wirbelsäule die damit einher gehende Verteilung von Dermatomen und Kennmuskeln auf die einzelnen Wirbelsäulensegmente, welche unter anderem durch Thawait et al. vorgestellt wird (Thawait et al., 2012). Eine solche Zuordnung wird zunächst für die Normwirbelsäule beschrieben, und anschließend die häufigsten bei lumbosakralen Übergangsanomalien zu beobachtenden Phänotypen genannt (Thawait et al., 2012).

### Normwirbelsäule:

#### **- L4:**

Kennmuskeln: M. vastus medialis und lateralis;

Dermatom: Oberschenkel anterior

#### **- L5:**

Kennmuskeln: M. tibialis anterior und M. extensor hallucis longus;

Dermatom: Oberschenkel lateral, Wade anterolateral, Fußrücken

#### **- S1:**

Kennmuskeln: M. gastrocnemius (medialer Kopf) und M. soleus;

Dermatom: Oberschenkel dorsal, Wade, Fußsohle

Bei Vorliegen einer lumbosakralen Übergangsanomalie kann diese Zuordnung variieren (Thawait et al., 2012). So zeigt sich laut McCulloch & Waddell der kaudalste mobile Wirbel bezüglich der Innervation in der Regel als funktioneller L5 (McCulloch & Waddell, 1980). Laut Kim et al. weist bei Vorliegen eines lumbalisierten S1 eben jener S1 ein typisches S1-Innervationsmuster, und L5 das typische L5-Innervationsmuster auf, während bei sakralisiertem L5 das typische L5-Innervationsmuster L4 zuzuordnen war (Kim et al., 2008). Zu beachten ist jedoch, dass bei Patienten mit lumbosakraler Übergangsanomalie häufig verschieden ausgeprägte Überlappungen bei der Innervation zu beobachten waren (Thawait et al., 2012). Die von Thawait et al. erwähnten klinischen Aspekte bzgl. des Verhaltens von Dermatomen und Kennmuskeln bei anatomischen Varianten der Wirbelsäule, mögen zwar von großem klinischem Interesse sein, können aber aufgrund ihrer klinischen Natur kaum Berücksichtigung finden in einem nach radiologischen Aspekten erstellten Klassifikationssystem, wie es in dieser Arbeit vorliegt (Liebl et al., 2021).

## 5.3 Zervikothorakaler Übergang

Da die von Brewin et al. formulierten Kriterien (vgl. Kap. 3.1.2) für das Vorliegen von Halsrippen als gleichzeitiges Kriterium für das Vorliegen einer zervikothorakalen Übergangsanomalie (Brewin et al., 2009) in keinerlei Widerspruch zu den in unserem Datensatz vorhandenen Fällen standen, wurden sie als solche in unseren

Goldstandard übernommen (Liebl et al., 2021).

## 5.4 Thorakolumbaler Übergang

Wie im Kapitel 3.1.3 dieser Arbeit bereits erwähnt, wurde die von Russell Wigh vorgeschlagene, vier Kategorien umfassende Einteilung von Wirbeln im Bereich des thorakolumbalen Übergangs (regulärer thorakaler Wirbel, regulärer lumbaler Wirbel, thorakolumbaler Übergangswirbel vom thorakalen sowie vom lumbalen Typ) aus dem Grund, da sie einerseits die tatsächlichen morphologischen Verhältnisse am besten beschreibt, andererseits aber auch eine durch eine automatisierte Segmentierung praktikable Einteilung vornimmt, in das Klassifikationssystem der Arbeitsgruppe integriert (Liebl et al., 2021); (Wigh, 1980). Um die Funktionsweise des Segmentierungsalgorithmus zu verbessern, wurden die Kriterien der Differenzierung zwischen dem thorakalen und dem lumbalen Typ des thorakolumbalen Übergangswirbels im Rahmen des Projektes um einige eigenständig in der Arbeitsgruppe erarbeitete Kriterien erweitert (vgl. hierzu Kap. 3.1.3) (Liebl et al., 2021).

Eine ähnliche Kategorisierung wie Russell Wigh nehmen auch Park et al. vor (Park et al., 2016), die jedoch anstatt von vier Kategorien zunächst die drei folgenden Hauptkategorien nennen:

- Normaler thorakaler Wirbel:  
Ein Wirbel am thorakolumbalen Übergang ist dann als normaler thorakaler Wirbel anzusehen, wenn beidseits artikulierende Rippen normaler Länge (= 38 mm oder > 38 mm) vorliegen (Park et al., 2016).
- Normaler lumbaler Wirbel:  
Ein Wirbel am thorakolumbalen Übergang kann dann als normaler lumbaler Wirbel bezeichnet werden, wenn er beidseits über knöchern verschmolzene Processus transversi und über keine Rippen verfügt (Park et al., 2016).
- Thorakolumbaler Übergangswirbel:  
Ein Wirbel am thorakolumbalen Übergang ist dann als thorakolumbaler



Übergangswirbel anzusehen, wenn die o.g. Kriterien sowohl für den normalen thorakalen als auch für den normalen lumbalen Wirbel nicht erfüllt sind. Ein solcher Übergangswirbel kann dann anhand morphologischer Eigenschaften näher subklassifiziert werden (Park et al., 2016).

Eine Reihe weiterer Kriterien ermöglicht nach dem Klassifikationssystem von Park et al. die Subklassifikation thorakolumbalen Übergangswirbel (s.o.) (Park et al., 2016). Diese Kriterien sind im Folgenden aufgelistet.

- Typ I)

Mindestens eine Stummelrippe (< 38 mm, artikuliert mit dem Wirbel im Wirbelkörper oder im Pedikel) liegt vor, kontralateral liegt entweder Stummelrippe oder eine Normalrippe (= 38 mm oder > 38 mm) vor (Park et al., 2016).

- Typ IIa)

Eine Normalrippe liegt vor, kontralateral liegt entweder ein akzessorisches Ossifikationszentrum oder ein gewöhnlicher Querfortsatz vor (Park et al., 2016).

- Typ IIb)

Eine Stummelrippe liegt vor, kontralateral liegt entweder ein akzessorisches Ossifikationszentrum oder ein gewöhnlicher Querfortsatz vor (Park et al., 2016).

- Typ III)

Eine oder zwei "Mixed-type-Rippen" liegen vor (Rippenähnliche Strukturen, für die weder die Definition eines akzessorischen Ossifikationszentrums noch diejenige für eine Normal- oder Stummelrippe zutrifft) (Park et al., 2016).

- Typ IV)

Mindestens ein akzessorisches Ossifikationszentrum (nach Park et al. definiert als eine knöcherne Struktur, die einem verkürzten Proc. costalis anliegt oder mit ihm artikuliert) liegt vor, kontralateral liegt entweder ein weiteres akz. Ossifikationszentrum oder ein gewöhnlicher Querfortsatz vor (Park et al., 2016).

Darüber hinaus wird ein Vorschlag genannt zur Einordnung des jeweiligen Übergangswirbels entweder zum thorakalen oder zum lumbalen Abschnitt der Wirbelsäule (Park et al., 2016). Hierbei gilt Typ I eher als thorakal, Typ IV eher als lumbal, und die Typen II und III als unklar (Park et al., 2016).

Die Klassifikation des thorakolumbalen Übergangs in der von uns erarbeiteten Systematik fußt im Wesentlichen auf den von Russell Wigh eingeführten Grundkategorien (Liebl et al., 2021). Als ergänzende Gesichtspunkte sind die in Kapitel 3 dargestellten weichen Kriterien von uns eingeführt worden (Liebl et al., 2021). Die Einbindung morphologischer Eigenschaften des Wirbels an sich in den Klassifikationsprozess erschien uns weniger fehleranfällig als die Methode von Park et al., zu diesem Zweck die Stummelrippen in der beschriebenen Weise näher zu subcharakterisieren, weshalb diese zuletzt genannten Kriterien nicht in unser Klassifikationssystem übernommen wurden (Liebl et al., 2021).

## 5.5 Lumbosakraler Übergang

Anders als beim thorakolumbalen Übergang erarbeitet Russell Wigh beim lumbosakralen Übergang lediglich morphologische Kriterien, die Aufschluss darüber geben, ob es sich um einen Übergangswirbel handelt oder nicht (Wigh, 1980). Dabei wird primär nicht die Frage gestellt, welchem Wirbelsäulenabschnitt der Übergangswirbel zugerechnet werden soll (Wigh, 1980). Vielmehr betont Wigh, dass die untere Begrenzung der Lendenwirbelsäule unterhalb des letzten wahren Lendenwirbels liegt, welcher laut Wigh der kaudalste Wirbel ist, dessen Processus transversus nicht mit einem benachbarten Wirbel knöchern verschmolzen ist oder artikuliert (Wigh, 1980). Demnach soll ein lumbosakraler Übergangswirbel weder der Lendenwirbelsäule noch dem Sakrum zugeordnet werden (Wigh, 1980). Die folgenden Merkmale deuten laut Wigh auf das Vorhandensein eines lumbosakralen Übergangswirbels hin (Wigh, 1980):

- i) Processus transversus uni- oder bilateral mit dem Sakrum knöchern verschmolzen
- ii) Processus transversus uni- oder bilateral mit dem Sakrum artikulierend
- iii) Processus transversus auf einer Seite knöchern mit dem Sakrum verschmolzen, kontralateral Artikulation zwischen dem Processus transversus

und dem Sakrum.

- iv) Squaring: Squaring beschreibt eine Reihe morphologischer Eigenschaften, die ein vermeintlicher S1 (Erster Sakralwirbel von kranial), aufweisen kann. Dazu zählen ein quadratisches Erscheinungsbild des Wirbelkörpers in coronarer Ansicht, und der Verlust der konischen Form des Wirbelkörpers in sagittaler Ansicht (zu ermitteln als Ratio [Tiefe der sup. Fläche / Tiefe der inf. Fläche] = 1,37 oder kleiner)
- v) Flaring: Flaring beschreibt eine bulbäre Konfiguration lateral und superior an der Ala ossis sacri eines vermeintlichen S1, die wiederum morphologische Ähnlichkeiten zu einem Processus costalis aufweist. Ein positives Flaring gibt aufgrund hoher Koinzidenz Anlass zur gezielten Suche nach dem "Squaring-sign" (s.o.), im Sinne einer Abklärung eines möglichen Übergangswirbels.
- vi) Schräge Verlaufsform der superioren Fläche des Wirbelkörpers bei vermeintlichem letztem Lumbalwirbel.

Eine alternative Möglichkeit zur Beschreibung lumbosakraler Übergangswirbel, die zudem eine Einordnung derselben in verschiedene Subkategorien ermöglicht, bieten Castellvi et al. (Castellvi et al., 1984) (s. Kapitel 3.1.5). Die Castellvi – Klassifikation des lumbosakralen Übergangs wurde in ihrer Gesamtheit in unser Klassifikationssystem integriert (Castellvi et al., 1984; Liebl et al., 2021). Die von Wigh et al. erarbeiteten Kriterien spielen hier lediglich eine untergeordnete Rolle, wobei allenfalls das zuvor beschriebene "Squaring" als Aspekt eines weichen Kriteriums in unsere Methoden übernommen wurde (Liebl et al., 2021).

## 5.6 Prävalenzen anatomischer Varianten der Wirbelsäule

Tins et al. untersuchten separat die Prävalenzen von numerischen Aberrationen und Übergangswirbeln der LWS (Tins & Balain, 2016). In diesem Ansatz wurde die HWS als 7-gliedrig und die BWS als 12-gliedrig prädefiniert (Tins & Balain, 2016). Bei Vorliegen eines lumbosakralen Übergangswirbels galt dieser als sakralisierter Lendenwirbel, wenn er gemäß kraniokaudaler Zählung und unter Annahme einer 12-gliedrigen BWS unterhalb von L4 positioniert war (Tins & Balain, 2016). Befand er sich unterhalb von L5, so wurde er als lumbalisierter Sakralwirbel bezeichnet (Tins & Balain, 2016). Die im Rahmen der Studie ermittelten Prävalenzen betragen für

numerische Aberrationen ca. 7,7 % und für Übergangswirbel ca. 3,3 % (Tins & Balain, 2016). Somit kommen Tins et al. zu dem Schluss, dass im Bereich der LWS numerische Aberrationen mehr als doppelt so häufig in der unselektierten Bevölkerung anzutreffen sind als Übergangswirbel (Tins & Balain, 2016). Ferner weisen Tins et al. darauf hin, dass eine korrekte Beurteilung der Wirbelsäule im Hinblick auf numerische Aberrationen nur unter Zuhilfenahme einer Bildgebung gelingt, welche die gesamte Wirbelsäule abbildet (Wholespine-imaging) (Tins & Balain, 2016). Gerade die Annahme, dass der zwanzigste mobile Wirbel in kraniokaudaler Zählweise dem ersten Lendenwirbel entspricht, die explizit nicht in unser Klassifikationssystem aufgenommen wurde (Liebl et al., 2021), erklärt, dass die Beurteilung der Wirbelsäule nach dieser Methode ein Wholespine-imaging erfordert. Im Gegensatz dazu wurden in dieser Arbeit feine Kriterien (s.o.) zur exakten Klassifikation jedes einzelnen Wirbels erarbeitet, die diese auch dann ermöglichen, wenn nur ein Teilabschnitt der Wirbelsäule abgebildet ist (Liebl et al., 2021). Die im Rahmen dieser Arbeit geschätzte Prävalenz numerischer Aberrationen der LWS beträgt 10,23 % und unterscheidet sich somit nur unwesentlich von der mit 7,7 % angegebenen Prävalenz nach Tins et al. (Tins & Balain, 2016). Allerdings besteht eine große Diskrepanz hinsichtlich der Prävalenz lumbosakraler Übergangswirbel, die von Tins et al. auf 3,3 %, und in unserer Arbeit auf 12,43 % geschätzt wird (Tins & Balain, 2016). Vergleicht man diese beiden Ergebnisse hinsichtlich ihrer Plausibilität, so muss von unserer Seite eingeräumt werden, dass die aus 185 Patienten bestehende Stichprobe, die im Rahmen dieser Arbeit zur Berechnung der Prävalenz lumbosakraler Übergangswirbel verwendet wurde, deutlich kleiner ist als die Stichprobe, die Tins et al. zu diesem Zweck betrachtet haben, welche 418 Patienten umfasst (Tins & Balain, 2016). Gleichzeitig muss jedoch die durch Tins et al. berechnete Prävalenz numerischer Varianten der LWS, trotz der starken Ähnlichkeit zu dem in unserer Arbeit berechneten Wert, wiederum angezweifelt werden, zumindest nach den in dieser Arbeit verwendeten Definitionen. Wird nämlich lediglich untersucht, ob jenseits des 19. mobilen Wirbels eine numerische Aberration vorliegt, so ist dies gleichbedeutend mit der Suche nach dem Anteil an Patienten, die im Allgemeinen von der 24-Gliedrigkeit der mobilen Wirbelsäule abweichen. Auf diese Weise werden also diejenigen Fälle, bei denen zwar eine 24-Gliedrigkeit der Wirbelsäule vorliegt, jedoch eine anomale Verteilung zwischen Brust- und Lendenwirbelsäule besteht (11/6; 13/4) nicht erfasst. Da wir

nach den von uns erarbeiteten Kriterien jedoch solche Fälle ebenfalls als numerische Aberrationen (unter anderem) der LWS verstehen, besteht die Möglichkeit, dass der von Tins et al. angegebene Wert den wahren Wert nach unseren Definitionen unterschätzt. Thawait et al. nennen für lumbosakrale Übergangsanomalien eine Prävalenz von ca. 4-30 % (Thawait et al., 2012). Aufgrund dieses breiten Intervalls besteht kein Widerspruch zu dem in dieser Arbeit ermittelten Prävalenzwert sowohl numerischer Aberrationen als auch lumbosakraler Übergangswirbel. Apazidis et al. untersuchen die Prävalenz anatomischer Varianten der LWS in einem Kollektiv von 211 Patienten (Apazidis et al., 2011). Für lumbosakrale Übergangswirbel beschreiben sie eine geschätzte Prävalenz von 35,6 %, welche wiederum deutlich höher liegt als die in unserem Projekt ermittelte Prävalenz von 12,43 % (Apazidis et al., 2011). Dieser Unterschied lässt sich womöglich durch die Tatsache erklären, dass Apazidis et al. auch Patienten mit einem Übergangswirbel nach Castellvi Grad I gewertet haben, während in dieser Studie mindestens der Castellvi Grad II vorliegen musste, um die Voraussetzungen für die Klassifikation als Übergangswirbel zu erfüllen (Apazidis et al., 2011). Die Prävalenz numerischer Aberrationen der LWS geben Apazidis et al. mit 6,6 % an, womit sie nur relativ schwach von dem in unserem Projekt ermittelten Wert von 10,23 % abweicht (Apazidis et al., 2011).

## 5.7 VerSe – 20 Challenge

Sekuboyina et al. beschreiben einen deutlichen Anstieg der insgesamt zu beobachtenden Performance der verschiedenen Segmentierungsalgorithmen bei der VerSe – 20 Challenge im Vergleich zur VerSe – 19 Challenge (Sekuboyina et al., 2021). Den Grund für diese Verbesserung sehen Sekuboyina et al. in der Tatsache, dass aufgrund der Verwendung eines Datensatzes mit überrepräsentativ großem Anteil an anatomischen Varianten der Wirbelsäule, die Prozessierung eben dieser Anomalien durch die Algorithmen gezielt trainiert wurde (Sekuboyina et al., 2021). Dies bedeutet im Umkehrschluss, dass die fehlerhafte Performance an Scans, die anatomische Varianten der Wirbelsäule darstellen, mitursächlich für die relativ höhere Fehlerquote der VerSe – 19 – Challenge war (Sekuboyina et al., 2021). Schließlich sind anatomische Varianten der Wirbelsäule ein häufiges Phänomen, das bei der Diagnostik sorgfältig berücksichtigt werden muss (Wigh, 1980). Aus den beiden

VerSe – Challenges ging ebenfalls hervor, dass bei ausreichendem Dateninput die suffiziente Erkennung und Prozessierung von Scans mit enthaltenem Th13 (bei 13-gliedriger BWS) oder L6 (bei 6-gliedriger LWS) durch Segmentierungsalgorithmen gewährleistet werden kann (Sekuboyina et al., 2021). Es ist davon auszugehen, dass auf diesem Wege noch weiteres Potential besteht für die Verbesserung der Performance von Segmentierungsalgorithmen, und so geben Sekuboyina et al. einen Ausblick, dass künftige Arbeiten dieses Potential in Angriff nehmen könnten (Sekuboyina et al., 2021).

## 6. Zusammenfassung

Anatomische Varianten der Wirbelsäule sind ein häufig in der unselektierten Bevölkerung anzutreffendes Phänomen (Wigh, 1980). Dabei ist die Studienlage bezüglich der Prävalenzen der einzelnen Subtypen unklar. So variieren beispielsweise die Schätzungen zur Häufigkeit lumbosakraler Übergangswirbel, die nach den im Rahmen dieser Arbeit angestellten Berechnungen bei 12,43% liegt, in der präexistenten Literatur durchaus stark (Apazidis et al., 2011; Thawait et al., 2012; Tins & Balain, 2016). Unabhängig davon kann jedoch konstatiert werden, dass anatomische Varianten der Wirbelsäule von großer Bedeutung für eine gute klinische Praxis sind, und ein besonderes Augenmerk bei der Diagnostik verdienen, wie bereits mehrfach andernorts in der Fachliteratur eingefordert wird (Park et al., 2016; Thawait et al., 2012; Wigh, 1980). Überdies gehen wir aufgrund der in Kapitel 4 dargestellten Berechnungen davon aus, dass eine hohe Koprävalenz zwischen den verschiedenen Formen von Wirbelsäulenomalien untereinander besteht (vgl. Kapitel 4). Es liegt also der Gedanke nahe, dass insbesondere bei Vorliegen einer solchen, ein zweiter Blick im Sinne des Ausschlusses einer weiteren Anomalie, von Seiten des Diagnostikers lohnt. Von nicht geringerer Bedeutung sind anatomische Varianten der Wirbelsäule für die Funktionsweise automatisierter Segmentierungsalgorithmen, deren Leistungsfähigkeit und Genauigkeit allein seit 2019 deutlich zugenommen haben (Sekuboyina et al., 2021). Insbesondere die Verbesserung der Performance der an der VerSe – 20 Challenge teilnehmenden Algorithmen im Vergleich zum Vorjahr lässt sich ganz wesentlich auf die Tatsache zurückführen, dass die Algorithmen im diesjährigen Wettbewerb gezielt auf die Erkennung und Prozessierung anatomischer Varianten der Wirbelsäule trainiert wurden (Sekuboyina et al., 2021). Durch den in Kapitel 4 dargestellten Vergleich zwischen der Performance verschiedener Algorithmen einerseits an solchen Scans, die keine Anomalien der Wirbelsäule enthalten, und andererseits an solchen, die mindestens eine Anomalie enthalten, wird jedoch ersichtlich, dass noch immer Potential besteht, über die Verbesserung der Erkennung und Prozessierung anatomischer Varianten der Wirbelsäule durch eben diese Algorithmen, deren durchschnittliche Leistung erheblich zu verbessern.

## 7. Literaturverzeichnis

- Apazidis, A., Ricart, P. A., Diefenbach, C. M., & Spivak, J. M. (2011). The prevalence of transitional vertebrae in the lumbar spine. *Spine J*, 11(9), 858-862. doi:10.1016/j.spinee.2011.08.005
- Becker, L., Taheri, N., Haffer, H., Muellner, M., Hipfl, C., Ziegeler, K., Diekhoff, T., & Pumberger, M. (2022). Lumbosacral Transitional Vertebrae Influence on Acetabular Orientation and Pelvic Tilt. *J Clin Med*, 11(17). doi:10.3390/jcm11175153
- Brewin, J., Hill, M., & Ellis, H. (2009). The prevalence of cervical ribs in a London population. *Clin Anat*, 22(3), 331-336. doi:10.1002/ca.20774
- Carrino, J. A., Campbell, P. D., Jr., Lin, D. C., Morrison, W. B., Schweitzer, M. E., Flanders, A. E., Eng, J., & Vaccaro, A. R. (2011). Effect of spinal segment variants on numbering vertebral levels at lumbar MR imaging. *Radiology*, 259(1), 196-202. doi:10.1148/radiol.11081511
- Castellvi, A. E., Goldstein, L. A., & Chan, D. P. (1984). Lumbosacral transitional vertebrae and their relationship with lumbar extradural defects. *Spine (Phila Pa 1976)*, 9(5), 493-495. doi:10.1097/00007632-198407000-00014
- Choi, R. Y., Coyner, A. S., Kalpathy-Cramer, J., Chiang, M. F., & Campbell, J. P. (2020). Introduction to Machine Learning, Neural Networks, and Deep Learning. *Transl Vis Sci Technol*, 9(2), 14. doi:10.1167/tvst.9.2.14
- Glocker, B., Zikic, D., Konukoglu, E., Haynor, D. R., & Criminisi, A. (2013). Vertebrae localization in pathological spine CT via dense classification from sparse annotations. *Med Image Comput Comput Assist Interv*, 16(Pt 2), 262-270. doi:10.1007/978-3-642-40763-5\_33
- Griffith, J. F., Xiao, F., Hilkens, A., Griffith, I. H. Y., & Leung, J. C. S. (2022). Increased vertebral body area, disc and facet joint degeneration throughout the lumbar spine in patients with lumbosacral transitional vertebrae. *Eur Radiol*, 32(9), 6238-6246. doi:10.1007/s00330-022-08736-0
- Hanhivaara, J., Maatta, J. H., Karppinen, J., Niinimäki, J., & Nevalainen, M. T. (2022). The Association of Lumbosacral Transitional Vertebrae with Low Back Pain and Lumbar Degenerative Findings in MRI: A Large Cohort Study. *Spine (Phila Pa 1976)*, 47(2), 153-162. doi:10.1097/BRS.0000000000004244
- Harris, G., Ball, J., & De Decker, S. (2019). Lumbosacral transitional vertebrae in cats and its relationship to lumbosacral vertebral canal stenosis. *J Feline Med Surg*, 21(4), 286-292. doi:10.1177/1098612X18774449
- Hussain, D., & Han, S. M. (2019). Computer-aided osteoporosis detection from DXA imaging. *Comput Methods Programs Biomed*, 173, 87-107. doi:10.1016/j.cmpb.2019.03.011
- Kaka, H., Zhang, E., & Khan, N. (2021). Artificial Intelligence and Deep Learning in Neuroradiology: Exploring the New Frontier. *Can Assoc Radiol J*, 72(1), 35-44. doi:10.1177/0846537120954293
- Kim, Y. H., Lee, P. B., Lee, C. J., Lee, S. C., Kim, Y. C., & Huh, J. (2008). Dermatome variation of lumbosacral nerve roots in patients with transitional lumbosacral vertebrae. *Anesth Analg*, 106(4), 1279-1283, table of contents. doi:10.1213/ane.0b013e3181605635
- Konin, G. P., & Walz, D. M. (2010). Lumbosacral transitional vertebrae: classification, imaging findings, and clinical relevance. *AJNR Am J Neuroradiol*, 31(10), 1778-1786. doi:10.3174/ajnr.A2036



- Lessmann, N., van Ginneken, B., de Jong, P. A., & Isgum, I. (2019). Iterative fully convolutional neural networks for automatic vertebra segmentation and identification. *Med Image Anal*, 53, 142-155. doi:10.1016/j.media.2019.02.005
- Liebl, H., Schinz, D., Sekuboyina, A., Malagutti, L., Loffler, M. T., Bayat, A., El Hussein, M., Tetteh, G., Grau, K., Niederreiter, E., Baum, T., Wiestler, B., Menze, B., Braren, R., Zimmer, C., & Kirschke, J. S. (2021). A computed tomography vertebral segmentation dataset with anatomical variations and multi-vendor scanner data. *Sci Data*, 8(1), 284. doi:10.1038/s41597-021-01060-0
- Loffler, M. T., Jacob, A., Scharr, A., Sollmann, N., Burian, E., El Hussein, M., Sekuboyina, A., Tetteh, G., Zimmer, C., Gempt, J., Baum, T., & Kirschke, J. S. (2021). Automatic opportunistic osteoporosis screening in routine CT: improved prediction of patients with prevalent vertebral fractures compared to DXA. *Eur Radiol*. doi:10.1007/s00330-020-07655-2
- Loffler, M. T., Sekuboyina, A., Jacob, A., Grau, A. L., Scharr, A., El Hussein, M., Kallweit, M., Zimmer, C., Baum, T., & Kirschke, J. S. (2020). A Vertebral Segmentation Dataset with Fracture Grading. *Radiol Artif Intell*, 2(4), e190138. doi:10.1148/ryai.2020190138
- McCulloch, J. A., & Waddell, G. (1980). Variation of the lumbosacral myotomes with bony segmental anomalies. *J Bone Joint Surg Br*, 62-B(4), 475-480. doi:10.1302/0301-620X.62B4.7430228
- O'Driscoll, C. M., Irwin, A., & Saifuddin, A. (1996). Variations in morphology of the lumbosacral junction on sagittal MRI: correlation with plain radiography. *Skeletal Radiol*, 25(3), 225-230. doi:10.1007/s002560050069
- Park, S. K., Park, J. G., Kim, B. S., Huh, J. D., & Kang, H. (2016). Thoracolumbar junction: morphologic characteristics, various variants and significance. *Br J Radiol*, 89(1064), 20150784. doi:10.1259/bjr.20150784
- Payer, C., Stern, D., Bischof, H., & Urschler, M. (2019). Integrating spatial configuration into heatmap regression based CNNs for landmark localization. *Med Image Anal*, 54, 207-219. doi:10.1016/j.media.2019.03.007
- Rabau, O., Smorgick, Y., Tal, S., Tamir, E., Levshin, M., Mirovsky, Y., & Anekstein, Y. (2021). Association between lumbosacral transitional vertebrae and spinal pathologies based on T2 whole-spine sagittal magnetic resonance imaging. *Skeletal Radiol*, 50(12), 2503-2508. doi:10.1007/s00256-021-03809-5
- Sekuboyina, A., Hussein, M. E., Bayat, A., Loffler, M., Liebl, H., Li, H., Tetteh, G., Kukacka, J., Payer, C., Stern, D., Urschler, M., Chen, M., Cheng, D., Lessmann, N., Hu, Y., Wang, T., Yang, D., Xu, D., Ambellan, F., Amiranashvili, T., Ehlke, M., Lamecker, H., Lehnert, S., Lirio, M., Olaguer, N. P., Ramm, H., Sahu, M., Tack, A., Zachow, S., Jiang, T., Ma, X., Angerman, C., Wang, X., Brown, K., Kirszenberg, A., Puybareau, E., Chen, D., Bai, Y., Rapazzo, B. H., Yeah, T., Zhang, A., Xu, S., Hou, F., He, Z., Zeng, C., Xiangshang, Z., Liming, X., Netherton, T. J., Mumme, R. P., Court, L. E., Huang, Z., He, C., Wang, L. W., Ling, S. H., Huynh, L. D., Boutry, N., Jakubicek, R., Chmelik, J., Mulay, S., Sivaprakasam, M., Paetzold, J. C., Shit, S., Ezhov, I., Wiestler, B., Glocker, B., Valentinitich, A., Rempfler, M., Menze, B. H., & Kirschke, J. S. (2021). VerSe: A Vertebrae labelling and segmentation benchmark for multi-detector CT images. *Med Image Anal*, 73, 102166. doi:10.1016/j.media.2021.102166
- Sun, J., Chhabra, A., Thakur, U., Vazquez, L., Xi, Y., & Wells, J. (2021). The association of lumbosacral transitional vertebral anomalies with acetabular dysplasia in adult patients with hip-spine syndrome : a cross-sectional

- evaluation of a prospective hip registry cohort. *Bone Joint J*, 103-B(8), 1351-1357. doi:10.1302/0301-620X.103B8.BJJ-2020-2481.R1
- Thawait, G. K., Chhabra, A., & Carrino, J. A. (2012). Spine segmentation and enumeration and normal variants. *Radiol Clin North Am*, 50(4), 587-598. doi:10.1016/j.rcl.2012.04.003
- Tins, B. J., & Balain, B. (2016). Incidence of numerical variants and transitional lumbosacral vertebrae on whole-spine MRI. *Insights Imaging*, 7(2), 199-203. doi:10.1007/s13244-016-0468-7
- Tureli, D., Ekinci, G., & Baltacioglu, F. (2014). Is any landmark reliable in vertebral enumeration? A study of 3.0-Tesla lumbar MRI comparing skeletal, neural, and vascular markers. *Clin Imaging*, 38(6), 792-796. doi:10.1016/j.clinimag.2014.05.001
- Vinha, A., Bartolo, J., Lemos, C., Cordeiro, F., & Rodrigues-Pinto, R. (2022). Lumbosacral transitional vertebrae: prevalence in a southern European population and its association with low back pain. *Eur Spine J*. doi:10.1007/s00586-022-07415-4
- Wigh, R. E. (1980). The thoracolumbar and lumbosacral transitional junctions. *Spine (Phila Pa 1976)*, 5(3), 215-222. doi:10.1097/00007632-198005000-00003
- Yun, S., Park, S., Park, J. G., Huh, J. D., Shin, Y. G., & Yun, J. H. (2018). Spinal Enumeration by Morphologic Analysis of Spinal Variants: Comparison to Counting in a Cranial-To-Caudal Manner. *Korean J Radiol*, 19(6), 1140-1146. doi:10.3348/kjr.2018.19.6.1140