

Diplomarbeit

**Klinische Evaluierung semiautomatischer bild-
basierter Unterkiefersegmentierung mithilfe eines
open-source Softwarealgorithmus**

**Möglichkeit der praktischen Durchführbarkeit einer
neuen Vorgehensweise**

eingereicht von

Dr.med. univ. Reinhard Potz

zur Erlangung des akademischen Grades

Doktor der Zahnheilkunde

(Dr. med. dent.)

an der

Medizinischen Universität Graz

ausgeführt am 17.01.2018

an der klinischen Abteilung für Mund-, Kiefer- und Gesichtschirurgie

unter der Anleitung von

DDr. Jürgen Wallner

Priv.Doz.DDr. Wolfgang Zemann

Graz, 17.01.2018

Eidesstattliche Erklärung

Ich erkläre ehrenwörtlich, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst habe, andere als die angegebenen Quellen nicht verwendet habe und die den benutzten Quellen wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Graz, am 17.01.2018

Reinhard Potz eh.

1 Danksagungen

Diese Seite möchte ich nutzen um mich bei all jenen Personen zu bedanken, die mich im Laufe meines Studiums, sowie bei der Anfertigung der Diplomarbeit begleitet und unterstützt haben.

Ein besonderer Dank gilt meinem Diplomarbeitbetreuer DDr. Jürgen Wallner, der mir zur jeder Zeit mit seiner fachlichen Kompetenz und seiner außerordentlichen Hilfsbereitschaft zur Seite stand. Es war mir eine große Freude an dieser vorliegenden Studie mitwirken zu können.

Ein Dank gilt auch Priv. Doz. DDr. Wolfgang Zemann für die Unterstützung der Studie und die hilfreichen fachlichen Inputs.

Ein tiefer Dank gilt meiner Familie, insbesondere meinen Eltern, Dagmar und Erich Potz, für eure fortwährende emotionale und finanzielle Unterstützung. Danke an meine Brüder, Christian und Thomas Potz, für eure stete Motivation und eure hilfreichen Ratschläge. Euer Rückhalt hat mir viel Kraft gegeben.

Ein großer Dank geht an meine Verlobte Lisa für ihr Verständnis während des Studiums und ihre unablässige Unterstützung in allen Bereichen.

Zuletzt ein Dank an meine Freunde und Kollegen/innen, die mich schon im Humanmedizin Studium begleitet haben und die die Zeit im Studium und in der Freizeit bunter gestalteten.

2 Zusammenfassung

Einleitung: Computerunterstützte Softwareprogramme zur bildbasierten Segmentierung von radiologischen Computertomographie (CT) oder Magnetresonanztomographie (MRT) Datensätzen finden in vielen Teilbereichen der Medizin zunehmend Verwendung. Bildbasierte Segmentierungsmethoden bilden unter anderem die Grundlage für die Visualisierung von virtuellen, dreidimensionalen (3D) anatomischen Strukturen oder für die Herstellung von realen 3D Modellen. Im Bereich der Mund-, Kiefer- und Gesichtschirurgie beziehen sich die Einsatzgebiete dieser Methode vor allem auf die Unterstützung in der Diagnosestellung, in der chirurgischen Planung oder in der Bilddatenanalyse, insbesondere bei komplexen chirurgischen Fällen. Für die bildbasierte Segmentierung von anatomischen Strukturen stehen verschiedene Softwarealgorithmen zur Verfügung. Viele davon sind im Rahmen von lizenzgeschützten Softwarepaketen mit hohem finanziellem Aufwand kommerziell zu erwerben. Weiters ist der Einsatz dieser Softwarepakete in der klinischen Routine oft nicht funktionsstabil, mit hohem Arbeitsaufwand verbunden oder benötigt zusätzliches Personal. Zudem fallen laufende kostenpflichtige Updates und/ oder teure Hardwareerweiterungen an. Das Ziel dieser Studie war daher, eine lizenzfreie, konventionell erhältliche und frei zugängliche (open-source) Segmentierungsmethode in Bezug auf die Genauigkeit, Funktionalität und potentielle Anwendbarkeit im klinischen Bereich zu testen.

Material und Methoden: In dieser retrospektiven, randomisierten, kontrollierten Studie wurden die Genauigkeit und die Funktionalität einer konventionellen, lizenzfreien semiautomatischen Segmentierungsmethode (GrowCut), im Vergleich zur echten visualisierten anatomischen Struktur derselben Region (Ground Truth) überprüft. Diese Ground Truth Modelle wurden durch zweifache manuelle schichtweise Segmentierung auf Grundlage von CT-Datensätzen erstellt, um eine entsprechende Kontrollgruppe zu generieren. Für eine vergleichende Beurteilung der open-source Segmentierungsmethode durch den GrowCut-Algorithmus wurden 10 CT-Datensätzen des Unterkiefers aus der klinischen Routine zufällig ausgewählt. Die Auswahl wurde anhand von definierten Ein- und Ausschlusskriterien durchgeführt. Die beiden Gruppen (Segmentierungsgruppe-GrowCut und Ground Truth Kontrollgruppe) wurden untereinander anhand objektiv

messbarer Parameter mit Hilfe von statistischen Methoden verglichen: Zeit der Segmentierung, Dice-Score-Koeffizient, Hausdorff- Distanz, Volumen, Anzahl der Voxel.

Ergebnisse: Die Zeit für die vollständige semiautomatische Segmentierung durch den GrowCut-Algorithmus betrug für die anatomische Struktur eines kompletten Unterkiefers rund 1 Minute. Vergleicht man die Übereinstimmung der Segmentierung durch die semiautomatische Methode (Segmentierungsgruppe) mit der der manuellen segmentierten Ground Truth derselben anatomischen Struktur (Kontrollgruppe), zeigt sich ein mittlerer Dice-Score Wert von über 85% und im Schnitt eine Hausdorff-Distanz von unter 33.5 Voxel. Die ermittelten Parameter zwischen den beiden Gruppen waren statistisch nicht signifikant unterschiedlich ($p < 0.05$). Der Korrelationskoeffizient war für sämtliche Vergleiche der beiden Gruppen untereinander nahe dem Wert 1 ($r > 0.94$).

Diskussion: Die in dieser retrospektiven, kontrollierten Studie vorgestellte, lizenzfreie, konventionelle, semiautomatische Segmentierungsmethode mit Hilfe des GrowCut-Algorithmus liefert vielversprechende Ergebnisse in Bezug auf die Genauigkeit, Funktionalität und eine potentielle klinische Anwendung. Die Unterstützung von Diagnosestellungen und chirurgischen Therapieplanungen sowie die Erstellung von virtuellen 3D Visualisierungen oder gedruckten 3D Modellen könnten mit der untersuchten Methode mit hoher Genauigkeit durchgeführt werden. Durch den geringen Zeitaufwand und die Einfachheit der Segmentierungsmethode wäre die Anwendung ohne zusätzlichen Mehraufwand durchführbar. Im Bereich der Mund- Kiefer- und Gesichtschirurgie stellt der untersuchte Algorithmus eine frei zugängliche und schnell verfügbare Alternative zur Bildsegmentierung im Vergleich zu kommerziell erhältlichen Softwarepaketen dar. Zusätzliche monetäre Aspekte oder lizenzbasierte Einschränkungen sind durch die Anwendung dieser open-source Segmentierungsmethode nicht zu erwarten. Dies steigert sowohl die Funktionalität als auch die Praktikabilität in einer möglichen klinischen Anwendung. Zusätzlich können die generierten Ground Truth Daten aus dieser Studie für die Prüfung von anderen bildbasierten Segmentierungsmethoden verwendet werden.

3 Abstract

Introduction: Computer-aided software programs for segmentation processes of computer tomography or magnetic resonance tomography datasets are an increasing topic of interest in many areas of medicine. Image-based segmentation methods form the basis for the visualization of virtual three-dimensional (3D) anatomical structures or for the production of real 3D models. In the field of oral and maxillofacial surgery this methods can help in assistance of diagnosis, in surgical planning or image data analysis. Various software algorithms for the image-based segmentation of anatomical structures are available, but many of them are commercially available, not license-free and with high financial costs. The use of these software packages is often not functional stable in clinical practice and requires a great deal of work or additional staff. In addition updates and / or expensive hardware extensions sometimes have to be paid. The aim of this study was therefore to test a license-free, conventionally available and freely accessible segmentation method with regard to accuracy, functionality and potential clinical practice.

Methods and material: In this retrospective, controlled trial the accuracy and practicability of a conventional, license-free semiautomatic segmentation method (GrowCut) was compared to the true anatomical structure of the same region (Ground Truth). The Ground Truth was created by two-fold manual slice-by-slice segmentation of the datasets to generate an appropriate control group. For a comparative assessment of the open-source segmentation method by the GrowCut algorithm, 10 CT datasets of the mandible were randomly selected. All this datasets were collected during the clinical routine. The selection was made on the basis of defined inclusion and exclusion criteria. The two groups (Segmentation Group - GrowCut and Ground Truth Control Group) were compared with each other using objectively measurable parameters using statistical methods: time of segmentation, Hausdorff distance, Dice-score coefficient, volume and number of voxels.

Results: The time for semiautomatic segmentation of a complete mandible by the GrowCut algorithm was approximately 1 minute. An average Dice-score of over 85% and a Hausdorff distance of less than 33.5 voxels could be achieved by comparing the accordance of the segmentation by the semiautomatic method

(segmentation group) with the segmentation of the manual segmented ground-truth of the same anatomical structure (control group).

The statistically determined parameters were not significantly different between the two groups ($p < 0.05$). The correlation coefficient for all comparisons between the two groups was close to 1 ($r > 0.94$).

Discussion: The license-free, conventional available semiautomatic segmentation method presented in this study using the GrowCut algorithm provides good results in accuracy, functionality and potential clinical practice. The support of diagnoses and surgical therapy planning, as well as 3D virtual visualizations for the patient and the creation of real printed 3D models would be possible with high accuracy. Due to the simple implementation and the less need of time using this segmentation method, the application in the clinical area would be possible without additional effort. In the field of Oral and Maxillofacial Surgery the algorithm represents a freely available and open-sourced alternative to image segmentation, compared to commercially available software packages. Additional monetary or license-based restrictions are not to be expected with this open-source segmentation method. This enhances both functionality and practicability in a potential clinical application. In addition, the generated ground truth data from this study can be used to test other image-based segmentation methods.

4 Inhaltsverzeichnis

1	Danksagungen	iii
2	Zusammenfassung.....	iv
3	Abstract	vi
4	Inhaltsverzeichnis.....	viii
5	Glossar und Abkürzungen	9
6	Abbildungsverzeichnis.....	10
7	Tabellenverzeichnis.....	11
8	Einleitung.....	12
8.1	Hintergrund und Relevanz des Sachverhaltes.....	12
8.2	Grundlagen und Definitionen.....	16
8.2.1	Medizinische Bildbearbeitung.....	16
8.2.2	Merkmalsextraktion	17
8.2.3	Bildbasierte Segmentierung	17
8.2.4	Hybride Segmentierungsverfahren	19
8.2.5	Bildbasierte Segmentierung mit Hilfe des GrowCut-Algorithmus	19
8.2.6	3D Slicer.....	20
8.2.7	Mevislab	20
8.2.8	Ground Truth.....	21
8.3	Zielsetzung.....	22
9	Material und Methoden.....	23
9.1	Studienablauf und Datenselektion.....	23
9.2	Zielgrößen und Endpunkte.....	27
9.3	Statistische Methoden	29
10	Ergebnisse	30
11	Diskussion.....	43
12	Zusammenfassung der wichtigsten Fakten der Arbeit	54
13	Literaturverzeichnis.....	55

5 Glossar und Abkürzungen

2D... zweidimensional

3D... dreidimensional

bzw... beziehungsweise

CBCT... Cone-beam-computed-tomography; Kegelstrahl- Computertomographie

CT... Computertomographie

DICOM... Digital Imaging and Communications in Medicine

DSC... Dice-Score-Koeffizient;

Ground Truth... Grundwahrheit; tatsächliche Kartographie der echten visualisierten 3D Struktur; Kontrollgruppe

DVT... Digitale-Volumentomographie

HD... Hausdorff-Distanz

HE... Hounsfield Einheit

Max... Maximum

Mevislab... MeVis Medical Solutions AG, lizenzfreie medizinische Plattform zur Bildbearbeitung

Min... Minimum

mm³... Kubikmillimeter

MRT... Magnetresonanztomographie

MUG... Medizinische Universität Graz

μ... Mü

o.ä... oder ähnliche(s)

OP... Operation

p... Signifikanz, Signifikanzniveau

PACS... Picture Archiving and Communications Systems

Pearson Korrelationskoeffizient... Produkt-Moment-Korrelationskoeffizient

Pixel... Bildpunkt

r... Pearson-Korrelationskoeffizient

u.ä... und ähnliche

σ... Standardabweichung

6 Abbildungsverzeichnis

Abbildung 1 Vorgang der semiautomatischen Segmentierung _____	13
Abbildung 2 Framework Mevislab _____	25
Abbildung 3 Segmentierungsvorgang Mevislab _____	26
Abbildung 4 Schichtweise Segmentierung _____	26
Abbildung 5 Dice-Score Darstellung _____	28
Abbildung 6: Visualisierung der Segmentierungen _____	41
Abbildung 7: Genauigkeitsbewertung der Ground Truth _____	41

7 Tabellenverzeichnis

Tabelle 1 Ein- und Ausschlusskriterien _____	24
Tabelle 2 Manuelle Segmentierungen im direkten Vergleich _____	30
Tabelle 3 Manuelle Segmentierungsergebnisse mit Min, Max, $\sigma \pm \mu$ _____	31
Tabelle 4 Ergebnisse der Ground Truth A und des GrowCut-Algorithmus _____	32
Tabelle 5 Vergleich Ground Truth B mit GrowCut-Algorithmus _____	33
Tabelle 6 Anzahl der Voxel Ground Truth (A, B), GrowCut _____	34
Tabelle 7 Vergleich der Ergebnisse Ground Truth (A,B) vs. Algorithmus _____	35
Tabelle 8 Volumen Vergleichsberechnungen _____	36
Tabelle 9 Voxel Vergleichsberechnungen _____	36
Tabelle 10 Regressionsanalyse Volumen _____	37
Tabelle 11 Regressionsanalyse Voxel _____	38
Tabelle 12 Dice-Score Werte _____	39
Tabelle 13 Hausdorff-Distanzen (HD, Voxel) Werte _____	40

8 Einleitung

8.1 Hintergrund und Relevanz des Sachverhaltes

Digitale Bilder, die aus routinemäßig durchgeführten radiologischen Computertomographie (CT)- und/oder Magnetresonanztomographie (MRT)-Aufnahmen entstehen, bestehen aus Pixel oder Voxel. Diesen Pixel oder Voxel sind bestimmte Farb-, Kontrast- oder Helligkeitswerte zugeordnet. Durch ein einheitliches Format kann das radiologisch entstandene Bild über Kommunikationsnetzwerke an verschiedenen Orten gleichzeitig begutachtet und bearbeitet werden. Bildverarbeitungsprogramme erweitern dabei die Möglichkeiten für den/die Behandler/in in den verschiedensten medizinischen Fachgebieten. Die medizinische Bildverarbeitung beinhaltet verschiedene Teilbereiche. Das Bild muss zunächst aufgenommen und in eine digitale Bildmatrix umgewandelt werden. Wir sprechen von der Bilderzeugung. Auf diese folgt die Bilddarstellung, die von der Beleuchtung, Schattierung und anderen Faktoren abhängt. Sie dient der Ausgabe des Bildes. Eine Bildspeicherung ist zur Archivierung und effizienten Datenübertragung notwendig. Als letzter, aber immer wichtiger werdender Punkt, muss die Bildauswertung genannt werden. Hierzu zählen beispielsweise die bildbasierte Merkmalsextraktion und die bildbasierte Segmentierung, auf die in Kapitel 8.2 und 8.3 genauer eingegangen wird. Diese Vorgänge werden computerunterstützt mit einer speziellen Software durchgeführt. (1)

Aufgrund der voranschreitenden computerunterstützten Entwicklung von dreidimensionalen (3D) Technologien und fortschrittlicher Computer-Software finden Bildbearbeitungsprogramme in den vergangenen Jahren im medizinischen Bereich zunehmend immer mehr Verwendung. Dies trifft vor allem auf die westlichen Industrieländer zu. Die Anwendungsgebiete sind vielfältig und spielen unter anderem in der Unterstützung der Diagnosestellung, der Datenanalyse oder der chirurgischen Planung eine wichtige Rolle. (2) Die Verwendung einer computerbasierten Software zur Analyse von CT- und/oder MRT-Bildern zur Unterstützung der Diagnosestellung oder Behandlungsplanung ist in vielen Teilbereichen der Medizin heute nicht mehr wegzudenken. (3) Dies trifft vor allem auf die Unterstützung der Therapieplanung in komplexen chirurgischen Fällen zu. Ein wesentlicher Bestandteil der Bildbearbeitung ist die Bildsegmentierung. Unter Bildsegmentierung versteht man, dass bildlich zusammenhängende Bereiche

voneinander digital unterschieden werden, indem Pixel, Kanten, Regionen und weitere Strukturen beurteilt und entweder dem einen oder dem anderen Segment zugeordnet werden. (4) Klinisch finden im Bereich der Mund-, Kiefer- und Gesichtschirurgie derartige computerunterstützte Softwareprogramme zur Bildsegmentierung vor allem bei komplexen Trümmerbrüchen des Gesichtsschädels, angeborenen Fehlbildungen oder skelettalen Fehlstellungen von Ober- und/oder Unterkiefer Anwendung.

Grundsätzlich unterscheidet man zwischen manueller und automatischer Segmentierung. Letztere kann zusätzlich in eine vollautomatische und semiautomatische Segmentierung untergliedert werden. Semiautomatische Segmentierung bedeutet, dass der/die Benutzer/in die Software aktiv unterstützt, indem er/sie die zu segmentierenden Regionen und Bereiche, die im Hintergrund liegen, grob markiert (siehe Abbildung 1). Die anschließende Segmentierung erfolgt automatisch durch einen definierten Algorithmus. Die semiautomatische Segmentierung ist demnach eine Kombination aus der vollautomatischen und der manuellen Segmentierung. (4–6)

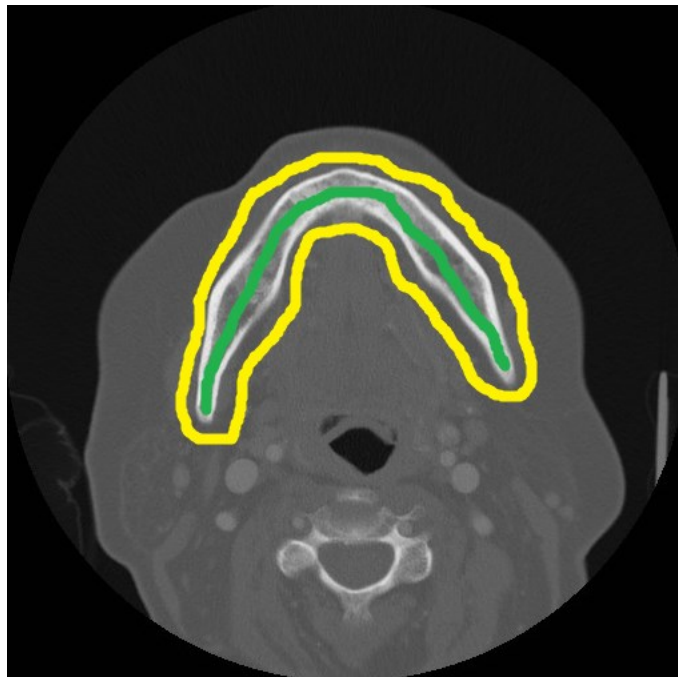


Abbildung 1 Vorgang der semiautomatischen Segmentierung:

Der Unterkiefer wird dabei grob markiert (grün). Anschließend markiert der/die Benutzer/in die Umgebung (gelb). Nun kann der Algorithmus die Segmentierung selbstständig durchführen.

(Beispielbild CT Unterkiefer (7))

Für ein computerunterstütztes Segmentierungsverfahren muss also zuvor ein definierter Algorithmus vorhanden sein. Dieser folgt, je nach Programmierung, einem Pixel-, Kanten-, Regionen- oder Texturverfahren. Auch eine Kombination aus diesen Verfahren ist möglich. So können Objekte von ihrem Hintergrund in relativ kurzer Zeit getrennt werden. (4–6)

Die Segmentierung eines Bildobjektes, obgleich sie vollautomatisch oder semiautomatisch läuft, stellt in Bezug auf die Genauigkeit und Funktionalität eine Herausforderung für den Algorithmus und das gesamte Softwarepaket dar und ist in der Medizin aufgrund der biologischen und pathologischen Vielfalt der anatomischen Strukturen als besonders schwierig zu betrachten. Vollautomatische Segmentierungsverfahren liefern oftmals unzureichende Ergebnisse, sodass eine zusätzliche manuelle Nachbearbeitung notwendig ist. (8),(9) Funktionsstabilen Algorithmen zur Abgrenzung anatomischer Strukturen und anderer Regionen wird daher großes Interesse zugeschrieben.

Ein wesentlicher Vorteil der computergestützten (semi)automatischen Segmentierung ist der Zeitfaktor. Eine große Anzahl von Fällen bzw. radiologischen Datensätzen kann bei entsprechender Funktionsstabilität der Segmentierung auf diese Weise mit hoher Genauigkeit und in kurzer Zeit bearbeitet werden. (3)

Gerade im Bereich der Mund-, Kiefer- und Gesichtschirurgie gibt es viele komplexe klinisch relevante chirurgische Fälle, die eine sorgfältige Vorbereitung und Planung benötigen. Dreidimensionale Aufnahmen vor schwierigen chirurgischen Eingriffen sind mittlerweile Standard. In den letzten Jahrzehnten haben sich zahlreiche wissenschaftliche Zentren und große Firmen darauf spezialisiert, geeignete Softwarepakete zur Segmentierung medizinischer Bilder und Bilddatensätze zu entwickeln, mit dem Ziel eine virtuelle Operationsplanung zu ermöglichen. (10)

Mit verschiedenen Software Tools kann eine bestimmte anatomische Struktur aus dem vorliegenden radiologischen Bilddatensatz virtuell dreidimensional dargestellt und bearbeitet werden. Das auf diese Weise entstehende virtuelle 3D Modell kann anschließend bei Bedarf ebenfalls gedruckt werden. (11) Der/die Chirurg/in hat die Möglichkeit, präoperativ verschiedene Operationsmöglichkeiten am segmentierten 3D Modell zu simulieren und zu vergleichen. Weiters kann dem/der Patienten/in das Operationsgebiet und - je nach Eingriff - auch die Operation visualisiert werden. Sofern die Software funktionsstabil ist, schafft sie die Grundlage eines

möglicherweise weniger invasiven Eingriffs durch genauere OP Planung, eine kürzere Operationsdauer oder ein verbessertes Operationsergebnis. (12) Die Anwendungsgebiete sind vielfältig und reichen von virtuell erstellten 3D Modellen bis zu anatomisch angepassten, individuell gestalteten, Schädelknochen-Implantaten, beispielsweise nach Gesichtstraumata oder bei Frakturen der Augenhöhle. (6)

Bislang wird in der Literatur eine Vielzahl an konventionell verfügbaren semiautomatischen und automatischen Methoden zur Segmentierung von Bilddateien beschrieben und laufend evaluiert. (13) Dennoch kommen semi- oder vollautomatische Segmentierungsalgorithmen in vielen klinischen Einrichtungen nicht routinemäßig für die Unterstützung von Diagnosestellungen oder Behandlungsverfahren zum Einsatz. Die Gründe dafür sind unter anderem eine mangelhafte Funktionsstabilität der zur Verfügung stehenden Softwarepakete, eine teilweise unzureichende Rechenleistung der Hardware oder zeitraubende und komplexe Bedienungsprozesse der Software. Ein Mangel an benutzerfreundlichen Oberflächen sowie industriegetriebene hohe Kosten für Lizenzgebühren sind weitere Erklärungen, warum zahlreiche kommerzielle bildbasierte Segmentierungspakete in der täglichen klinischen Routine nicht ubiquitär zur Anwendung kommen. (14)

Neben käuflich erwerblichen lizenzgestützten Segmentierungsmethoden, die in der Regel in großen Softwarepaketen angeboten werden, existieren zudem lizenzfreie Algorithmen, die für einfache Anwendungen konstruiert wurden. Über die Funktionalität, Genauigkeit und klinische Praktikabilität dieser frei zugänglichen Segmentierungs-Algorithmen ist in der Literatur bislang weniger bekannt. Vor allem jedoch wurde die Anwendung von frei verfügbaren Segmentierungsmethoden auf einer open-source Grundlage im Bereich der Mund,- Kiefer- und Gesichtschirurgie unzureichend untersucht. Die Genauigkeit und praktische Funktionalität eines lizenzfreien open-source Algorithmus zur Segmentierung von Bilddatensätzen im Bereich des Gesichtsschädels wurde bislang ebenfalls nicht objektiv beurteilt.

8.2 Grundlagen und Definitionen

8.2.1 Medizinische Bildbearbeitung

In der Medizin sind digitale Bildaufnahmen wie die Computer- oder die Magnetresonanztomographie aus der klinischen Routine unverzichtbar. Dabei spielt die Möglichkeit einer qualitativ hochwertigen und wenig zeitaufwendigen Bildbearbeitung für weiterführende Diagnosestellungen oder Behandlungsplanungen ebenfalls eine große Rolle.

Ein digitales Bild setzt sich aus einzelnen Bildpunkten (Pixeln) zusammen, denen bestimmte Helligkeits- oder Farbwerte zugeordnet werden. Mit diversen Softwareprogrammen kann dieses bearbeitet oder über Kommunikationsnetze wie beispielsweise PACS (Picture Archiving and Communications Systems (15)) an verschiedenen Orten gleichzeitig zur Verfügung gestellt werden.

Eine wesentliche Grundlage in der medizinischen Bildbearbeitung spielt die computerunterstützte bildbasierte Segmentierung mit Hilfe von Softwarealgorithmen. Dabei werden benachbarte Pixel oder Voxel einem Homogenitätskriterium entsprechend zusammengefasst und unterschiedliche Strukturen erkannt. Dies kann sowohl manuell als auch automatisch geschehen. Für die automatische Segmentierung gibt es unterschiedliche Ansätze (1):

- Pixelorientierte Verfahren
- Kantenorientierte Verfahren
- Regionenorientierte Verfahren
- (Modellbasierte Verfahren)
- Texturorientierte Verfahren

Für jedes dieser Verfahren muss zuerst die Ebene der jeweiligen Pixel betrachtet werden. Die Pixelebene ist die „Ansammlung“ von definierten Bildpunkten. Die Kantenebene spiegelt eindimensionale Strukturen wider, während die Texturebene zweidimensionale Strukturen ohne eine sichtbare Umrandung „repräsentiert“. Anders ist es bei der Regionenebene, die für die zweidimensionalen Strukturen eine Umrandung benötigt. (16)

8.2.2 Merkmalsextraktion

Die Vorstufe der bildbasierten Segmentierung in der intelligenten Bildauswertung ist die Merkmalsextraktion. Bildinformationen (Merkmale) werden auf den Ebenen, auf denen Algorithmen arbeiten, betont. Je nach Verfahren werden zusammenhängende Merkmale wie Pixel, Kanten, Regionen (=homogene Bereiche) oder eine einheitliche Textur erkannt. Der/die Untersucher/in kann die Software beim Segmentierungsvorgang unterstützen. Gerade in der medizinischen Bildbearbeitung spielt die Erfahrung des/der Untersuchers/in eine wesentliche Rolle. Die Schwierigkeit bei der automatischen Bildbearbeitung besteht darin, die Kenntnisse eines/r erfahrenen Untersuchers/in in einen automatischen Bildbearbeitungsalgorithmus zu integrieren. In der Literatur wird dieses Phänomen als „Semantic Gap“, also die „Diskrepanz zwischen der kognitiven Interpretation eines Bildes durch den menschlichen Betrachter (high level) und den pixelbasierten Merkmalen, mit denen Computerprogramme ein Bild repräsentieren (low level)...“, beschrieben. Gerade bei medizinischen Bildern ist es besonders schwierig, den Semantic Gap möglichst gering zu halten. Menschliche Organe stellen sich aufgrund der individuellen Anatomie von Patient/in zu Patient/in unterschiedlich dar. Sie können zu verschiedenen Zeitpunkten an Größe und Form variieren oder auch gelegentlich Anomalien aufweisen. Ebenso lassen sich Gewebegrenzen vor allem in Weichgeweben oftmals schwer digital darstellen. Dies verkompliziert die Programmierung eines ubiquitär anwendbaren funktionsstabilen Segmentierungsalgorithmus. Meist ist eine Kombination aus verschiedenen bildbasierten Segmentierungsverfahren notwendig. (1)

Zudem gilt im medizinischen Bereich ein hoher Anspruch für die Anwendung von bildbasierten Segmentierungsmethoden bzw. -Algorithmen der Zuverlässigkeit, Funktionalität und Genauigkeit.

8.2.3 Bildbasierte Segmentierung:

Unter bildbasierter Segmentierung versteht man die Unterteilung eines Bildes in örtlich zusammenhängende Regionen mithilfe der Merkmalsextraktion. In der medizinischen Bildbearbeitung sollen anatomische Strukturen automatisch erkannt und unterschieden werden. (1)

Pixelbasierte Verfahren arbeiten mit den Farbwerten, beziehungsweise Grauwerten, einzelner Pixel. Die Umgebung des Pixels wird dabei nicht

berücksichtigt. Der Nachteil ist, dass lediglich eindimensionale Werte bestimmt werden und eine zusammenhängende Segmentierung mit allein diesem Verfahren nicht möglich ist. Eine Nachbearbeitung ist hier in jedem Fall erforderlich. (1, 17)

Bei *Kantenorientierten Verfahren* kommt es auf den lokalen Kontrast benachbarter Pixel an, das heißt Bereiche, in denen sich die Grauwerte der nebeneinanderliegenden Pixel unterscheiden, werden erkannt. Dieses Verfahren ist für die vollautomatische Segmentierung daher nur bei Bildern möglich, die Objekte mit klar definierten Umrandungen beinhalten. Bei medizinischen Bildern ist dies selten der Fall, ein Beispiel würden metallische Materialien oder Implantate darstellen. Eine kantenorientierte Segmentierung ist daher nur bei der halbautomatischen Segmentierung sinnvoll. Dabei bestimmt der/die Benutzer/in einige Stützpunkte am Rand des zu segmentierenden Objektes, der Algorithmus vergleicht die Grauwerte des Objektes mit den Grauwerten der umgebenden Regionen und legt eine dünne Linie um das zu segmentierende Objekt. Der/die Benutzer/in kann nun die Korrektheit direkt überprüfen und gegebenenfalls korrigieren. (1)

Texturbasierte beziehungsweise texturanalytische Verfahren sind eine Herausforderung für die automatische Bildbearbeitung. Es wird versucht eine Homogenität einer unregelmäßigen Struktur, die sich subjektiv als regelmäßig darstellt, zu quantifizieren. (1, 18) Pixel und Bildregionen werden beim texturorientierten Algorithmus auf charakteristische Merkmale hin analysiert. Pixel, die eine ähnliche Textur aufweisen, werden demselben Bildsegment zugeordnet. (19, 20) Die Textur kann sich hinsichtlich der Farb- beziehungsweise Grauwerte, der räumlichen Anordnung, der Rauigkeit, der Feinheit sowie der Körnung der Oberfläche unterscheiden. (19) Sie werden meist ergänzend zu anderen Methoden eingesetzt.

Bei *regionenbasierten Verfahren* wird mit dem Homogenitätsprinzip gearbeitet. Gruppen von Bildpunkten, die sich entsprechend eines Kriteriums als homogen erweisen, werden zusammengefasst. Die entstehenden Regionen sollen die Fläche des Objektes darstellen. (21) Beispiele hierfür sind das Region-Growing in der zweidimensionalen oder das Volume-Growing in der dreidimensionalen Bildbearbeitung. Beide Verfahren zählen zu den agglomerativen Verfahren und sind eine Untergruppe der regionenbasierten Segmentierung. Es werden zunächst Saatpixel (2-dimensionales Bild) / Voxel (3-dimensionales Bild) automatisch oder interaktiv gewählt. Diese werden mit den Nachbarpixel oder Voxel verglichen.

Liegt der Vergleichswert oder die Distanz unter einer definierten Schwelle, wird der Pixel/Voxel dem ursprünglichen Keimpunkt zugeordnet. Dies erfolgt solange, bis eine Verschmelzung nicht mehr möglich ist. Das Ergebnis hängt von der Anzahl der ursprünglich definierten Keimpunkte, dem definierten Vergleichswert, der definierten Distanz sowie der definierten Schwelle ab. (1, 22)

Neben dem agglomerativen Verfahren zählt auch das divisive Verfahren zur regionenbasierten Segmentierung. Regionen werden durch vertikale und horizontale Trennlinien so lange gesplittet, bis die Bereiche ausreichend homogen sind. Saatpunkte müssen nicht definiert werden, da das Splitting das gesamte Bild betrifft. (1)

8.2.4 Hybride Segmentierungsverfahren

Aufgrund der Fehleranfälligkeit einzelner Verfahren werden in der medizinischen Bildbearbeitung Segmentierungsverfahren meist kombiniert. Bei den daraus resultierenden sogenannten hybriden Segmentierungsverfahren werden am häufigsten kanten- und regionenorientierte Algorithmen miteinander verknüpft. Die Nachteile der jeweiligen Verfahren sollen auf diese Weise minimiert werden.

8.2.5 Bildbasierte Segmentierung mit Hilfe des GrowCut-Algorithmus

Der GrowCut¹-Algorithmus ist eine interaktive bildbasierte Segmentierungsmethode, um ein Bildobjekt bzw. eine Struktur von seiner Umgebung zu trennen. Der Algorithmus arbeitet dabei mit zellulären Automaten, denen Pixel zugeordnet sind. Pixel sind Bildpunkte in zweidimensionalen Bildern. Je größer die Zahl der Bildpunkte ist, desto höher ist die Auflösung des Bildes. (23) Jeder zelluläre Automat ist beschriftet und entweder als Objekt (Vordergrund), benachbartes Objekt (Hintergrund) oder als undefinierter Raum gekennzeichnet. Bevor die Segmentierung gestartet werden kann, muss der/die Benutzer/in zunächst einige Initialpunkte manuell bestimmen. Die gekennzeichnete Matrix soll den Bereich im Vordergrund und den Bereich im Hintergrund des Objektes markieren. Nach der Kennzeichnung kann die semiautomatische Segmentierung gestartet werden. Die Pixel werden nun entweder dem Objekt oder dem Hintergrund zugeordnet. (24) Die Methode des GrowCut-Algorithmus ist für diverse Bildgrößen sowie zur bildbasierten Segmentierung von graustufigen aber auch

¹ GrowCut 3.0, www.growcut.com

kolorierten Bilddateien geeignet. Im medizinischen Bereich ist der GrowCut-Algorithmus weiters in der Lage, unscharfe Regionen mithilfe der Expertise eines/r Klinikers/in zu definieren. Er arbeitet schnell, nach einem einfachen Prinzip und gilt zudem als ein zuverlässiges bildbasiertes Segmentierungsverfahren für schwierige 2D und 3D Objekte. (5),(24)

Der/die Anwender/in kann das Ergebnis beeinflussen, indem er/sie während der Segmentierung durch interaktives Eingreifen zusätzliche Bildpunkte bestimmt (semiautomatische Segmentierung). (25)

Der GrowCut-Algorithmus ist eine open-source Segmentierungsmethode, deswegen frei verfügbar und mit frei-zugänglichen Software Plattformen oder Grafikeditoren kompatibel.

8.2.6 3D Slicer

3D Slicer² ist eine open-source Softwareplattform für medizinische Bildinformatik, Bildverarbeitung und dreidimensionale Visualisierung. In die Plattform können bildbasierte Segmentierungsalgorithmen problemlos eingespielt und verwendet werden. Die Software ist frei-konventionell verfügbar und nicht kommerziell. (26)

8.2.7 Mevislab

Mevislab³ ist eine modulbasierte Softwareplattform (modulares Framework), welches für die Bildverarbeitung, Forschung und Entwicklung von Bilddatensoftware eingesetzt werden kann. Neu programmierte Algorithmen können unkompliziert integriert und entsprechend ihres Anwendungsbereiches getestet werden. Die Plattform bietet weitere Softwaremodule für die volumetrische, morphologische und funktionelle Analyse von Bilddatensätzen an. (27)

Mevislab wurde von der Firma MeVis Medical Solutions AG in Zusammenarbeit mit dem Forschungsinstitut Fraunhofer MEVIS für die Testung und Prüfung von unterschiedlichen Softwarepaketen und Algorithmen im klinischen Bereich entwickelt. Mevislab wird ebenfalls nicht kommerziell zum Download angeboten.

² Slicer 4.4.0 Software (Surgical Planning Lab, Harvard Medical School, Harvard University, Boston, USA) (26)

³ MeVisLab 2.5.2. Software (Medical imaging prototyping platform, MeVis Medical Solutions AG, Fraunhofer Institut, Bremen, Deutschland) (27)

8.2.8 Ground Truth

Die Ground Truth ist ein bildbasiertes, segmentiertes, virtuelles 3D Modell, das die echte, tatsächliche, visualisierte Größe oder das Volumen einer bestimmten Struktur hat. Die echte, tatsächliche, visualisierte Größe oder das Volumen richten sich dabei nach der Art des Bilddatensatzes, der die bestimmte Struktur beinhaltet. Die Ground Truth kann mit einer 3D Kartographie (mapmaking) einer bestimmten Struktur verglichen werden. In der medizinischen Visualisierung stellt die Ground Truth bzw. das Ground Truth Modell eine hoch qualitative bzw. elegante Möglichkeit dar, um bildbasierte Softwares bzw. Segmentierungsmethoden zu beurteilen und miteinander zu vergleichen. Dies ist darauf zurückzuführen, dass die Ground Truth als objektiver, valider Vergleichswert beziehungsweise als Kontrollgruppe in wissenschaftlichen Untersuchungen angenommen werden kann. Schwierigkeiten ergeben sich lediglich in der Erstellung von objektiven 3D Ground Truth Modellen, da dabei der Faktor der Subjektivität in der Erstellung vermieden, beziehungsweise so gering wie möglich gehalten werden muss. Objektiv erstellte Ground Truth Modelle sind von hoher Wichtigkeit, da ansonsten ein systematischer Fehler in einem vergleichenden Beurteilungsverfahren mit beispielsweise einem zu testenden Softwareprozess oder einer Segmentierungsmethode resultieren kann. (28–30)

8.3 Zielsetzung

Ziel dieser Arbeit ist es, die Genauigkeit, die Funktionsstabilität (Funktionalität), und die praktisch- klinische Anwendbarkeit einer freizugänglichen bildbasierten open-source Segmentierungsmethode im Bereich der Mund-, Kiefer- und Gesichtschirurgie zu beurteilen. Die Segmentierungsergebnisse sollen mit der echten visualisierten anatomischen Struktur verglichen und anhand dieser 3D Ground Truth Modelle evaluiert und bewertet werden. Die Möglichkeit einer einfachen, schnellen und lizenzfreien Segmentierungsmethode zur Erfassung anatomischer Strukturen aus CT und/oder MRT Datensätzen könnte unter anderem effektive Unterstützung in der Diagnostik und der chirurgischen Therapieplanung liefern.

Dementsprechend wird die Hypothese für diese Studie folgendermaßen definiert: Die Genauigkeit, Übereinstimmung und Funktionalität einer konventionellen, lizenzfreien, semiautomatischen open-source Segmentierungsmethode unterscheiden sich nicht signifikant von der Ground Truth derselben anatomischen Struktur.

9 Material und Methoden

In dieser Studie wurde eine semiautomatische, frei-zugängliche und bildbasierte Segmentierungsmethode an CT-Datensätzen des Unterkiefers getestet. Die Basis der Segmentierung bildete ein open-source Algorithmus. Dieser sollte auf die Genauigkeit, Funktionalität und praktische Anwendbarkeit im Gesichtsschädel für eine klinische Anwendung geprüft werden. Die CT-Datensätze stammten aus der klinischen Routine und wurden aus bereits vorhandenen Datensätzen, die auf Grund einer medizinischen Indikation zur Diagnosesicherung oder Krankheitsabklärung erstellt wurden, ausgewählt.

Die vorliegende Studie wurde von der Ethikkommission der Medizinischen Universität Graz zur Durchführung freigegeben (EK-Nummer 29-143 ex 16/17). Es handelt sich bei dieser Studie um keine Medizinproduktstudie nach dem Medizinproduktegesetz, sondern um eine retrospektive Datenanalyse mit Hilfe einer neuen Methode.

9.1 Studienablauf und Datenselektion

Zunächst wurde für diese kontrollierte, randomisierte, retrospektive Studie ein entsprechendes Datenkollektiv erstellt. Hierfür wurden 45 CT-Datensätze (DICOM-Dateien⁴) vorselektiert, die an der klinischen Abteilung für Mund Kiefer- und Gesichtschirurgie an der Medizinischen Universität Graz im Rahmen der klinischen Routine vorhanden waren. Die Daten wurden im Zeitraum von 2015-2017 akquiriert. Sämtliche Daten wurden vor der Verwendung anonymisiert. Für die Auswahl der 45 CT-Datensätze wurden nur Datensätze mit hoher Auflösung sowie einer Schichtdicke von höchstens 1.0 mm und eine Pixelgröße von maximal 0.25mm herangezogen. Außerdem wurden ausschließlich zahnlose Unterkiefer gewählt, um mögliche Störfaktoren, wie beispielsweise radiologische Artefakte, ausgehend von zahnärztlichen Restaurationen zu eliminieren. Die ausgewählten Unterkieferdatensätze sollten physiologische Strukturen aufweisen und ohne knöcherne Pathologie sein. Atrophierte Unterkiefer kamen für die Selektion genauso in Frage wie nicht-atrophe Unterkiefer. Unvollständige, iatrogen- oder pathologisch bedingt veränderte Unterkiefer (Kiefernekrosen o.ä.) wurden ebenso

⁴ Medizinische Bilder werden zur Archivierung als Dicom-Datei gespeichert. Dicom steht für „Digital Imaging and Communications in Medicine“. (31)

ausgeschlossen wie frakturierte Unterkiefer oder Unterkiefer, die Osteosynthesematerial beinhalteten.

Entsprechend diesen Ein- und Ausschlusskriterien konnten aus den zur Verfügung stehenden 45 CT-Datensätzen 20 Datensätze für diese Studie ausgewählt werden. Mit Hilfe einer automatischen Randomisierungssoftware (Randomizer®⁵) wurden folglich 10 Datensätze aus den vorliegenden 20 zufällig gewählt. Von diesen 10 Datensätzen waren 6 Datensätze männliche Unterkiefer und 4 Datensätze weibliche Unterkiefer.

<i>Einschlusskriterien:</i>	<i>Ausschlusskriterien:</i>
<i>zahnloser Unterkiefer</i>	<i>bezahnte Unterkiefer / Implantate</i>
<i>Alter ≥18 Jahre</i>	<i>Alter < 18 Jahren</i>
<i>Schichtdicke höchstens 1mm</i>	<i>Unterkiefer mit Osteosynthesematerial</i>
<i>Pixelgröße bis maximal 0.25mm</i>	<i>Kiefernekrosen/Kieferfrakturen</i>
<i>physiologische Knochenanatomie</i>	<i>Unvollständige Datensätze</i>

Tabelle 1 Ein- und Ausschlusskriterien

Aus diesen 10 Datensätzen wurde zunächst eine Kontrollgruppe (Ground Truth) definiert. Hierfür musste für jeden der 10 Datensätze die tatsächliche anatomische Knochenstruktur des Unterkiefers manuell schichtweise definiert werden. Um dies zu erreichen, segmentierten zwei klinische Experten für Mund-, Kiefer- und Gesichtschirurgie (A und B), unterstützt durch einen Facharzt für Radiologie, die 10 Datensätze manuell. Auf diese Weise wurde jeder Datensatz zweimal schichtweise manuell segmentiert. Die manuelle Segmentierung erfolgte mithilfe der öffentlich zugänglichen medizinischen Plattform Mevislab. Zur Durchführung musste dafür ein individuelles modulares Framework (siehe Abbildung 2) erstellt werden, welches in die Plattform Mevislab eingespielt wurde. Die Erstellung und Einspielung der Module in Mevislab wurde mit Hilfe der technischen Universität Graz⁶ durchgeführt. Die Unterkieferdatensätze wurden nun von den Klinikern auf axialer Ebene, Schicht für Schicht, zweifach manuell segmentiert, je ein Datensatz von einem klinischen Experten A und B (siehe Abbildung 3 und 4). So konnten zwei virtuelle Ground Truth Modelle bzw. Volumen (Kontrollvolumen, Kontrollmodelle A und B) je Datensatz festgelegt werden.

⁵ Randomizer: Randomisierung für klinische und nicht-klinische Studien; Graz, Austria (32)

⁶ Institut für Computergraphik und Visualisierung, TU Graz; Graz, Österreich

Für die semiautomatische Segmentierung wurde der frei verfügbare und konventionell zugängliche GrowCut-Algorithmus gewählt. Für die semiautomatische Segmentierung mit Hilfe des GrowCut-Algorithmus wurden die ausgewählten Datensätze in das Programm 3D Slicer eingespielt. Die Software 3D Slicer ist ebenfalls frei verfügbar und enthält viele verschiedene Funktionen zur Bearbeitung und Erstellung von 3D Visualisierungen auf Grundlage von medizinischen Bilddatensätzen. Das Ergebnis der semiautomatischen Segmentierungsmethode mithilfe des GrowCut-Algorithmus wurde als virtuelles 3D Modell gespeichert (Segmentierungsmodell bzw. Segmentierungsvolumen). Nach den erfolgreichen manuellen und semiautomatischen Segmentierungen der zufällig ausgewählten Datensätze waren nun 10 Datensätze von je einem klinischen Experten, die die Ground Truth (A, B) definierten sowie weitere 10 Datensätze der semiautomatischen Segmentierung mithilfe des GrowCut-Algorithmus vorhanden.

Für eine entsprechende Beurteilung der open-source Segmentierungsmethode mithilfe des GrowCut-Algorithmus wurden die Datensätzen folgendermaßen untereinander verglichen:

- Algorithmus (GrowCut) : Ground Truth A (Kliniker A)
- Algorithmus (GrowCut) : Ground Truth B (Kliniker B)
- Ground Truth A (Kliniker A): Ground Truth B (Kliniker B)

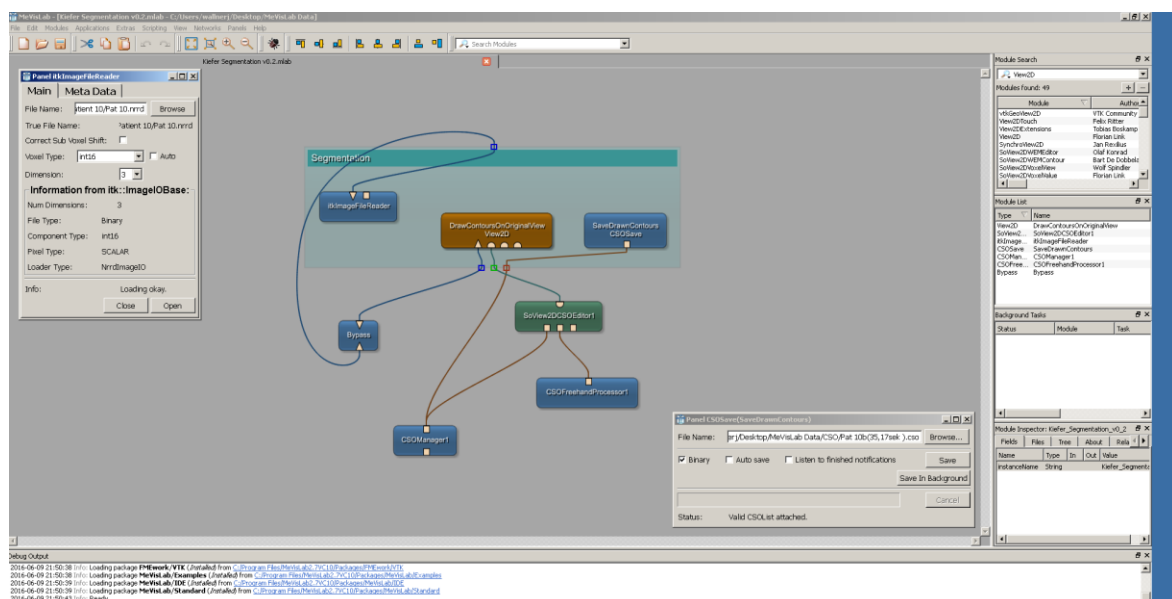


Abbildung 2 Framework Mevislab. Die Abbildung zeigt das modulare Framework zur manuellen Segmentierung in der Softwareplattform Mevislab.

(Quelle: Privataarchiv: Klinische Abteilung für Mund-, Kiefer- und Gesichtschirurgie, MUG)

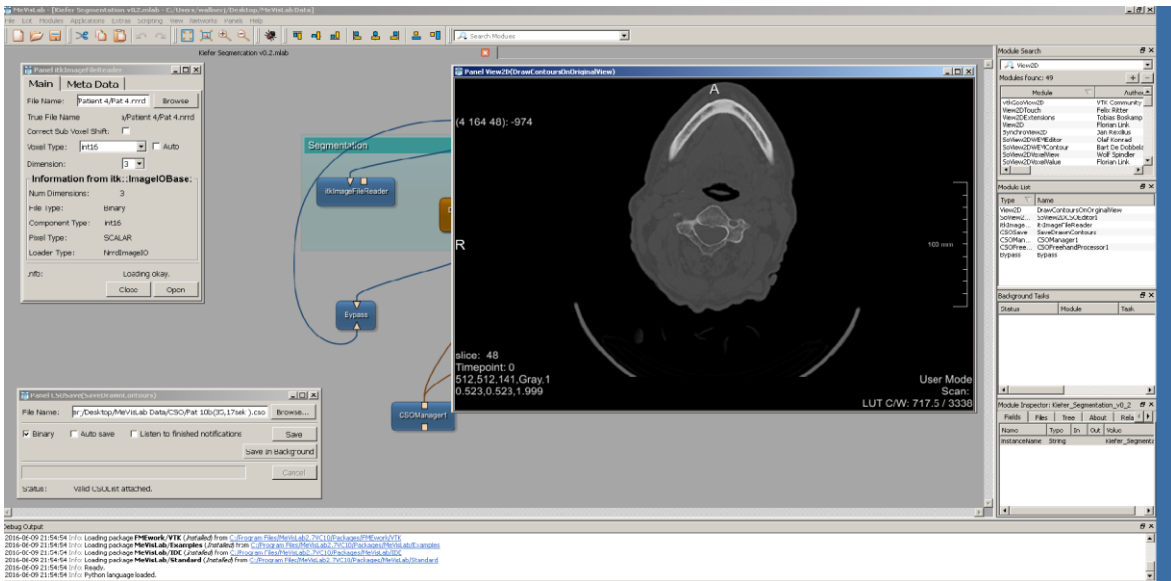


Abbildung 3 Segmentierungsvorgang Mevislab. Die Abbildung zeigt einen Schnitt eines CT-Bildes, nach dem Einspielen der CT-Daten in das Programm Mevislab. Anschließend wurde der Unterkiefer manuell segmentiert.

(Quelle: Privatarchiv: Klinische Abteilung für Mund-, Kiefer- und Gesichtschirurgie, MUG)

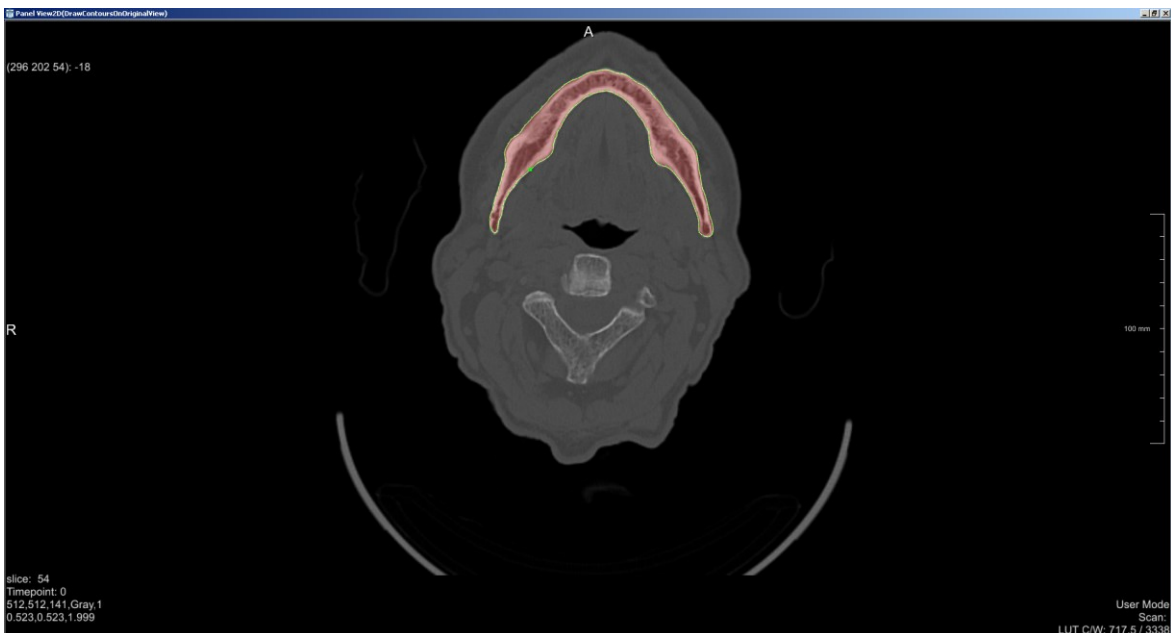


Abbildung 4 Manuelle schichtweise Segmentierung. Die Abbildung zeigt den Schnitt eines CT-Bildes. In dieser Schicht wurde der Unterkiefer bereits manuell segmentiert.

(Quelle: Privatarchiv Klinische Abteilung für Mund-, Kiefer- und Gesichtschirurgie, MUG)

9.2 Zielgrößen und Endpunkte

Zur vergleichenden Beurteilung der semiautomatischen Segmentierung hinsichtlich Genauigkeit, Übereinstimmung, Funktionalität und Praktikabilität wurden die Zielgrößen 1) benötigte Zeit der Segmentierung (t), 2) die Hausdorff-Distanz (HD), 3) der Dice-Score-Koeffizient (DSC), 4) das Volumen (mm³) und 5) die Anzahl der Voxel der virtuellen Segmentierungsmodelle gewählt. Der Vergleich wurde zwischen den semiautomatisch erstellten Segmentierungsmodellen mit Hilfe des GrowCut-Algorithmus und den zweifach manuell erstellten Ground Truth Modellen derselben anatomischen Struktur anhand der definierten Zielgrößen erstellt.

Ad 1: Die Zeit, die der Benutzer für die vollständige Segmentierung mit dem GrowCut-Algorithmus benötigte, wurde gemessen. Ebenfalls wurde die Zeit gemessen, die die klinischen Experten für die manuelle schichtweise Segmentierung benötigten. Hierfür wurde als Startpunkt das Einlesen des Datensatzes und als Endpunkt der Zeitmessung das Speichern des Segmentierungsvolumens als virtuelles 3D Modell festgesetzt.

Ad 2: Die Hausdorff-Distanz ist ein weit verbreiteter Parameter um zwei Punktmengen miteinander vergleichen zu können. Dabei wird der geringste Abstand eines Punktes der Punktmenge A zu einem Punkt aus der Punktmenge B ermittelt. Je geringer die Hausdorff-Distanz ist, umso geringer ist der Abstand zwischen den Punkten aus der Punktmenge A und den Punkten aus der Punktmenge B. Folglich decken sich bei einer geringen Hausdorff-Distanz die Punktmengen zu einem großen Teil und weisen eine hohe Übereinstimmung auf.
(33)

Ad 3: Der Dice-Score (DSC) ist ein wichtiger Parameter, um die räumliche Überlappung bzw. Überschneidung zweier Bildsegmente beurteilen zu können. Er ist definiert als $DSC(A, B) = 2(A \cap B) / (A + B)$, wobei \cap die Überschneidung definiert.
(34)

Ad 4: Die Volumina der semiautomatischen Segmentierung der Unterkieferdatensätze werden mit den Volumina der manuellen Segmentierungen verglichen.

Ad 5: Die Anzahl der Voxel der segmentierten Unterkieferdatensätze werden ebenfalls miteinander verglichen.

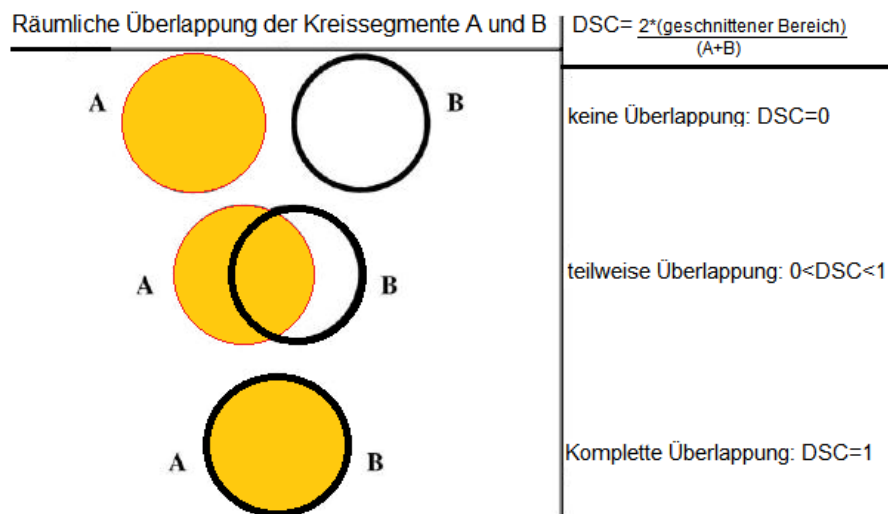


Abbildung 5 Dice-Score Darstellung. Die Abbildung stellt die räumliche Überlappung bzw. Überschneidung anhand von Kreissegmenten dar.

$$DSC = \frac{2 \cdot (\text{geschnittener Bereich})}{(\text{Summe von Region A und B})}$$

Die definierten Parameter 1 - 4 wurden sowohl zur Bewertung und zum Vergleich der manuellen Segmentierungsmodelle (A, B) untereinander, als auch zur vergleichenden Bewertung zwischen der manuellen Segmentierung (A, B) und der semiautomatischen Segmentierung (GrowCut) herangezogen.

9.3 Statistische Methoden

Zur statistischen Auswertung der Studie wurden deskriptive Berechnungen der gemessenen Werte durchgeführt und das Minimum, das Maximum, die Standardabweichungen und die Durchschnittswerte berechnet. Weitere analytische statistische Methoden wurden für die Berechnungen der Unterschiede zwischen den Gruppen verwendet. Diese beinhalteten (A-C):

- A. Korrelationskoeffizienten (Pearson)
- B. Zweistichproben t-Test
- C. Regressionsanalysen

Zunächst erfolgte die statistische Auswertung zum Vergleich der Ground Truth Daten A und B untereinander. Anschließend wurden die Segmentierungsergebnisse des GrowCut-Algorithmus mit den Ergebnissen der Ground Truth A und denen der Ground Truth B verglichen.

Die gesamte statistische Auswertung wurde in SPSS®⁷ durchgeführt.

⁷ IBM SPSS® Statistics 16, MS Windows, <https://spss-16.de.softonic.com/>

10 Ergebnisse

Ziel dieser Studie war es, eine frei- zugängliche konventionelle Methode zur bildbasierten semiautomatischen Segmentierung des Unterkiefers zu testen. Dabei wurde das Hauptaugenmerk auf die praktische Durchführbarkeit in der klinischen Anwendung, die Genauigkeit der Segmentierung und den möglichen Zeitvorteil gelegt. Die untersuchte Segmentierungsmethode wurde mit der Ground Truth bzw. den Ground Truth Daten derselben anatomischen Struktur verglichen. Hierfür wurden die Ergebnisse der zwei manuellen Segmentierungen untereinander und die Ergebnisse der manuellen Segmentierung mit den Ergebnissen des GrowCut-Algorithmus verglichen. Als Untersuchungsparameter dienten 1) die benötigte Zeit der Segmentierung, 2) der Dice-Score, 3) die Hausdorff-Distanz, 4) das Volumen und 5) die Anzahl der Voxel der 3D Segmentierungsmodelle. Der Vergleich einzelnen Segmentierungen erfolgte anhand dieser Parameter. In Tabelle 2 werden die Ergebnisse der manuellen schichtweisen Segmentierungen (Ground Truth A, B) dargestellt.

Segmentierte Unterkiefer	Volumen der Unterkiefer (mm ³)		HD (Voxel)	DSC (%)	Zeit in Minuten (t)		
	Ground Truth A	Ground Truth B			Ground Truth A	Ground Truth B	Algorithmus
	1	30507.8			29413.4	3.16	
2	17333	17730.4	5.2	91.72	46	40	1
3	19356.9	20067.2	3.16	92.65	38	39	1
4	46506.9	47508.8	6.32	94.66	38	38	1
5	39813.6	39733	3.32	93.68	37	35	1
6	30861.2	31283.1	4.12	94.48	43	40	1
7	45792.7	45492.8	4.69	94.11	38	42	1
8	31525.1	32288.9	2.24	94.23	36	37	1
9	18150.5	18686.3	4.24	92.53	38	38	1
10	32951.8	31296.5	3.46	93.73	36	35	1

Tabelle 2 Manuelle Segmentierungen im direkten Vergleich. Die Tabelle zeigt den direkten Vergleich der beiden manuellen Segmentierungen des Unterkiefers (Ground Truth A, B). Die manuelle Segmentierung wurde von zwei klinischen Experten (A, B) bei jedem der zehn Unterkiefer Schicht für Schicht durchgeführt. DSC(%), HD (Voxel)

Die untersuchte Segmentierungsmethode mithilfe des GrowCut-Algorithmus konnte alle Datensätze erfolgreich segmentieren. Der Benutzer definierte im Programm lediglich Punkte im Vordergrund der digitalen Schichtbilder, sowie Punkte im Hintergrund. Dies war für den Benutzer nach einer fünfminütigen Einweisung möglich. Der zeitliche Aufwand lag im Mittel bei 1 Minute. Bei der manuellen Segmentierung lag der Zeitaufwand im Schnitt bei 38.6 ± 3.31 Minuten beim Kliniker A (Ground Truth A) und 38.4 ± 2.27 beim Kliniker B (Ground Truth B). Das mittlere Volumen bei der manuellen Segmentierung betrug 31.28 ± 10.69 cm³ (Ground Truth A) und 31.35 ± 10.59 cm³ (Ground Truth B). Das Segmentierungsvolumen des semiautomatischen GrowCut-Algorithmus lag im Durchschnitt bei 32.18 ± 13.02 cm³. (Tabellen 1-5).

	Volumen der Unterkiefer (cm ³)		Ground truth A vs. Ground truth B		Zeit in Minuten (t)		
	Ground Truth A	Ground Truth B	HD (Voxel)	DSC (%)	Ground Truth A	Ground Truth B	Algorithmus
Min	17.33	17.73	2.24	91.72	36	35	1
Max	46.51	47.51	6.32	94.66	46	42	1
μ	31.28	31.35	3.99	93.61	38.6	38.4	1.0
$\pm \sigma$	± 10.69	± 10.59	± 1.18	± 0.98	± 3.31	± 2.27	± 0.0

Tabelle 3 Manuelle Segmentierungsergebnisse mit Min, Max, $\sigma \pm \mu$. Die Tabelle zeigt die Segmentierungsergebnisse der beiden manuellen Segmentierungen (Ground Truth A, B) mit Minimum, Maximum, Mittelwert (μ) und der Standardabweichung (σ) der 10 Unterkiefer. DSC (%), HD (Voxel)

Segmentierte Unterkiefer	Volumen der Unterkiefer (mm^3)		HD (Voxel)	DSC (%)	Zeit in Minuten (t) - Ground Truth A
	Ground Truth A	Algorithmus			
1	30507.8	26710.9	29.22	83.26	36
2	17333	21200.4	51.39	80.73	46
3	19356.9	19033.5	21.35	82.73	38
4	46506.9	47028.9	19.65	88.42	38
5	39813.6	50087.4	57.46	80.81	37
6	30861.2	30118.9	29.1	87.8	43
7	45792.7	52090.8	29.45	86	38
8	31525.1	30556.2	49.49	88.27	36
9	18150.5	16548.2	19.87	90.33	38
10	32951.8	28474.4	28.14	86.28	36

Tabelle 4 Ergebnisse der Ground Truth A und des GrowCut-Algorithmus. Die Tabelle zeigt den direkten Vergleich zwischen der manuellen Schicht-für-Schicht-Segmentierung der Ground Truth A und der semiautomatischen Segmentierung mithilfe des GrowCut-Algorithmus. Als Parameter dienen die Hausdorff-Distanz, der Dice-Score-Koeffizient, die Zeit(t) und das Volumen(mm^3). Die Zeit der semiautomatischen Segmentierung betrug im Mittel 1 Minute, während die Zeiten der manuellen Segmentierung in der äußersten Spalte dargestellt werden. DSC (%), HD (Voxel)

Segmentierte Unterkiefer	Volumen der Unterkiefer (mm ³)		HD (Voxel)	DSC (%)	Zeit in Minuten (t) - Ground Truth B
	Ground Truth B	Algorithmus			
1	29413.4	26710.9	27.91	83.6	40
2	17730.4	21200.4	50.96	80.66	40
3	20067.2	19033.5	20.71	83.77	39
4	47508.8	47028.9	19.34	88.69	38
5	39733	50087.4	57.46	80.59	35
6	31283.1	30118.9	28.86	88.79	40
7	45492.8	52090.8	33.65	86.34	42
8	32288.9	30556.2	47.84	87.76	37
9	18686.3	16548.2	19.34	89.85	38
10	31296.5	28474.4	28.25	87.49	35

Tabelle 5 Vergleich Ground Truth B mit GrowCut-Algorithmus. Die Tabelle zeigt den direkten Vergleich zwischen der manuellen Schicht-für-Schicht-Segmentierung der Ground Truth B und der semiautomatischen Segmentierung mithilfe des GrowCut-Algorithmus. Als Parameter dienen die Hausdorff-Distanz und der Dice-Score-Koeffizient sowie die Zeit und das Volumen (mm³). Die Zeit der semiautomatischen Segmentierung betrug rund 1 Minute, die Zeiten der manuellen Segmentierung werden in der äußersten Spalte dargestellt. DSC (%), HD (Voxel)

Die Anzahl der Voxel lag im Durchschnitt bei 119147 ± 46957.5 für die Ground Truth A und bei 119200.7 ± 45568.9 für die Ground Truth B. Die durchschnittliche Anzahl der Voxel lag für die Segmentierungsmethode des GrowCut-Algorithmus lag bei 123613.5 ± 58013.4 Voxel (siehe Tabelle 6).

<i>Segmentierte Unterkiefer</i>	<i>Anzahl der Voxel</i>		
	<i>Ground Truth A</i>	<i>Ground Truth B</i>	<i>Algorithmus</i>
<i>1</i>	166749	160767	145996
<i>2</i>	118277	120989	144668
<i>3</i>	54887	56901	53970
<i>4</i>	84897	86726	85850
<i>5</i>	153211	152901	192747
<i>6</i>	96836	98160	94507
<i>7</i>	211925	210537	241072
<i>8</i>	77436	79312	75056
<i>9</i>	123856	127512	112922
<i>10</i>	103396	98202	89347
<i>Minimum</i>	54887	56901	53970
<i>Maximum</i>	211925	210537	241072
μ	119147	119200.7	123613.5
$\pm \sigma$	± 46957.5	± 45568.9	± 58013.4

Tabelle 6 Anzahl der Voxel Ground Truth (A,B), GrowCut. Die Tabelle zeigt die Anzahl der Voxel, Minimum-, Maximum-, Mittelwert (μ) und die Standardabweichung (σ) von den jeweiligen manuellen Segmentierungen (Ground Truth A und Ground Truth B) sowie der semiautomatischen Segmentierung (GrowCut-Algorithmus).

Anschließend wurde die Übereinstimmung zwischen den manuellen Segmentierungen untereinander (A, B) und jeweils eine manuelle Segmentierung mit der semiautomatischen Segmentierung verglichen (A, GrowCut; B, GrowCut). Die Übereinstimmung ergab einen Dice-Score-Koeffizienten im Mittel von $93.61 \pm 0.98\%$ zwischen den manuellen Segmentierungen der klinischen Experten (A, B). Im Vergleich zur semiautomatischen Segmentierung (GrowCut) wurde ein Dice-Score-Koeffizient im Durchschnitt von $85.46 \pm 3.38\%$ für die Ground Truth A und ein Dice-Score von $85.75 \pm 3.39\%$ für die Ground Truth B erreicht. Die Hausdorff-Distanz betrug im Schnitt 3.99 ± 1.18 Voxel-Einheiten zwischen der Ground Truth A und der Ground Truth B sowie 33.51 ± 13.98 Voxel-Einheiten zwischen der Ground Truth A und der semiautomatischen Segmentierung (GrowCut) und 33.43

± 13.86 Voxel-Einheiten zwischen der Ground Truth B und der semiautomatischen Segmentierung (GrowCut) (siehe Tabelle 3 und 7).

	<i>Volumen der Unterkiefer (cm³)</i>			<i>Ground Truth A, B vs. Algorithmus</i>			
	<i>Ground Truth A</i>	<i>Ground Truth B</i>	<i>Algorithmus</i>	<i>DSC (%)</i>		<i>HD (voxel)</i>	
				<i>Ground Truth A</i>	<i>Ground Truth B</i>	<i>Ground Truth A</i>	<i>Ground Truth B</i>
<i>Min</i>	17.33	17.73	16.55	80.73	80.59	19.65	19.34
<i>Max</i>	46.51	47.51	52.09	90.33	89.85	57.46	57.46
μ	31.28	31.35	32.18	85.46	85.75	33.51	33.43
$\pm\sigma$	± 10.69	± 10.59	± 13.02	± 3.38	± 3.39	± 13.98	± 13.86

Tabelle 7 Vergleich der Ergebnisse Ground Truth (A,B) vs. Algorithmus. Die Tabelle zeigt zusammenfassend die Ergebnisse der errechneten Volumina der Ground Truth (A, B) sowie mit der GrowCut-Segmentierungsmethode mithilfe von Minimum, Maximum, Standardabweichung (σ) und Mittelwerten (μ). Weiters werden die Hausdorff-Distanz und der Dice-Score der Ground Truth (A, B) im Vergleich zur GrowCut-Segmentierungsmethode mithilfe von Minimum, Maximum, der Standardabweichung (σ) und dem Mittelwert (μ) angegeben. DSC (%), HD (Voxel)

Die berechneten Unterschiedswerte zwischen den beiden manuellen Segmentierungs-Gruppen, sowie zwischen der jeweiligen manuellen und der semiautomatischen Segmentierung waren nicht signifikant unterschiedlich zwischen den definierten Parametern, Volumen und Voxel ($p < 0.05$). Der Korrelationskoeffizient (Pearson) war im Vergleich zwischen den jeweiligen Gruppen nahe 1 mit einem Mittelwert von 0.94 ($r > 0.94$) (siehe Tabelle 8,9).

<i>Volumen Vergleich</i>	<i>Signifikanz (p)</i>	<i>Koeffizient (r)</i>
<i>Ground Truth A vs. Ground Truth B</i>	$p = 0.803$	$r = 0.997$
<i>Ground Truth A vs. Algorithmus</i>	$p = 0.550$	$r = 0.943$
<i>Ground Truth B vs. Algorithmus</i>	$p = 0.571$	$r = 0.948$

Tabelle 8 Volumen Vergleichsberechnungen. Die Tabelle zeigt die unterschiedlichen Werte des gepaarten t-Tests sowie die Produkt-Moment Korrelationswerte (Pearson, r) für den Vergleich der Volumina der manuellen Segmentierungen (A, B) und der Volumina der semiautomatischen Segmentierung (GrowCut-Algorithmus). Es ist statistisch kein signifikanter Unterschied ($p < 0.05$) zwischen den segmentierten Volumina nachweisbar. Eine hohe direkt proportionale Korrelation (r) der segmentierten Volumina nahe dem Wert 1 ist feststellbar.

<i>Voxel Vergleich</i>	<i>Signifikanz (p)</i>	<i>Koeffizient (r)</i>
<i>Ground Truth A vs. Ground truth B</i>	$p = 0.960$	$r = 0.998$
<i>Ground Truth A vs. Algorithmus</i>	$p = 0.502$	$r = 0.948$
<i>Ground Truth B vs. Algorithmus</i>	$p = 0.493$	$r = 0.957$

Tabelle 9 Voxel Vergleichsberechnungen. Die Tabelle zeigt die unterschiedlichen Werte des gepaarten t-Tests sowie die Produkt-Moment Korrelationswerte (Pearson, r) für den Vergleich der Voxel der manuellen Segmentierung und den Voxel der semiautomatischen Segmentierung (GrowCut-Algorithmus). Es ist statistisch kein signifikanter Unterschied ($p < 0.05$) zwischen der Anzahl der Voxel feststellbar. Eine hohe direkt proportionale Korrelation (r) der segmentierten Volumina nahe dem Wert 1 ist erkennbar.

Die Regressionsanalyse zeigt, dass die Messwerte der Volumen und der Voxel nahe der konstruierten Regressionslinie liegen (siehe Tabelle 10, 11). Der Unterschied zwischen dem Gradienten der Regressionsgeraden in den konstruierten Regressionsmodellen war für die Vergleiche zwischen manueller-

(Ground Truth A, B) und semiautomatischer (GrowCut) Segmentierung nicht signifikant unterschiedlich ($p < 0.001$). Die Tabellen 10 und 11 stellen in den Boxplots die Ähnlichkeit der Messwerte in den einzelnen Segmentierungsgruppen für das Volumen und die Anzahl der Voxel grafisch dar. Besonders ähnlich sind die Werte der Ground Truth Modelle A und B.

Ein ähnliches Ergebnis zeigte der Vergleich der Parameter Hausdorff-Distanz und Dice-Score-Koeffizient. Die Parameter unterschieden sich beim Vergleich zwischen der semiautomatischen Segmentierung (GrowCut-Algorithmus) und der manuellen Segmentierung (A, B) nur geringfügig.

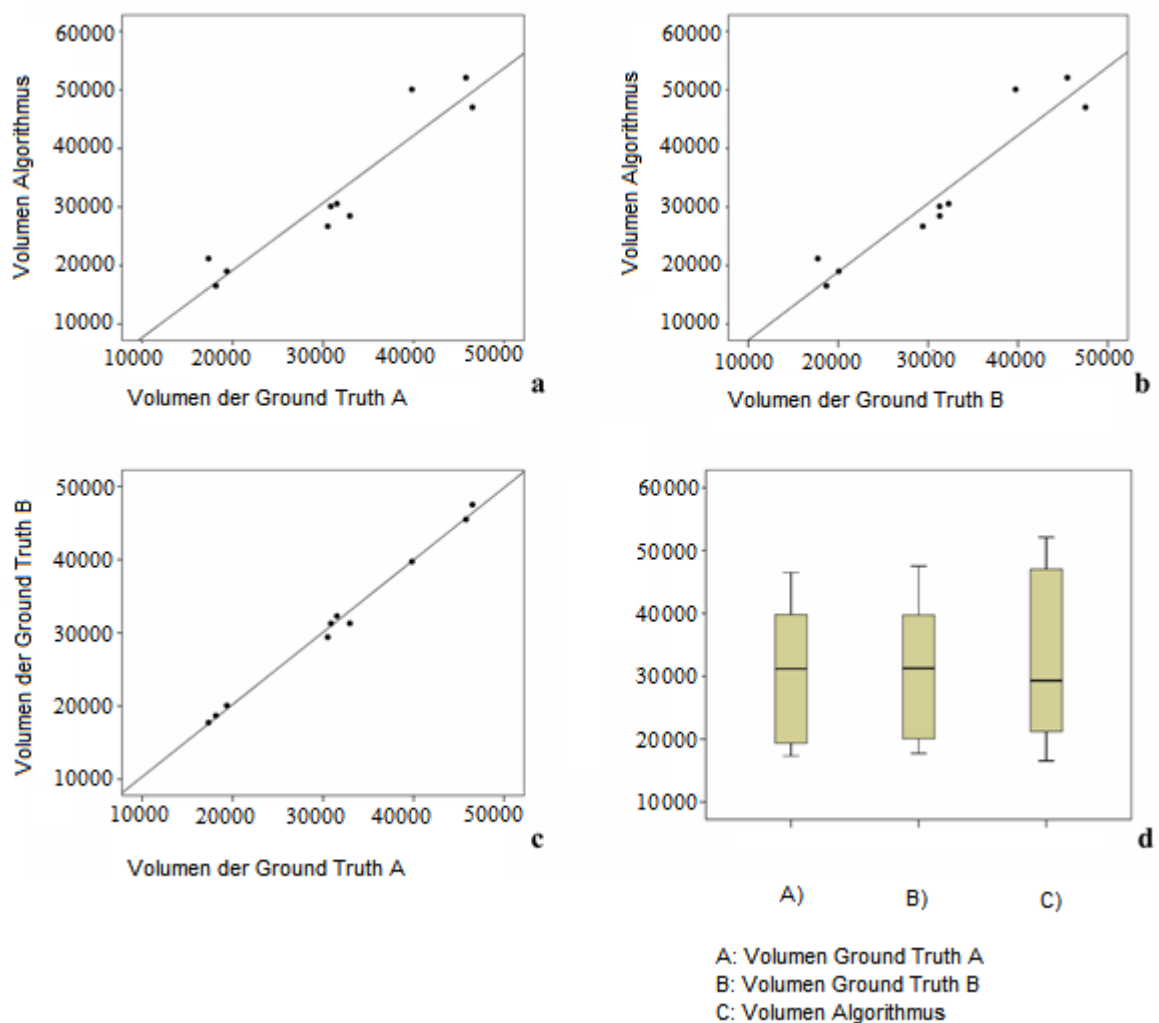


Tabelle 10 Regressionsanalyse Volumen. Die Tabelle zeigt die Regressionsanalyse der Volums-Messungen. Die Werte der Volums-Messungen liegen bei der Ground Truth Segmentierung (A und B) und der semiautomatischen Segmentierung (GrowCut) nahe der Regressionslinie (a, b). Die Werte der Volumenmessungen der Ground Truth A und der Ground Truth B liegen nahe der Regressionslinie (c). Die Ergebnisse der

Volumenmessungen der Segmentierung sind im Boxplot Diagramm dargestellt (d). Die Differenzwerte zwischen dem Gradienten der Regressionslinie waren statisch nicht signifikant unterschiedlich. ($p < 0.001$).

(Die Werte der Volumina sind in mm^3 angegeben)

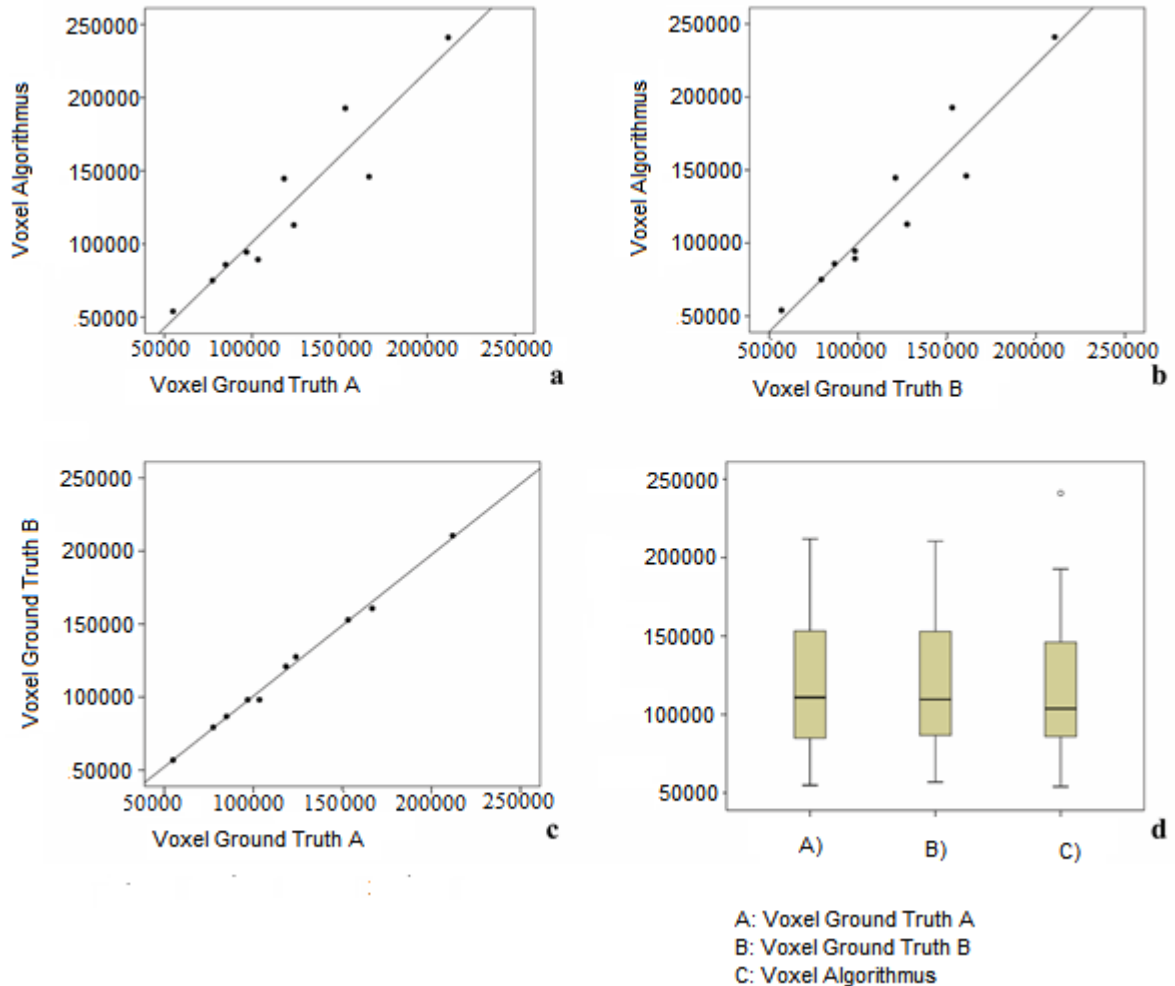
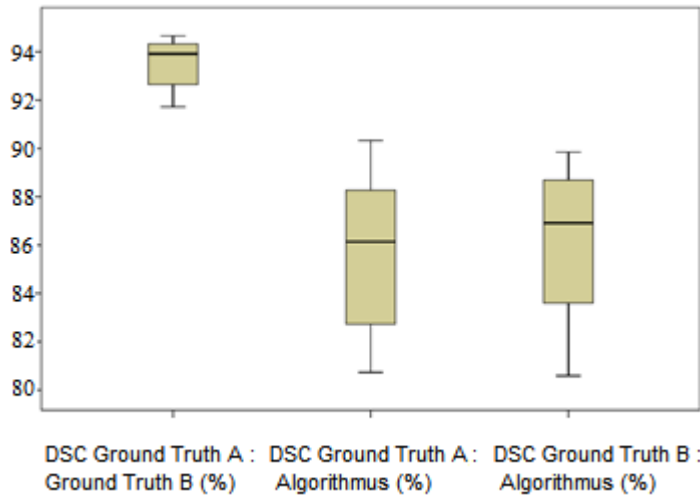
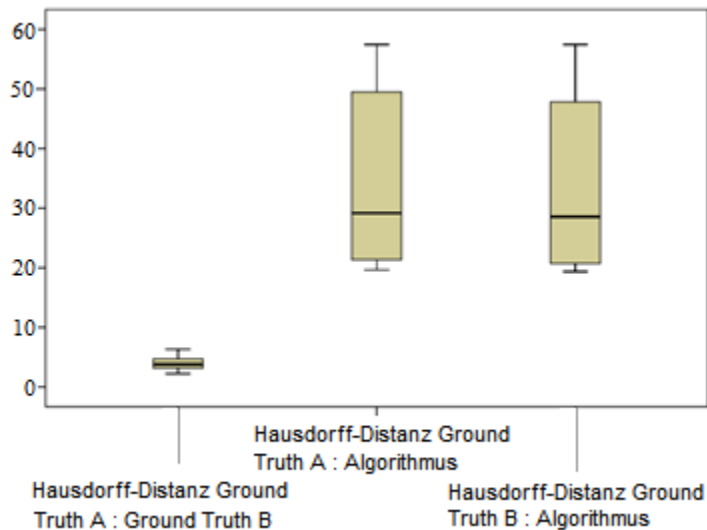


Tabelle 11 Regressionsanalyse Voxel. Die Tabelle zeigt die Regressionsanalyse der Messwerte der Voxel. Die Werte der Voxelmessungen liegen bei der Ground Truth Segmentierung (A, B) und der semiautomatischen Segmentierung (GrowCut) nahe der Regressionslinie (a, b). Die Werte der Voxelmessungen der Ground Truth A und der Ground Truth B liegen nahe der Regressionslinie (c). Die Ergebnisse der Voxelmessungen der Segmentierungen sind im Boxplot Diagramm dargestellt (d). Die Differenzwerte zwischen dem Gradienten der Regressionslinie waren statisch nicht signifikant unterschiedlich ($p < 0.001$).



<i>DSC (%) Vergleich</i>	<i>Min</i>	<i>Max</i>	μ	σ
<i>Ground Truth A vs. Ground truth B</i>	91.7	94.7	93.6	1.0
<i>Ground Truth A vs. Algorithmus</i>	80.7	90.3	85.5	3.4
<i>Ground Truth B vs. Algorithmus</i>	80.6	89.9	85.8	3.4

Tabelle 12 Dice Score Werte. Die Tabelle zeigt den Vergleich der Werte des Dice-Score-Koeffizienten (DSC in %) mit Minimum- (Min), Maximum- (Max), Mittelwert (μ) und Standardabweichung (σ). Die Ergebnisse sind im Boxplot Diagramm dargestellt. Die Plots für den Vergleich der DSC-Werte der Ground Truth Segmentierungen (A, B) mit den DSC-Werten der semiautomatischen Segmentierung (GrowCut) sind ähnlich. Die Werte der Ground Truth Segmentierungen (A, B) wurden von zwei klinischen Experten (A, B) unabhängig voneinander erstellt.



<i>HD (Voxel) Vergleich</i>	<i>Min</i>	<i>Max</i>	μ	σ
<i>Ground Truth A vs. Ground truth B</i>	2.2	6.3	4.0	1.2
<i>Ground Truth A vs. Algorithmus</i>	19.7	57.5	33.5	14.0
<i>Ground Truth B vs. Algorithmus</i>	19.3	57.5	33.4	13.9

Tabelle 13 Hausdorff-Distanzen (HD, Voxel) Werte. Die Tabelle zeigt den Vergleich der berechneten Werte für die Hausdorff-Distanz (HD in Voxel) mit Minimum- (Min), Maximum- (Max), Mittelwert (μ) und Standardabweichung (σ). Die Ergebnisse sind im Boxplot Diagramm dargestellt. Die Plots für den Vergleich der HD-Werte der Ground Truth Segmentierungen (A, B) mit den HD-Werten der semiautomatischen Segmentierung (GrowCut) sind ähnlich. Die Werte der manuellen Segmentierungen (Ground Truth A, B) wurden von zwei klinischen Experten (A, B) unabhängig voneinander erstellt.

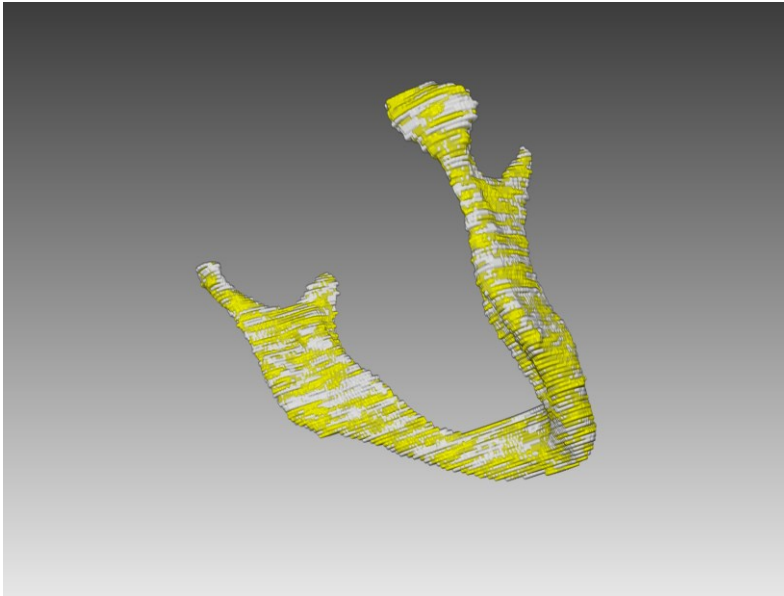


Abbildung 6: Visualisierung der virtuellen Segmentierungsmodelle. Die Segmentierungsergebnisse der manuellen Segmentierung (Ground Truth) in Weiß und das Segmentierungsergebnis des semiautomatischen Algorithmus in Gold werden übereinandergelegt dargestellt. Als Bewertungsgrundlage werden der Dice- Score, die Hausdorff-Distanz, das Segmentierungsvolumen und die Anzahl der Voxel herangezogen. (Quelle: Privatarchiv Klinische Abteilung für Mund-, Kiefer- und Gesichtschirurgie, MUG)

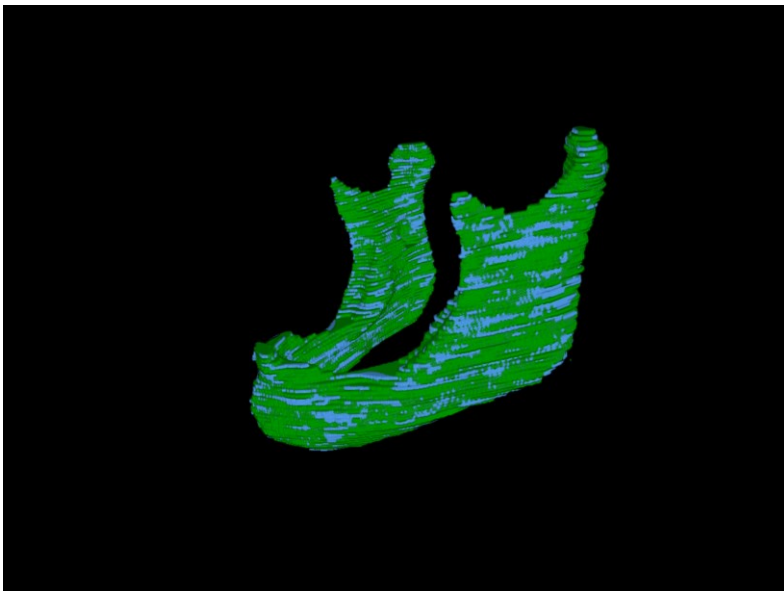


Abbildung 7: Visualisierung der virtuellen Ground Truth Modelle (A, B). Die Abweichungen der Ground Truth Segmentierungen (blau A, grün B) wurden für jeden Datensatz überprüft. Die Ground Truth Daten wurden durch manuelle schichtweise Segmentierung von zwei klinischen Experten (A, B) für jeden Datensatz zweifach erzeugt (A, B). Die Ergebnisse des jeweiligen Kliniklers (A, B) werden übereinandergelegt

dargestellt (blau A, grün B). Als Bewertungsgrundlage werden der Dice- Score, die Hausdorff-Distanz, das Segmentierungsvolumen und die Anzahl der Voxel herangezogen.
(Quelle: Privataarchiv Klinische Abteilung für Mund-, Kiefer- und Gesichtschirurgie, MUG)

11 Diskussion

Das Aufkommen von computerunterstützten bildbasierten 3D Softwaretechnologien für die Anwendung im medizinischen Bereich hat sich in der Unterstützung der Diagnosestellung und Planung von ärztlichen Behandlungen als nützlich erwiesen. Deshalb ist der Einsatz dieser Technologien mittlerweile Standard in zahlreichen medizinischen Fachgebieten. (35) Computergestützte Softwaretechnologien können die medizinische Diagnostik unterstützen und folglich die Behandlungszeit reduzieren oder ebenfalls unterstützend in der Therapieplanung in komplexen Fällen mitwirken. Komplexe anatomische Strukturen können 3D visualisiert und das therapeutische Vorgehen an virtuellen und realen Modellen simuliert werden. (3) Die Grundlage für die Erstellung dieser virtuellen und realen 3D-Modelle, die ebenfalls zu Lehrzwecken eingesetzt werden können, bilden in klinischen Zentren funktionsstabile bildbasierte Segmentierungsmethoden. (36)

Zur Erstellung von 3D-Modellen einer anatomischen Zielstruktur ist die Segmentierung der anatomischen Region aus einem entsprechenden Bilddatensatz erforderlich. Aus diesem Grund existieren zahlreiche semi- und vollautomatische bildbasierte Segmentierungsmethoden, die je nach ihrem zu Grunde liegenden Algorithmus selbstständig entsprechende anatomische Zielstrukturen segmentieren können. Da diese bildbasierten 3D Softwaretechnologien zunehmende Verwendung in der Medizin finden (35), sind in den letzten Jahren zahlreiche komplexe kommerziell-erwerbbar lizenzgestützte Softwarepakete mit vielseitigen Anwendungsmöglichkeiten in der radiologischen Bildbearbeitung entstanden. Jedoch sind diese Softwarepakete meistens Einrichtungen vorbehalten, die die entsprechenden finanziellen und humanen Ressourcen für eine ausreichende Bedienung und Anwendung aufbringen können. (36)

Eine frei zugängliche semiautomatische konventionelle Segmentierungsmethode für eine einfache, rasche und präzise Bildsegmentierung hat demnach einen hohen Stellenwert für klinische Einrichtungen, da sie unabhängig von deren Größe, deren Ressourcen oder finanziellen Aspekten eingesetzt werden kann. In dieser Studie wurde deshalb eine open-source Segmentierungsmethode zur semiautomatischen Segmentierung des Unterkiefers getestet.

Der Unterkiefer wurde als anatomische Zielstruktur für die Segmentierung ausgewählt, da er eine klar definierte Knochenstruktur aufweist, die einen objektiven Vergleich zwischen den zu untersuchenden virtuellen Segmentierungsmodellen mit einer entsprechenden Kontrollgruppe derselben anatomischen Struktur ermöglicht (Ground Truth). Zudem spielt der Unterkiefer im Bereich der Mund-, Kiefer- und Gesichtschirurgie eine klinisch relevante Rolle, da er bei Gesichtsschädelfrakturen mit ungefähr 40% der am häufigsten betroffene Knochen ist. (37,38) Klinisch sind diese Frakturen oft mit zeitintensiven, komplexen, chirurgischen Eingriffen verbunden. Bildbasierte Softwareprogramme und digitale Bildbearbeitungen werden bei diesen komplexen Fällen im Unterkiefer daher, wie in anderen anatomischen Bereichen ebenfalls, unterstützend eingesetzt.

Klinisch-praktisch beweisen dies verschiedene Arbeiten in der Literatur. (39) (40) (41) Thor et al. zeigten eindrucksvoll die Vorteile einer virtuellen OP-Planung mithilfe von bildbasierten Segmentierungsprozessen. Sie beschrieben die virtuelle Planung eines chirurgischen Eingriffs bei einer dislozierten, bilateralen Unterkieferfraktur. Eine herkömmliche Osteosynthese scheiterte und der Patient entwickelte einen offenen Biss. Mithilfe eines CBCT's wurde die aktuelle Situation festgehalten und der Unterkiefer anschließend mit einer Computersoftware segmentiert. Durch virtuelle Osteotomien im Bereich der Frakturen konnte am Computer die korrekte Okklusion simuliert und wiederhergestellt werden. Der Biss wurde virtuell geschlossen. Man ließ nun für die am Computer geplante Situation eine patientenspezifische Osteosyntheseplatte anfertigen. Im Rahmen des chirurgischen Eingriffs wurde die Rekonstruktionsplatte unter Zuhilfenahme der präoperativen Planung sowie einer präoperativ angefertigten Bohrschablone verschraubt. Eine gute und komplikationsfreie Heilung wurde erzielt. Die virtuelle Planung und die auf diese Weise angefertigten patientenspezifischen Implantate erwiesen sich als enorm wertvoll für den chirurgischen Eingriff. (39)

Nakao et al. beschrieben Methoden zur automatisierten, virtuellen, präoperativen Planung von Unterkieferrekonstruktionen mittels Fibulatransplantaten. In ihrer retrospektiven Studie wurden 120 Fälle virtuell geplant und mit den Planungsentscheidungen des/r Chirurgen/in verglichen. Die Ergebnisse zeigten eine hohe Übereinstimmung der beiden Planungen. (40) Für eine prospektive effektive klinische Anwendung derartiger virtueller Operationsplanungen sind

präzise zeitsparende automatische Segmentierungsvorgänge, die in Funktionsstabilität und Genauigkeit überzeugen, Voraussetzung.

In einer anderen Studie von Abdolali et al. wurde der Mandibularkanal automatisch segmentiert. Dies kann für präoperative implantat-chirurgische Planungen im Unterkiefer ein wesentlicher Vorteil sein. Die Segmentierungsmethode basierte auf einer Kombination aus statistischen und anatomischen Informationen. Die Ergebnisse waren auch bei unscharfen CT-Bildern sowie bei Unterkiefern mit leichter Knochenresorption zufriedenstellend. Laut dieser Studie war der klinische Einsatz der verwendeten Segmentierungssoftware in der computergesteuerten Planung von Zahnimplantaten zuverlässig möglich. (41)

In der vorliegenden Studie wurde die untersuchte semiautomatische open-source Segmentierungsmethode in Bezug auf ihre Genauigkeit, Funktionalität und Möglichkeit der klinischen Anwendung geprüft. Im Rahmen der klinischen Routine wurden vorhandene CT-Datensätze des Gesichtsschädels ausgewählt und anschließend der gesamte Unterkiefer segmentiert. Die Segmentierung erfolgte zum einen manuell schichtweise, um eine Kontrollgruppe (Ground Truth) zu erstellen, und zum anderen semiautomatisch mithilfe des GrowCut-Algorithmus, der auf zellulären Automaten basiert.

Die manuellen und die semiautomatischen Segmentierungen wurden mithilfe von definierten Parametern vergleichend beurteilt (Zeit, Volumen, Anzahl der Voxel, Dice-Score, Hausdorff-Distanz). Zunächst wurden die manuellen Ground Truth Segmentierungen (A, B) untereinander verglichen, um mögliche Ungenauigkeiten oder systemische Fehler in der Kontrollgruppe zu erkennen. Anschließend wurden die beiden manuellen Segmentierungen den semiautomatischen GrowCut-Segmentierungen gegenübergestellt. Die zur Hypothesenprüfung herangezogenen Parameter sind valide Parameter für den Vergleich der Übereinstimmung zweier Segmentierungsvolumen bzw. virtueller Segmentierungsmodelle und wurden bereits in anderen Studien zur Prüfung der Genauigkeit und Qualität von bildbasierten Segmentierungsprozessen verwendet. (34, 42, 43)

In der Literatur wird die Qualität und Genauigkeit einer Segmentierungsmethode gerne bewertet, indem virtuell erstellte 3D-Modelle mit realen, gedruckten 3D Modellen verglichen werden. (44, 45) Es gibt bislang jedoch keine Arbeit, die die Genauigkeit und klinische Anwendbarkeit einer open-source Segmentierungsmethode im cranio-mandibulo-maxillären Bereich anhand der

virtuellen Ground Truth Modelle derselben anatomischen Struktur und somit anhand der tatsächlichen visualisierten Anatomie vergleicht. Das Erstellen dieser Ground Truth Volumina bzw. dieser virtuellen Ground Truth Modelle erfolgte durch unabhängiges zweimaliges manuelles Segmentieren durch zwei klinische Experten. Mit Hilfe dieses Vorgehens konnten Differenzen und störende Subjektivitätsfaktoren im manuellen Segmentierungsprozess aufgezeigt und eine objektive Kontrollgruppe zur Hypothesenprüfung erstellt werden.

Eine mögliche Ursache für das Fehlen derartiger experimenteller Vergleiche mit vollständigen Ground Truth Modellen für bereits bestehende Softwarepakete ist die aufwendige und zeitraubende Erstellung dieser Kontrollgruppen (Ground Truth) durch die manuellen Segmentierungsvorgänge.

Anders als in dieser Arbeit verwendete Szymor et al. gedruckte 3D-Modelle für die Beurteilung einer Segmentierungsmethode im Bereich der inneren Augenhöhle. Die gedruckten Modelle wurden von derselben Region, die virtuell segmentiert wurde, erstellt. Die fertigen 3D-Modelle wurden anschließend mit einem optischen 3D-Scanner digitalisiert und erneut als virtueller Datensatz für eine vergleichende Beurteilung gespeichert. Anschließend wurde das virtuelle 3D-Modell, das im Rahmen der Segmentierung für den 3D-Druck generiert wurde, mit dem eingescannten erneut abgespeicherten 3D-Modell verglichen. (44) Der Nachteil bei diesem Vorgehen ist, dass durch die reale Herstellung des 3D-Modells zahlreiche Einflussfaktoren mitwirken, die die Beurteilung der Segmentierungsmethode beeinflussen können. So können unter anderem diese zusätzlichen Faktoren wie die Qualität des 3D-Druckers, die für das 3D-Modell verwendeten Materialien, die druckereigene Software und nicht zuletzt auch der 3D-Scanner, der das reale Modell wieder in eine virtuelle Form bringt, zu Fehlern oder zu ungenauen Beurteilungen und Endergebnissen führen. In der vorliegenden Arbeit konnte auf diese zusätzlichen Arbeitsschritte und einflussnehmenden Faktoren, wie die Verwendung eines 3D-Scanners o.ä., verzichtet werden. Durch die Erstellung der Ground Truth Modelle als Kontrollgruppe lagen die Segmentierungsmodelle ständig in virtueller Form vor und zusätzliche störende Einflussfaktoren wurden somit limitiert.

Naumovich et al. präsentierten eine Software, die es ermöglicht, aus CT-Bildern Zähne zu segmentieren und anschließend daraus ein 3D-Modell zu erstellen. Laut dieser Studie ist der Einsatz im klinischen Betrieb vorstellbar, da die Software einfach und intuitiv gestaltet ist. Die verwendete Software ist zudem mit Bildern

verschiedener CT-Scans kompatibel. Der Zeitaufwand betrug im Schnitt jedoch etwa 10-15 Minuten pro segmentiertem Datensatz und ist im Durchschnitt wesentlich höher als der Zeitbedarf der in dieser Arbeit untersuchten Segmentierungsmethode mithilfe des GrowCut-Algorithmus. (46)

Fu Y et al. wiesen in ihrer Studie auf die Vorteile einer präzisen Segmentierung der menschlichen Knochen in vielen medizinischen Disziplinen hin. Eine funktionsstabile Segmentierung wäre im Bereich der Diagnostik von Knochenerkrankungen, der Bewertung eines Therapieerfolges und bei der Planung von chirurgischen oder strahlentherapeutischen Eingriffen ein wesentlicher Vorteil. Trotz guter Kontrastunterschiede zwischen Knochen und umgebendem Weichgewebe war eine vollautomatische Segmentierung aufgrund von Ungenauigkeiten der Segmentierungsalgorithmen sowie der Funktionsinstabilitäten schwierig. CT-Bilder stellten für eine präzise vollautomatische Segmentierung eine gewisse Limitation dar. Dieser Limitationsfaktor ergibt sich aus einem niedrigen Signal-Rausch-Verhältnis, einer unzureichenden Auflösung oder einer ununterscheidbaren Bildintensität zwischen Knochen und Weichgewebe. (47) Aus diesem Grund wurde in der vorliegenden Studie eine semiautomatische Segmentierungsmethode anstatt eines vollautomatischen Verfahrens verwendet. Weiters spricht der von Fu Y et al. beschriebene Limitationsfaktor der vollautomatischen Segmentierung von CT-Datensätzen (47) für eine fehlende Funktionsstabilität dieser vollautomatischen Verfahren.

Ein funktionsstabiler Algorithmus zur (semi)automatischen Segmentierung ist für die virtuelle Therapieplanung unerlässlich, da eine manuelle Segmentierung im klinischen Alltag zu viel Zeit in Anspruch nehmen würde. Funktionsinstabilitäten im Segmentierungsverfahren können dabei allerdings zu zahlreichen beschriebenen Limitation führen, wie unter anderem von Fu Y et al. bereits berichtet wurde (47).

Voraussetzung für die klinische Anwendbarkeit von bildbasierten Softwareprogrammen, unabhängig von ihren zusätzlichen Funktionen, ist die Möglichkeit einer zeitsparenden, unkomplizierten und funktionsstabilen Segmentierung der anatomischen Zielstruktur mithilfe eines ubiquitär verfügbaren, konventionellen Algorithmus. Denn so kann der Einsatz in der klinischen Routine leicht und barrierefrei Anwendung finden.

Bei der in dieser Studie vorgestellten Segmentierungsmethode mithilfe des GrowCut-Algorithmus markiert der/die Benutzer/in je einen Punkt in einem

sagittalen, axialen und coronalen Schnitt am Unterkiefer und in dessen Umgebung. Dies ist im Schnitt in weniger als einer Minute möglich. Die weitere Segmentierung erfolgt automatisch mithilfe des bildbasierten open-source Verfahrens (Growcut). Eine Segmentierung ist daher in kurzer Zeit möglich. Diese Methode wurde in der Literatur bereits in anderen anatomischen Bereichen auf dem Gebiet der Neurochirurgie erfolgsversprechend angewandt. (48)

Die in der vorliegenden Studie erhaltenen Ergebnisse zeigen, dass nicht nur die Praktikabilität, sondern auch die Genauigkeit der semiautomatischen GrowCut-Segmentierung überzeugend ist. Beim Dice-Score-Koeffizienten konnte im Schnitt eine hohe Übereinstimmung zur Ground Truth Segmentierung von über 85% erreicht werden, während die Hausdorff-Distanz im Mittel bei unter 33.5 Voxel im Vergleich zur manuell segmentierten Kontrollgruppe (Ground Truth A, B) lag.

Die in dieser Studie erreichten Dice-Score-Koeffizienten und Hausdorff-Distanzen scheinen für eine klinische Anwendung der untersuchten Segmentierungsmethode akzeptabel zu sein, da - entsprechend der klinischen Erfahrung - Dice-Score-Koeffizienten von über 80% ausreichende Präzision für den Einsatz in klinischen Einrichtungen liefern. Das klinische Einsatzgebiet für die Anwendung der open-source Segmentierungsmethode mit GrowCut reicht dabei von 1) virtuellen 3D Visualisierungen der anatomischen Zielstruktur über die 2) virtuelle Vorbereitung und Modifikation von druckbaren 3D Modellen bis zum 3) 3D Druckvorgang selbst bzw. bis zur realen Modellherstellung für die 4) Adaptation von Osteosynthesematerialien oder patientenspezifischen Implantaten. 5)

Die Anwendung der Methode muss dabei nicht zwingend auf den Unterkiefer beschränkt sein, da der GrowCut-Algorithmus einerseits bereits erfolgreich zur Segmentierung anderer Strukturen in anderen medizinischen Fachgebieten eingesetzt wurde (48) und andererseits die in dieser Studie erhobenen Ergebnisse auf weitere anatomische Regionen im Gesichtsschädels übertragen werden können, sofern eine solide Knochenstruktur vorhanden ist. Die Ergebnisse der vorliegenden Studie unterstützend wurden ähnliche klinisch-ausreichende Übereinstimmungsparameter ebenfalls von Egger et al. bei der semiautomatischen Segmentierung von Glioblastomen unter der Anwendung des GrowCut-Algorithmus beobachtet. (48)

Weiterführend unterschieden sich die in der vorliegenden Studie erhobenen Werte des Volumens und der Anzahl der Voxel der semiautomatisch segmentierten Unterkiefer nicht signifikant ($p < 0.05$) von den Werten der Volumina und der Anzahl

der Voxel in der Kontrollgruppe (Ground Truth A, B). Eine hohe direkte Proportionalität nahe dem Wert 1 war ebenfalls erkennbar. Zudem konnte in der Regressionsanalyse kein signifikanter Unterschied hinsichtlich der Anzahl der Voxel und des Volumens beim Vergleich zwischen der semiautomatischen und der manuellen Segmentierung festgestellt werden ($p < 0.01$).

Der Vergleich der beiden Ground Truth Kontrollgruppen (A, B) miteinander zeigte eine sehr hohe Übereinstimmung durch hohe Dice-Score Werte im Mittel von 94% und durch eine niedrige Hausdorff-Distanz im Mittel von 4 Voxel. Die Anzahl der Voxel sowie das Volumen unterschieden sich beim direkten Vergleich der beiden Kontrollgruppen A und B ebenfalls nicht signifikant ($p < 0.05$). Eine hohe Übereinstimmung der beiden Ground Truth Segmentierungen kann somit bewiesen werden. Diese Ergebnisse bestätigen, dass die manuellen Segmentierungsmodelle, die von zwei klinischen Experten unabhängig voneinander durchgeführt wurden, als valide Ground Truth und somit als objektive kontrollierte Werte der tatsächlich-visualisierten Anatomie angenommen werden können. Aus diesem Grund hat die Kontrollgruppe ausreichend Validität für einen direkten Vergleich mit der untersuchten semiautomatischen Segmentierungsmethode.

Die Zeit für die manuelle Segmentierung lag im Durchschnitt bei 38.4 Minuten. Die Zeit für die semiautomatische Segmentierung betrug im Gegensatz nur 1 Minute. Die Ergebnisse in dieser Studie bestätigen neben der Validität der Ground Truth Daten, dass es möglich ist, mithilfe des GrowCut-Algorithmus die komplette anatomische Struktur des knöchernen Unterkiefers in kurzer Zeit mit hoher Genauigkeit und Präzision zu segmentieren und anschließend virtuell 3D darzustellen.

Aufgrund der in dieser Arbeit erhobenen Ergebnisse kann die Hypothese der Studie bestätigt werden. Die Genauigkeit, Übereinstimmung und Funktionalität einer konventionellen, lizenzfreien, semiautomatischen, open-source Segmentierungsmethode unterschieden sich nicht signifikant von der Ground Truth derselben anatomischen Struktur. Die semiautomatischen Segmentierungsergebnisse waren im Vergleich zu den erstellten validen Ground Truth Modellen in keinem der definierten Parameter signifikant unterschiedlich. Unterstützt wird dies anhand von hohen Übereinstimmungswerten durch hohe Dice-Score-Koeffizienten und niedrige Hausdorff-Distanzen zwischen der

untersuchten Methode (GrowCut) und der tatsächlichen visualisierten anatomischen Struktur (Ground Truth).

Neben diesen erhobenen Ergebnissen hinsichtlich der getesteten Methode wurden in dieser Studie aber auch bestimmte Limitationen festgestellt. 1) Es wurde versucht, die Ground Truth Daten so präzise wie möglich zu generieren, indem zwei klinische Experten den kompletten Unterkiefer jeweils einzeln segmentierten. Die erstellten Ground Truth Modelle wurden mithilfe von statistischen Berechnungen miteinander verglichen und können auf Grund des berechneten Ergebnisses als objektiv valide Kontrollgruppe angesehen werden. Dennoch bringt die Erstellung eines Ground Truth Modells eine bestimmte untersucherabhängige Ungenauigkeit und Subjektivität mit sich, die in dieser Studie jedoch so gering wie möglich gehalten wurde. 2) Obwohl die CT-Datensätze aus einem selektierten Datenpool zufällig anhand von definierten Ein- und Ausschlusskriterien ausgewählt wurden, wäre eine höhere Anzahl an vollständigen segmentierten Datensätzen hinsichtlich der Hypothesenprüfung möglicherweise aussagekräftiger. Da es sich jedoch um eine explorative Studie handelt, die in dieser Form bislang noch nicht durchgeführt wurde, wurden in dieser Untersuchung vorerst 10 randomisierte CT-Datensätze ausgewählt. 3) In dieser Studie wurden ausschließlich Datensätze mit zahnlosen Unterkiefen verwendet. Dies hat den Grund, dass Datensätze mit bezahnten Unterkiefen oft enorme radiologische Artefakte enthalten und deshalb unvollständige oder nicht verwendbare CT-Datensätze entstehen, aus denen eine große Menge an Schichten für eine korrekte Bildbearbeitung gelöscht werden müssen. Um sicherzustellen, dass sämtliche Schichten und die kompletten anatomischen Zielstrukturen in jedem Datensatz segmentiert werden können sowie ebenfalls ein homogenes Datenkollektiv in der Erstellung der Ground Truth Modelle erzeugt werden kann, wurden rein zahnlose, vollständige, physiologisch-erhaltene Unterkiefer CT-Scans ausgewählt. 4) Ein Problem für die Segmentierung stellte die anatomische Region der Unterkieferkondylen dar, da diese Region mit den umgebenden anatomischen Strukturen (Schädelbasis etc.) in sehr enger Lagebeziehung liegt und eine präzise Trennung eine Herausforderung an das Segmentierungsverfahren stellt. Daher ist es vor allem bei der manuellen Segmentierung zur Erstellung der Ground Truth Modelle schwer, die korrekten Grenzen zu definieren. Die auf diese Weise entstandenen Ungenauigkeiten waren entsprechend der durchgeführten Berechnungen statistisch jedoch nicht

signifikant. 5) In dieser Studie wurden lediglich nicht CE-geprüfte Softwareprogramme verwendet. Diese kommerziellen lizenzgestützten Softwarepakete, wie zum Beispiel Brainlab®, Materialise®, CMF Module®, Maxilim® oder IPS Case Designer®, sind in großen klinischen Zentren zwar teilweise Goldstandard, wurden in dieser Studie aber bewusst nicht berücksichtigt, um die Genauigkeit, Funktionalität und praktische Durchführbarkeit einer einfachen open-source Segmentierungsmethode zu untersuchen. Derartige Untersuchungen liegen in der Literatur weit weniger vor, als jene in denen große CE geprüfte Softwarepakete evaluiert werden, obwohl die Zugänglichkeit, der Erwerb und die Nutzbarkeit einer konventionellen lizenzfreien Methode wesentlich einfacher und unaufwendiger ist. Das gegenwärtige bestehende Interesse in der klinischen Anwendung von nicht-CE geprüften Softwareprogrammen zeigt ebenfalls die Literatur, da durchaus Arbeiten existieren, die, ähnlich wie in der vorliegenden Studie, die Genauigkeit von open-source Methoden zur Segmentierung evaluieren (44).

Trotz der beschriebenen Limitationen korrelieren die in dieser Studie gefundenen Ergebnisse mit anderen Untersuchungen in der Literatur, in denen Segmentierungsmethoden mit ausreichender Genauigkeit in anderen anatomischen Regionen getestet wurden. (49–51) Ähnlich wie Szymor et al., die in ihrer Untersuchung ebenfalls die Software-Plattform 3D-Slicer anhand gedruckter 3D-Modelle auf Genauigkeit evaluierten (44), empfiehlt die vorliegende Studie, dass die Genauigkeit und praktische Durchführbarkeit des bildbasierten open-source Segmentierungsverfahrens GrowCut für die klinische Anwendung ausreichend ist.

Zusammenfassend bieten bildbasierte (semi)automatische Segmentierungsverfahren und die Erstellung virtueller 3D-Modelle mithilfe von Softwareprogrammen viele gegenwarts- und zukunftsorientierte Möglichkeiten in der klinischen Anwendung der Bilddatenanalyse. Da die Segmentierungsmethode in dieser Studie mit einem frei zugänglichen open-source Softwareverfahren durchgeführt wurde, entstehen für den/die Benutzer/in keine monetären Abgaben im Sinne von Lizenzgebühren oder Folgekosten für laufende Anwendungen. Aufgrund der konventionellen Verfügbarkeit ist die Anwendung in weiteren klinischen Einrichtungen oder in weiteren wissenschaftlichen Arbeitsgruppen möglich. Die Anwendungsmöglichkeiten der Methode sind vielfältig. So können

mithilfe der untersuchten bildbasierten Segmentierungsmethode 3D Visualisierungen und virtuelle oder reale 3D-Modelle unkompliziert und zeitsparend erstellt werden und zur Simulation von Operationstechniken in der Therapieplanung oder im Lehrbetrieb genutzt werden. Weiters besteht die Möglichkeit der präoperativen anatomischen Adaptation von individuellen Osteosynthesematerialien oder patientenspezifischen Implantaten. Ein zusätzliches intraoperatives Biegen oder Anpassen der Materialien ist folglich nur mehr in geringem Maße notwendig. Auf diese Weise wird Operationszeit gespart, das postoperative Outcome und die Behandlungsqualität auf Grund von präziser anatomischer Materialadaptation verbessert. Weiters können dem/der Patienten/in mithilfe individuell gedruckter 3D-Modelle chirurgische Eingriffe, wie Umstellungsosteotomien der Kiefer, Osteosynthesen bei Frakturen oder das Setzen von Implantaten, individuell erklärt, veranschaulicht oder simuliert werden. Der Lehrbetrieb und der Trainingszweck sind weitere praktische Anwendungsgebiete, in denen virtuelle oder reale segmentierte 3D-Modelle anatomischer Zielstrukturen eingesetzt werden können.

Ausblickend könnten im Rahmen von weiterführenden zukünftigen Projekten andere anatomische Strukturen des Gesichtsschädels in weiteren Testreihen mit der untersuchten Methode segmentiert werden. Ebenso könnte das Datenkollektiv für eine höhere Aussagekraft der Ergebnisse vergrößert werden. Weiterführend könnte die GrowCut-Segmentierungsmethode mit anderen frei zugänglichen konventionellen oder auch kommerziellen lizenzierten Algorithmen verglichen werden.

Schlussfolgernd stellt der untersuchte lizenzfreie GrowCut-Algorithmus eine einfache und zeitsparende ubiquitär verfügbare Methode zur Segmentierung von soliden Knochenstrukturen dar. Die durchgeführten Segmentierungsergebnisse lassen auf hohe Genauigkeit und Präzision des Verfahrens schließen. Die Anwendung der Methode und der klinische Einsatz sowie die praktische Durchführung dieser Vorgehensweise können unabhängig von monetären oder lizenzierten Ressourcen reproduziert werden. Die untersuchte Methode und vor allem die erzeugten Ground Truth Datenmodelle bieten enormes Potenzial für weiterführende kontrollierte Untersuchungen und Überprüfungen auf diesem Gebiet. Zusätzlich könnten die erzeugten Ground Truth Datenmodelle von anderen Einrichtungen oder Forschungsgruppen als Kontrollgruppe zur

vergleichenden Beurteilung von Segmentierungsverfahren übernommen werden. Die Segmentierungsmethode könnte ebenfalls unter geringerer Strahlenbelastung, im Vergleich zur konventionellen CT, auf Basis von DVT-Datensätzen Anwendung finden. Durch den Einsatz von modernsten CT-Geräten und die Elimination von radiologischen Artefakten in Echtzeit könnten weiterführend auch Bilddatensätze mit dentogenen Implantaten, metallischen Zahnfüllungen oder Osteosynthesematerialien in die Evaluierung und Bewertung dieser oder ähnlicher Segmentierungsmethoden miteingeschlossen werden.

12 Zusammenfassung der wichtigsten Fakten der Arbeit:

- Durchführung von zweifacher manueller schichtweiser Segmentierung des Unterkiefers durch klinische Experten und Erstellung von virtuellen 3D Ground Truth Modellen (Kontrollgruppe).
- Durchführung von semiautomatischer open-source Softwaresegmentierung mithilfe des GrowCut-Algorithmus und Erstellung von virtuellen 3D Segmentierungsmodellen (Segmentierungsgruppe).
- Vergleichende Beurteilung der beiden Gruppen anhand objektiv messbarer Parameter. Überprüfung der Validität der Ground Truth Daten.
- Hohe Genauigkeit, Übereinstimmung, Funktionalität und geringer Zeitaufwand im Vergleich zur Kontrollgruppe sprechen für eine klinische Anwendung der untersuchten Segmentierungsmethode. Ground Truth Daten werden anhand von statistischer Beurteilung als valide und objektiv angesehen.
- In dieser Art und Weise wurde in der Literatur bislang keine lizenzfreie open-source Segmentierungsmethode im Bereich der Mund-Kiefer und Gesichtschirurgie auf Genauigkeit, Praktikabilität und klinische Anwendung getestet.
- Die Ground Truth Datensätze des Unterkiefers, die durch die zweifache manuelle Segmentierung erstellt wurden und statistisch als valide eingestuft wurden, könnten zu Forschungszwecken und Prüfung anderer bildbasierter Softwareprozesse und Segmentierungsverfahren verwendet werden.

13 Literaturverzeichnis

1. Kramme R, editor. Medizintechnik: Verfahren, Systeme, Informationsverarbeitung. 4. Aufl. Heidelberg: Springer; 2011.
2. Mazzoni S, Badiali G, Lancellotti L, Babbi L, Bianchi A, Marchetti C. Simulation-guided navigation: a new approach to improve intraoperative three-dimensional reproducibility during orthognathic surgery. *The Journal of craniofacial surgery* 2010; 21(6):1698–705.
3. Sharma N, Aggarwal LM. Automated medical image segmentation techniques. *Journal of medical physics / Association of Medical Physicists of India* 2010; 35(1):3–14.
4. Stefan Kalcher. Segmentierung in der medizinischen Bildverarbeitung [cited 2016 Jun 18]. Available from: URL: <http://www.student.tugraz.at/stefan.kalcher/ewa/segmentierung.pdf>.
5. Jiang H, He B, Ma Z, Zong M, Zhou X, Fujita H. Liver segmentation based on Snakes Model and improved GrowCut algorithm in abdominal CT image. *Computational and mathematical methods in medicine* 2013; 2013:958398.
6. Chen X, Xu L, Yang Y, Egger J. A semi-automatic computer-aided method for surgical template design. *Scientific reports* 2016; 6:20280.
7. Mandibula Beispielbild [cited 2017 Dec 17]. Available from: URL: <https://radiologypics.com/2013/01/19/mandibular-osteomyelitis/>.
8. Olabarriga SD, Smeulders AW. Interaction in the segmentation of medical images: a survey. *Medical image analysis* 2001; 5(2):127–42.
9. Egger J, Colen RR, Freisleben B, Nimsky C. Manual refinement system for graph-based segmentation results in the medical domain. *Journal of medical systems* 2012; 36(5):2829–39.
10. Cevidanes, Lucia H C, Tucker S, Styner M, Kim H, Chapuis J, Reyes M et al. Three-dimensional surgical simulation. *American journal of orthodontics and dentofacial orthopedics: official publication of the American Association of Orthodontists, its constituent societies, and the American Board of Orthodontics* 2010; 138(3):361–71.

11. Raith S, Wolff S, Steiner T, Modabber A, Weber M, Hölzle F et al. Planning of mandibular reconstructions based on statistical shape models. *International journal of computer assisted radiology and surgery* 2016.
12. Zhao L, Patel PK, Cohen M. Application of virtual surgical planning with computer assisted design and manufacturing technology to cranio-maxillofacial surgery. *Archives of plastic surgery* 2012; 39(4):309–16.
13. Feng Zhao and Xianghua Xie. An Overview of Interactive Medical Image Segmentation [cited 2016 Sep 6]. Available from: URL: <http://www.bmva.org/annals/2013/2013-0007.pdf>.
14. Egger J. Refinement-cut: user-guided segmentation algorithm for translational science. *Scientific reports* 2014; 4:5164.
15. Jehle R, Czeschik JC, Freund T. *Medizinische Informatik kompakt: Ein Kompendium für Mediziner, Informatiker, Qualitätsmanager und Epidemiologen*; 2015. (De Gruyter Studium) [cited 2017 Sep 15]. Available from: URL: <http://dx.doi.org/10.1515/9783110340259>.
16. Kramme R. *Informationsmanagement und Kommunikation in der Medizin*; 2017. Available from: URL: <http://dx.doi.org/10.1007/978-3-662-48778-5>.
17. Müller S. *Extraktion baulich geprägter Flächen aus Fernerkundungsdaten zur Qualitätssicherung flächenhafter Geobasisdaten [Zugl.: Hannover, Univ., Diss., 2007]*. Stuttgart: Ibidem-Verl; 2007. (Wissensbasierte Bildauswertung; vol 9).
18. Handels H, Horsch A, Lehmann T, Meinzer H. *Bildverarbeitung für die Medizin 2001: Algorithmen - Systeme - Anwendungen*. Berlin, Heidelberg: Springer; 2001. (Informatik aktuell). Available from: URL: <http://dx.doi.org/10.1007/978-3-642-56714-8>.
19. Chanda B, Dutta Majumder D. *Digital image processing and analysis*. Eastern economy ed. New Delhi: Prentice-Hall of India; 2002.
20. Tönnies KD. *Grundlagen der Bildverarbeitung*. München u.a.: Pearson Studium; 2005. (Informatik: Bildverarbeitung).
21. Jiang X, Bunke H. *Dreidimensionales Computersehen: Gewinnung und Analyse von Tiefenbildern*. Berlin, Heidelberg: Springer Berlin Heidelberg; 1997. Available from: URL: <http://dx.doi.org/10.1007/978-3-642-61447-7>.

22. Prokop M. Ganzkörper-Computertomographie: Spiral- und Multislice-CT; mit 328 Tabellen. 2., vollst. überarb. und erw. Aufl. Stuttgart: Thieme; 2007. (RRR - Referenz-Reihe Radiologie).
23. Handels H. Medizinische Bildverarbeitung: Bildanalyse, Mustererkennung und Visualisierung für die computergestützte ärztliche Diagnostik und Therapie. 2., überarb. und erw. Aufl. Wiesbaden: Vieweg+Teubner; 2009. (Studium). Available from: URL: <http://dx.doi.org/10.1007/978-3-8348-9571-4>.
24. Song H, Kang W, Zhang Q, Wang S. Kidney segmentation in CT sequences using SKFCM and improved GrowCut algorithm. BMC systems biology 2015; 9 Suppl 5:S5.
25. Egger J, Kapur T, Fedorov A, Pieper S, Miller JV, Veeraraghavan H et al. GBM volumetry using the 3D Slicer medical image computing platform. Scientific reports 2013; 3:1364.
26. Slicer [cited 2017 Nov 4]. Available from: URL: <https://www.slicer.org/>.
27. Mevislab [cited 2016 Sep 4]. Available from: URL: <http://www.mevislab.de/>.
28. Cheng I, Sun X, Alsufyani N, Xiong Z, Major P, Basu A. Ground truth delineation for medical image segmentation based on Local Consistency and Distribution Map analysis. Conference proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference 2015; 2015:3073–6.
29. Dou TH, Thomas DH, O'Connell DP, Lamb JM, Lee P, Low DA. A Method for Assessing Ground-Truth Accuracy of the 5DCT Technique. International journal of radiation oncology, biology, physics 2015; 93(4):925–33.
30. Lampert TA, Stumpf A, Gancarski P. An Empirical Study Into Annotator Agreement, Ground Truth Estimation, and Algorithm Evaluation. IEEE transactions on image processing: a publication of the IEEE Signal Processing Society 2016; 25(6):2557–72.
31. Bankman IN, editor. Handbook of medical image processing and analysis. 2. ed. Amsterdam u.a.: Elsevier / Acad. Press; 2009. (Academic Press series in biomedical engineering).

32. Randomizer [cited 2017 Oct 15]. Available from: URL: <https://www.randomizer.at/>.
33. Tian K, Yang X, Kong Q, Yin C, He RL, Yau SS. Two Dimensional Yau-Hausdorff Distance with Applications on Comparison of DNA and Protein Sequences. *PloS one* 2015; 10(9):e0136577.
34. Zou KH, Warfield SK, Bharatha A, Tempany CM, Kaus MR, Haker SJ et al. Statistical validation of image segmentation quality based on a spatial overlap index. *Academic Radiology* 2004; 11(2):178–89.
35. Tucker S, Cevidanes, Lucia Helena Soares, Styner M, Kim H, Reyes M, Proffit W et al. Comparison of actual surgical outcomes and 3-dimensional surgical simulations. *Journal of oral and maxillofacial surgery: official journal of the American Association of Oral and Maxillofacial Surgeons* 2010; 68(10):2412–21.
36. Byrne N, Velasco Forte M, Tandon A, Valverde I, Hussain T. A systematic review of image segmentation methodology, used in the additive manufacture of patient-specific 3D printed models of the cardiovascular system. *JRSM cardiovascular disease* 2016; 5:2048004016645467.
37. Yildirgan K, Zahir E, Sharafi S, Ahmad S, Schaller B, Ricklin ME et al. Mandibular Fractures Admitted to the Emergency Department: Data Analysis from a Swiss Level One Trauma Centre. *Emergency medicine international* 2016; 2016:3502902.
38. Ramisetty S, Gaddipati R, Vura N, Pokala S, Kapse S. Maxillofacial Injuries in Women: A Retrospective Study of 10 Years. *Journal of maxillofacial and oral surgery* 2017; 16(4):438–44.
39. Thor A. Preoperative Planning of Virtual Osteotomies Followed by Fabrication of Patient Specific Reconstruction Plate for Secondary Correction and Fixation of Displaced Bilateral Mandibular Body Fracture. *Craniofacial trauma & reconstruction* 2016; 9(2):188–94.
40. Nakao M, Aso S, Imai Y, Ueda N, Hatanaka T, Shiba M et al. Automated Planning with Multivariate Shape Descriptors for Fibular Transfer in Mandibular Reconstruction. *IEEE transactions on bio-medical engineering* 2016.
41. Abdolali F, Zoroofi RA, Abdolali M, Yokota F, Otake Y, Sato Y. Automatic segmentation of mandibular canal in cone beam CT images using conditional

statistical shape model and fast marching. *International journal of computer assisted radiology and surgery* 2017; 12(4):581–93.

42. Fabri D, Zambrano V, Bhatia A, Furtado H, Bergmann H, Stock M et al. A quantitative comparison of the performance of three deformable registration algorithms in radiotherapy. *Zeitschrift für medizinische Physik* 2013; 23(4):279–90.

43. Huttenlocher DP, Klanderman GA, Rucklidge WJ. Comparing images using the Hausdorff distance [cited 2017 Dec 17]. Available from: URL: https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/classes/cs294_people_places_things/2004sp/cs294/Huttenlocher93.pdf.

44. Szymor P, Kozakiewicz M, Olszewski R. Accuracy of open-source software segmentation and paper-based printed three-dimensional models. *Journal of cranio-maxillo-facial surgery: official publication of the European Association for Cranio-Maxillo-Facial Surgery* 2016; 44(2):202–9.

45. Huotilainen E, Jaanimets R, Valášek J, Marcián P, Salmi M, Tuomi J et al. Inaccuracies in additive manufactured medical skull models caused by the DICOM to STL conversion process. *Journal of cranio-maxillo-facial surgery: official publication of the European Association for Cranio-Maxillo-Facial Surgery* 2014; 42(5):e259-65.

46. Naumovich SS, Naumovich SA, Goncharenko VG. Three-dimensional reconstruction of teeth and jaws based on segmentation of CT images using watershed transformation. *Dento maxillo facial radiology* 2015; 44(4):20140313.

47. Fu Y, Liu S, Li H, Yang D. Automatic and hierarchical segmentation of the human skeleton in CT images. *Physics in medicine and biology* 2017; 62(7):2812–33.

48. Hu Y, Grossberg M, Mageras G. Semiautomatic tumor segmentation with multimodal images in a conditional random field framework. *Journal of medical imaging (Bellingham, Wash.)* 2016; 3(2):024503.

49. Abdullah M, Fraz MM, Barman SA. Localization and segmentation of optic disc in retinal images using circular Hough transform and grow-cut algorithm. *PeerJ* 2016; 4:e2003.

50. Neelapu BC, Kharbanda OP, Sardana V, Gupta A, Vasamsetti S, Balachandran R et al. A pilot study for segmentation of pharyngeal and sino-nasal

airway subregions by automatic contour initialization. *International journal of computer assisted radiology and surgery* 2017.

51. Kostopoulou E, Katsigiannis S, Maroulis D. 2D-gel spot detection and segmentation based on modified image-aware grow-cut and regional intensity information. *Computer methods and programs in biomedicine* 2015; 122(1):26–39.