



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원 저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리와 책임은 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)





자동 치아 분할 방식에 따른
디지털 치아 모델 분석의 정확도 및 효율성

연세대학교 대학원
치의학산업학과
유 혜 리

자동 치아 분할 방식에 따른
디지털 치아 모델 분석의 정확도 및 효율성

지도교수 차 정 열

이 논문을 석사 학위논문으로 제출함

2025년 06월

연세대학교 대학원

치의학산업학과

유 혜 리

자동 치아 분할 방식에 따른
디지털 치아 모델 분석의 정확도 및 효율성
유혜리의 석사 학위논문으로 인준함

심사위원장 유 형 석

심사위원 차 정 열

심사위원 황 충 주

연세대학교 대학원
치의학산업학과

2025년 06월

감사의 글

대학원 과정을 통해 회사 생활만으로는 얻을 수 없었던 소중한 배움을 쌓고 한 단계 성장할 수 있었음에 감사드립니다. 이 소중한 경험은 앞으로의 삶에 큰 도움이 될 것이라 생각합니다. 무엇보다 많은 분들의 따뜻한 격려와 도움이 있었기에 석사 과정을 잘 마칠 수 있었습니다.

먼저, 연구의 시작부터 끝까지 아낌없는 지도와 따뜻한 격려로 이끌어 주신 차정열 교수님께 진심으로 감사드립니다. 바쁘신 진료와 연구 중에도 귀한 시간을 내어 부족한 제 논문을 꼼꼼히 검토해 주시고, 끝까지 완성할 수 있도록 이끌어 주신 교수님의 가르침 덕분에 이 논문을 완성할 수 있었습니다.

다음으로, 학교와 회사에서 언제나 아낌없는 조언과 응원으로 힘이 되어 주시는 황충주 교수님께 감사드리며, 별 때마다 따뜻하게 맞아 주시고 늘 격려해 주신 유형석 교수님께도 감사의 마음을 전합니다.

그리고 늘 곁에서 믿고 응원해 준 엄마, 아빠, 그리고 동생 혜지에게도 고마움을 전합니다. 가족의 따뜻한 사랑과 지지 덕분에 이 논문을 끝까지 마칠 수 있었습니다.

마지막으로, 바쁜 일상 속에서도 한결같은 응원과 믿음으로 늘 힘이 되어 준 하나님뿐인 남편에게도 고마운 마음을 전합니다. 덕분에 이 논문을 잘 마무리할 수 있었고, 이 순간의 기쁨을 함께 할 수 있어 행복합니다.

2025년 6월

유혜리



차례

표 차례	ii
그림 차례	iii
국문 요약	iv
제1장 서론	1
제2장 연구 대상 및 방법	4
2.1. 연구 대상	4
2.1.1. 포함 기준	4
2.1.2. 제외 기준	4
2.2. 연구 방법	5
2.2.1. 연구 대상 소프트웨어	5
2.2.2. 평가 방법	7
2.3. 데이터 분석 방법 및 통계 처리	10
제3장 연구 결과	11
3.1. 연구 참여자의 일반적 특성	11
3.2. 분할 성공률	11
3.3. 분할 소요 시간	13
3.4. 근원심 폭경	15
3.4.1. 치아별 비교	15
3.4.2. 수동 분할군(MS)과 비교	19
3.5. Bolton 비율	21
3.6. 초기 배열 정확도	23
제4장 고찰	25
제5장 결론	28
참고 문헌	29
영문 요약	32



LIST OF TABLES

<Table 1> Classification of tooth segmentation methods by software	6
<Table 2> Success rate of tooth segmentation by software	13
<Table 3> Comparison of tooth segmentation time by software	14
<Table 4> Mesiodistal widths of maxillary teeth (U1-U7).....	16
<Table 5> Mesiodistal widths of mandibular teeth (L1-L7).....	17
<Table 6> Anterior Bolton ratios and tooth size sums by software.....	22
<Table 7> Overall Bolton ratios and tooth size sums by software	22
<Table 8> Comparison of Irregularity Index by arch and software.....	24



LIST OF FIGURES

<Fig 1> Tooth segmentation workflows of DS1, DS2, LS.....	6
<Fig 2> Examples of segmentation errors in digital tooth models	7
<Fig 3> Little's Irregularity Index	9
<Fig 4> Segmentation results from each software.....	12
<Fig 5> Comparison of mesiodistal width by tooth and software	18
<Fig 6> Bland-Altman plot comparing REF with each segmentation software	21

국 문 요 약

자동 치아 분할 방식에 따른 디지털 치아 모델 분석의 정확도 및 효율성

본 연구는 교정 진단과 치료 계획 수립에서 디지털 셋업의 정확도와 효율성을 평가하기 위해, 세 가지 상업용 CAD 소프트웨어(Onyxceph [DS1], OneOrtho [DS2], Autolign [LS])의 자동 치아 분할 방식에 따른 디지털 치아 모델 분석과 자동 치아 배열 정확도를 비교·분석하였다. 투명교정 치료 환자의 상·하악 디지털 치아 모델 15 쌍을 대상으로 하였으며, 분할 성공률, 분할 소요 시간, 근원심 폭경(mesiodistal width), Bolton 비율(Bolton ratio), 초기 배열 정확도(initial alignment accuracy)를 수동 치아 분할군(MS)과 비교, 평가하였다.

분석 결과, DS1은 높은 분할 성공률(98.3%)과 짧은 분할 소요 시간(35.9 초)을 보여 효율성과 정확성이 높았으며, 근원심 폭경 측정에서도 수동 분할 작업군과 유사한 결과를 보였다. DS2는 높은 분할 성공률을 보였지만 일부 치아에서 계측값이 과대 측정되었고, 처리 시간의 변동 폭이 크게 나타났다. LS는 세 소프트웨어 중 가장 낮은 분할 성공률과 긴 처리 시간을 보였으며, 일부 치아에서 계측값이 과대 측정되는 경향을 보였다. Bolton 비율은 소프트웨어 간 유의한 차이가 없었으며, 초기 배열 정확도 분석에서는 DS1이 가장 낮은 Irregularity Index를 보여 상대적으로 우수한 배열 정확도를 나타냈다.

이러한 결과는 자동 치아 분할 기능이 디지털 셋업 과정의 진단과 배열에 직접적인 영향을 미치며, 임상 적용 시 소프트웨어별 알고리즘 구조와 특성을 충분히 이해하고 활용하는 것이 중요함을 보여준다. 본 연구는 자동 치아 분할 및 배열 기능의 임상적 활용 가능성을 확인하였으며, 향후 다양한 임상 조건과 복잡한 사례를



포함한 연구를 통해 자동 치아 분할 기술의 임상적 신뢰성과 활용 범위를 확장하는데 기여할 수 있을 것이다.

핵심 되는 말: 자동 치아 분할, 디지털 치아 모델 분석, 디지털 셋업, CAD 소프트웨어, 분할 성공률, 분할 소요 시간, 근원심 폭경, Bolton 비율, 배열 정확도

1. 서론

교정 치료에서 진단 셋업은 발치 여부, 치간 삭제(interproximal reduction), 고정원 설정, 치료 기법 설정 등의 사전 평가와 정밀한 치료 계획 수립에 활용될 수 있다(Kesling, 1946). 디지털 셋업은 진단과 치료계획 수립에서 유용하게 사용될 수 있으며(Macchi et al., 2006), 디지털 셋업 진단 중례 중 23.6%에서 실제 치료 계획이 변경되어 정확한 진단과 구체적인 치료 계획 수립에 기여하며, 특히 복잡한 중례에서 유용하고, 임상 경험 수준과 관계 없이 치료 계획의 일관성과 신뢰도를 높여 경험이 부족한 임상의도 안정적인 치료 결과를 도출할 수 있게 하는 것으로 보고되었다(Hou et al., 2020). 최근에는 인공지능 기술이 교정 치료 계획 수립, 악교정 진단, 매복 치아 진단, 성장 단계 판단 등 다양한 분야에 적용되며, 디지털 교정의 진단 및 치료 계획의 효율성을 향상시키고 있다(Kok et al., 2019; Suhail et al., 2020). 디지털 교정에 활용되는 상업용 CAD/CAM 소프트웨어로는 OrthoAnalyzer(3Shape, Denmark), SureSmile(Dentsply Sirona, USA), OrthoStudio(Maestro, Italy), ClinCheck(Align Technology, USA), Autolign(Diorco, Korea), OneOrtho(Osstem Implant, Korea) 등이 있으며, 이들 프로그램은 디지털 셋업 기반의 진단 및 시뮬레이션 기능을 제공하여 치료 계획 수립에 사용되고 있고(Lee et al., 2020; Xie et al., 2010), clear aligner 치료, IDB Jig 제작, 맞춤형 브라켓 및 와이어 설계 등 디지털 기반 교정 치료의 다양한 임상 적용에도 활용되고 있다(Hou et al., 2020).

디지털 셋업의 정확도를 결정짓는 핵심 요소는 정확한 치아 분할(tooth segmentation)이다. 분할 과정의 초기 단계에서 오류가 발생하면 이후 모델 분석, 치아 배열, 장치 제작 등 전체 진단 및 치료 결과에 영향을 미치며, 결과적으로 임상 적용의 신뢰성을 저하시킬 수 있다. 기존의 수동 치아 분할(manual tooth segmentation)은 높은 숙련도를 요구하고 작업 시간이 길며, 작업자간 편차가 생기는 한계를 가진다. 특히 치아의 해부학적 다양성, 치은과 치아 사이의 모호한 경계, 인접 치아 간의 긴밀한 접촉부 등은 정확한 치아 분할을 어렵게 만드는

대표적인 임상적 요인으로 작용한다. 이와 같은 한계를 극복하기 위해 딥러닝 기반의 자동 치아 분할 기술이 도입되어 임상 활용이 확대되고 있으며, 수작업에 비해 작업 효율을 높이면서도 일정 수준의 정확도를 유지할 수 있어 임상에서 유용하게 사용되고 있다.

반면, 자동화된 디지털 분할 시스템도 본질적인 한계를 가지고 있다. 디지털 셋업 과정에서 치아 분할 정확도가 최종 배열 결과 및 분석의 신뢰도에 큰 영향을 미치며, 이에 대한 사례로 디지털 셋업에서 근원심 폭경(MD width)이 실제보다 작게 측정되는 경향이 보고되었다(Gonzalez Guzman & Teramoto Ohara, 2019). 또한 디지털 셋업 모델의 치열궁 둘레(arch perimeter)가 석고 모형의 셋업보다 짧게 측정되는 경향이 있으며(Im et al., 2014), 디지털 모델에서 측정한 총생량이 석고 모형보다 일관되게 낮게 나타난다는 보고도 있다(Yoon et al., 2018). 디지털 모델이 석고 모형에 비해 계측 정확도에서 차이를 보일 수 있음이 보고되며 두 방식 간 비교 검토의 필요성도 언급되었다(Fleming et al., 2011). 이러한 문제들은 디지털 환경에서의 계측 누적 오차 및 스캔 과정의 한계로 발생하며, 디지털 셋업을 활용할 때 반드시 고려해야 할 요소이다. 이러한 환경적 요인 외에도, 치아의 해부학적 다양성, 치은과 치아 사이의 불분명한 경계, 인접 치아 간의 긴밀한 접촉 등은 계측 오차를 유발한다. 특히 디지털 모델은 석고 모형과 달리 표면 스캔 데이터를 기반으로 생성된 hollow image 형태를 가지기 때문에 인접면 영역과 같이 스캔이 어려운 부위를 정밀하게 재현하기 어렵고, 이러한 해부학적 요인들이 치아 분할의 정밀도를 저해하는 주요 원인으로 보고된 바 있다(Tian, 2019). 이와 같은 구조적 한계는 디지털 셋업의 정밀도와 효율성에도 영향을 미치며, 소프트웨어에 따라 분할 정확도와 처리 방식이 달라질 수 있다. 실제로 디지털 셋업의 정확성과 효율성은 적용된 알고리즘의 구조에 따라 크게 좌우된다.

본 연구에서는 세 가지 상업용 CAD 소프트웨어의 치아 분할 방식에 따른 정확도 및 효율성을 비교하였다. Onxceph과 OneOrtho는 치아 번호 지정 및 분할 방식(Tooth Designation and Segmentation, DS)을 사용하여 각각 DS1과 DS2로 구분하였으며, Autolign은 랜드마크 기반 분할 방식(Landmark-based Segmentation,



LS)을 사용하여 LS로 구분하였다. 이들 소프트웨어는 사용자로부터 기준점 입력을 요구하므로 기술적으로는 ‘반 자동 분할’에 해당하지만, 최근 문헌에서는 관행적으로 ‘자동 치아 분할(automatic tooth segmentation)’로 분류하는 경향이 있어(Im et al., 2022) 본 연구에서도 이를 반영하여 기술하였다.

본 연구는 각 소프트웨어의 분할 성공률, 소요 시간, 근원심 폭경(MD width), Bolton 비율(Bolton ratio), 초기 배열 정확도(initial alignment accuracy)를 중심으로 정량적 분석을 수행하여, 디지털 치아 분할 방식에 따른 디지털 모델의 정확도와 효율성을 객관적으로 평가하고자 하였다. 본 연구의 귀무 가설은 “세가지 상업용 CAD 소프트웨어(DS1, DS2, LS) 간 분할 성공률, 소요 시간, 근원심 폭경, Bolton 비율, 초기 배열 정확도에 통계적으로 유의한 차이가 없다”는 것이다.

2. 연구 대상 및 방법

2.1. 연구 대상

본 연구는 연세대학교 치과대학병원 연구윤리심의위원회(Institutional review board, IRB)로부터 심의 면제를 받은 후 진행되었다(IRB 번호: 2-2025-0037). 투명교정 치료 환자의 디지털 치아 모델(3D 스캔 데이터)을 대상으로 후향적으로 진행하였으며, 총 15쌍의 상악 및 하악 케이스(총 30개 모델)가 포함되었다. 모든 환자는 Trios 스캐너(3Shape, Denmark)를 이용해 구강 내 디지털 인상을 채득하였으며, 각 케이스는 네 가지 상업용 CAD 소프트웨어를 통해 치아 분할을 수행하였다.

2.1.1. 포함 기준(Inclusion criteria)

- 가) 제 2 대구치가 맹출한 만 14 세 이상의 교정 치료 환자
- 나) 치아 및 치은에 결함이 없는 디지털 치아 모델
- 다) 비발치 교정 치료 대상 환자

2.1.2. 제외 기준(Exclusion criteria)

- 가) 왜소치(microdontia) 등 치아 형태 이상이 있는 환자
- 나) 심한 우식 또는 마모 치아가 포함된 환자
- 다) 선천적 혹은 후천적 결손치가 있는 환자
- 라) 과잉치(supernumerary tooth)가 있는 환자
- 마) 심한 치열 부조화(arch length discrepancy > 5 mm)가 있는 환자

2.2 연구 방법

2.2.1. 연구 대상 소프트웨어

본 연구에서는 네 가지 상업용 CAD 소프트웨어를 사용하여 디지털 치아 모델의 치아 분할 방식을 평가하였다. 각 소프트웨어의 기준점 설정 방식 및 주요 특징은 Table 1에 요약하였다. 자동 분할군의 분할 과정은 orientation, set point, segmentation의 세 단계로 구성되어 있으며, 워크플로우는 Figure 1에 제시하였다.

- 수동 치아 분할군(MS: Manual Tooth Segmentation): OrthoAnalyzer v1.11.2.0 (3Shape, Denmark). 숙련된 치과기공사가 각 치아의 mesial 및 distal 기준점을 수동으로 설정하고, 분할선을 직접 보정하여 수동 분할 모델로 제작하였다.
- 자동 치아 분할군
 - ① DS1(Tooth Designation and Segmentation – type 1): Onyxceph v3.2.65.71 (Image Instruments, Germany). 중심점(center point)을 지정하면 이를 기반으로 자동으로 치아 분할이 이루어지는 방식
 - ② DS2(Tooth Designation and Segmentation – type 2): OneOrtho v1.0.0.7 (Osstem Implant, Korea). DS1과 동일하게 중심점(center point)을 기준으로 치아 분할이 수행되는 방식
 - ③ LS(Landmark-based Segmentation): Autolign v1.6.5.6 (Diorco, Korea). mesial 및 distal point를 수동으로 지정하여 분할 수행하는 방식

분석은 Windows 10 Pro 64비트 운영체제, Intel Core i7-9700 프로세서(3.00 GHz), 16 GB DDR4 메모리, NVIDIA GeForce RTX 2060 GPU(6 GB VRAM)를 갖춘 컴퓨터에서 수행되었다.

Table 1. Classification of tooth segmentation methods by software

Software Group	Software	Manufacturer	Segmentation Method
MS	OrthoAnalyzer	3Shape, Denmark	Manual segmentation
DS1	Onyxceph	Image Instruments, Germany	Center point segmentation
DS2	OneOrtho	Osstem Implant, Korea	Center point segmentation
LS	Autolign	Diorco, Korea	MD point segmentation

Note. MD = Mesial-Distal

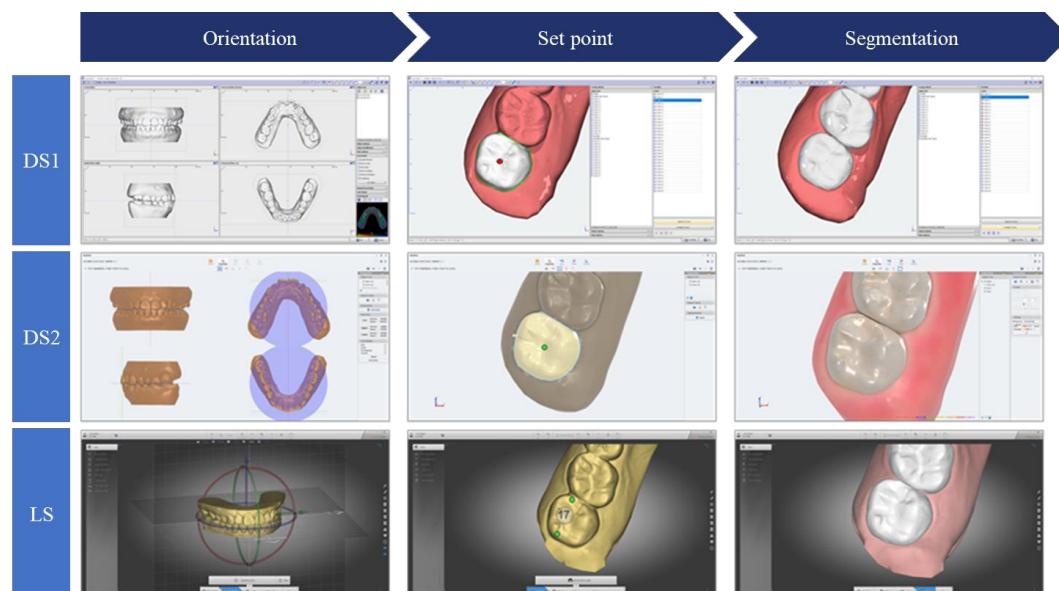


Figure 1. Tooth segmentation workflows of DS1, DS2, and LS. Each row (top to bottom) shows the software's three-step process: orientation, set point, and segmentation.

2.2.2. 평가 방법

본 연구는 네 가지 항목에 대해 소프트웨어 간 정량적 비교 분석을 수행하였다.

1) 분할 성공률

자동 치아 분할의 성공 여부는 다음 두가지 기준을 모두 만족하는 경우로 정의하였다.

- 소프트웨어가 해당 치아를 개별적으로 인식하고 번호를 부여한 경우
- 분할된 치아의 표면이 치은 경계(gingival margin)까지 정확히 포함되어 해부학적으로 완전한 형태를 유지한 경우(Yacout et al., 2024).

이 중 하나라도 충족하지 못한 경우에는 분할 실패로 간주하였다. 분할 실패 유형의 예시는 다음과 같다(Figure 2).

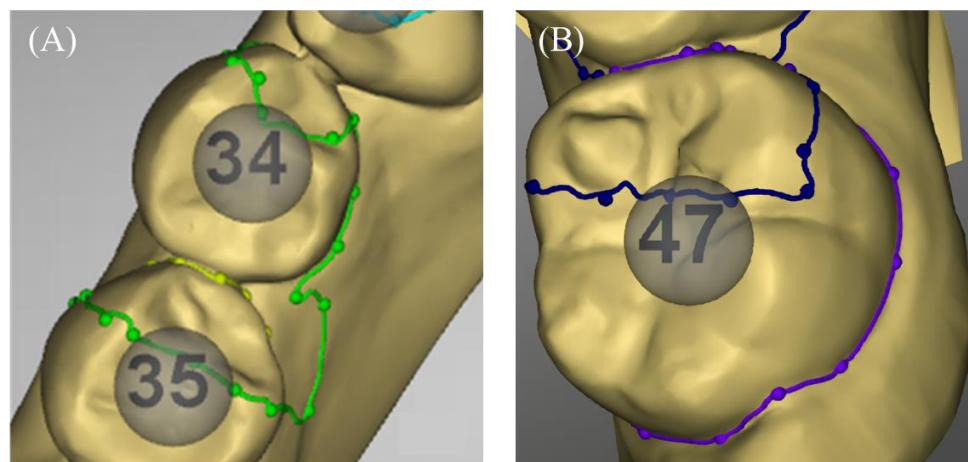


Figure 2. Examples of segmentation errors in digital tooth models.

- (A) Deviation of the segmentation line from the gingival-cervical margin.
(B) Deficiency segmentation on the occlusal surface of a posterior tooth.



2) 분할 소요 시간

각 소프트웨어에서 상악 및 하악 모델 전체를 분리하는 데 소요된 시간을 초 단위로 측정하였다. 타이머를 통해 분할 시작부터 종료 시점까지의 시간을 기록하였다.

3) 근원심 폭경

각 소프트웨어(DS1, DS2, LS)의 치아 분할 과정에서 자동으로 산출된 근원심 폭경 값을 수집하였다. 해당 측정값은 기준점 설정이나 수동 보정 없이 소프트웨어에서 제공된 값을 기반으로 분석에 활용하였다. 모든 대상 케이스에 대해 상·하악 각각 7개 치아를 기준으로 치아별 근원심 폭경의 평균값과 편차를 산출한 후, 세 소프트웨어(DS1, DS2, LS)간의 차이를 분석하였다. 이후, 각 소프트웨어의 측정값과 수동 분할군(MS) 간의 차이를 비교하였다.

4) Bolton 비율

소프트웨어에서 계산된 근원심 폭경 값을 바탕으로, 상악과 하악 치아의 폭경 합계를 계산한 후 전치부 및 전체 치아에 대해 다음 수식을 사용하여 Bolton 비율을 산출하였다.

$$\text{Anterior ratio (\%)} = \frac{\text{Sum of mandibular } 6}{\text{Sum of maxillary } 6} \times 100$$

$$\text{Overall ratio (\%)} = \frac{\text{Sum of mandibular } 12}{\text{Sum of maxillary } 12} \times 100$$

5) 초기 배열 정확도

각 소프트웨어의 자동 배열 기능(arch line generation)을 통해 전치부 치아의 배열을 수행한 후, Little's Irregularity Index를 적용하여 치아 정렬 상태의 불규칙성을 정량적으로 평가하였다(Little, 1975) (Figure 3).

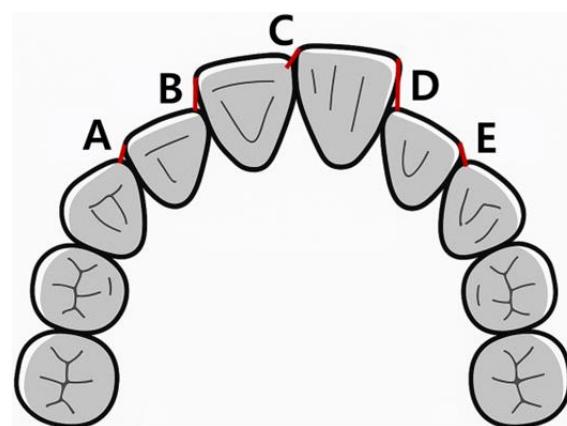


Figure 3. Little's Irregularity Index.

A-E represent the linear distances between the contact points of adjacent anterior teeth, which are measured as the red lines in the figure. The index is calculated by summing these values to represent the amount of anterior irregularity: Irregularity index = A + B + C + D + E.

2.3. 데이터 분석 방법 및 통계 처리

본 연구의 통계 분석은 IBM SPSS Statistics (Version 29.0; IBM Corp., Armonk, NY, USA)를 사용하여 수행하였다. 연구 대상자의 특성은 빈도와 백분율, 평균 및 표준편차를 이용하여 기술하였다.

치아 자동 분할 성공률과 자동 분할 소요 시간에 대한 소프트웨어 간 차이는 각각 Kruskal-Wallis H 검정을 통해 분석하였으며, 통계적으로 유의한 차이가 있는 경우 Bonferroni 보정을 적용한 사후 Wilcoxon signed-rank 검정을 실시하였다.

치아별 균원심 폭경은 각 치아에 대해 Friedman 검정을 실행하였고, 유의한 경우 동일하게 Bonferroni 보정을 적용한 Wilcoxon signed-rank 검정을 통해 사후 비교를 수행하였다. 수동 분할군(MS)과 세 가지 자동 분할 소프트웨어 간의 측정 일치도는 Bland-Altman 분석을 통해 평가하였으며, 상·하악 치아 각각에 대해 분석을 진행하였다. 또한 균원심 폭경의 반복 측정을 위해 무작위로 선택된 5쌍의 케이스에 대해 반복 측정을 실시하고, 결과를 바탕으로 Intraclass Correction Coefficient (ICC)를 산출하였다.

Bolton 비율은 전치부와 전체 치아에 대해 각 소프트웨어의 측정값을 기반으로 계산하였고, Friedman 검정으로 그룹간 간 차이를 비교하였다. 초기 배열 정확도는 각 소프트웨어의 자동 배열 기능으로 산출된 Irregularity Index 값을 바탕으로 상악과 하악으로 나누어 Friedman 검정으로 분석하였다. 모든 통계 분석에서 유의수준은 $p < 0.05$ 로 설정하였다.



3. 연구 결과

3.1. 연구 참여자의 일반적 특성

본 연구에는 제2대구치까지 맹출한 만 14세 이상의 환자 총 15명이 참여하였다. 이 중 여성은 13명(86.7%)으로 남성 2명(13.3%)보다 많았다. 대상자의 평균 연령은 23.9세($SD = 11.1$)였으며, 연령 범위는 15세에서 59세였다. 상악 치열의 평균 치열궁 길이 부조화(arch length discrepancy, ALD)는 -2.6 mm ($SD = 1.2$), 하악은 -3.3 mm ($SD = 1.1$)로, 상·하악 모두에서 공간 부족(crowding)이 관찰되었다.

3.2. 분할 성공률

치아 자동 분할의 성공률은 세 소프트웨어 간 통계적으로 유의한 차이를 보였다 ($p < 0.001$). 각 소프트웨어에서 치아 분할 결과를 정면, 측면, 교합면에서 표시하였다 (Figure 4).

DS1의 평균 성공률은 98.3%($SD = 3.3$)로 가장 높았고, DS2는 97.4%($SD = 3.2$)로 그 뒤를 이었다. 두 소프트웨어 간에는 통계적으로 유의한 차이가 나타나지 않았다 ($p = 1.000$). 반면, LS는 평균 73.6%($SD = 8.5$)의 낮은 성공률을 보여, DS1($p < 0.001$) 및 DS2($p < 0.001$)보다 유의하게 낮은 것으로 확인되었다. 최소값과 최대값은 각각 DS1(89.0–100.0), DS2(93.0–100.0), LS(61.0–89.0)로 나타났으며, LS는 표준편차가 가장 커 소프트웨어 간 성공률의 일관성 측면에도 차이를 보였다 (Table 2).



Figure 4. Segmentation results from each software.

(A) DS1 (B) DS2 (C) LS

Each figure shows segmentation results for anterior and posterior teeth. The top row presents an incisor, and the bottom row presents a molar.

From left to right, the three views are: front view (facial), side view (proximal), and top view (occlusal).

**Table 2.** Success rate of tooth segmentation by software (Unit: %)

Software	Mean ± SD	Min	Max	Post-hoc comparison
DS1	98.3 ± 3.3	89.0	100.0	DS1 vs DS2: $p = 1.000$ (ns)
DS2	97.4 ± 3.3	93.0	100.0	DS1 vs LS: $p < 0.001^*$
LS	73.6 ± 3.3	61.0	89.0	DS2 vs LS: $p < 0.001^*$

Note. Kruskal-Wallis test showed a significant difference in success rate across the three software groups, $\chi^2(2) = 32.079$, $p < 0.001$. Bonferroni-adjusted Wilcoxon signed-rank test was used for post-hoc pairwise comparisons.

* significance at the 0.05 level.

SD = standard deviation.

3.3. 분할 소요 시간

네 소프트웨어 간 상악 및 하악 모델 전체를 자동 분할하는 데 소요된 시간은 통계적으로 유의한 차이를 보였다($p < 0.001$). DS1의 평균 분할 시간은 35.9초(SD = 4.0)로 가장 짧았으며, DS2는 62.0초(SD = 20.0), LS는 73.5초(SD = 8.6)로 측정되었다. 수동 분할군인 MS는 평균 704.8초(SD = 62.0)로 가장 긴 분할 시간을 보였다.

Bonferroni 보정을 적용한 사후 분석 결과, MS는 모든 자동화 소프트웨어보다 유의하게 긴 소요 시간을 보였고($p < .001$), DS1은 LS($p < 0.001$) 및 MS($p < 0.001$)보다 유의하게 짧았다. 또한, LS는 MS보다 유의하게 짧은 시간을 보였다($p = 0.047$). DS1과 DS2간($p = 0.052$), DS2와 LS간($p = 0.468$)에는 통계적으로 유의한 차이가 없었다(Table 3).

DS와 LS는 DS1보다 평균 소요 시간이 길게 나타났으며, 특히 DS2는 표준편차가 가장 커 측정 일관성이 낮은 경향을 보였다. LS는 두 지점을 설정하는 기준 방식으로 인해 분할 시간이 상대적으로 오래 소요되는 것으로 나타났다.

**Table 3.** Comparison of tooth segmentation time by software (Unit: sec)

Software	Mean \pm SD	Min	Max	Post-hoc comparison
MS	704.8 \pm 62.0	610	890	vs All Others: $p < 0.001\dagger$
DS1	35.9 \pm 4.0	31	48	vs LS: $p < 0.001\dagger$ vs MS: $p < 0.001\dagger$
DS2	62.0 \pm 20.0	45	125	—
LS	73.5 \pm 8.6	64	95	vs DS1: $p < 0.001\dagger$ vs MS: $p = 0.047^*$

Note. Kruskal–Wallis test showed a significant difference in segmentation time across the four software groups, $\chi^2(3) = 52.757$, $p < 0.001$. Bonferroni-adjusted Wilcoxon signed-rank test was used for post-hoc pairwise comparisons.

* significance at the 0.05 level.

† significance at the 0.01 level.

SD = standard deviation; sec = seconds.

3.4. 근원심 폭경

3.4.1. 치아별 비교

네 가지 소프트웨어(MS, DS1, DS2, LS)를 통해 상악(U1-U7) 및 하악(L1-L7) 치아의 근원심 폭경을 분석하였으며 평균값, 표준편차 및 p -value는 Wilcoxon signed-rank 검정을 통해 산출되었다(Table 4, Table 5).

DS2와 LS는 대부분의 치아에서 MS 대비 근원심 폭경을 과대 측정하는 경향을 보였으며, 특히 하악 대구치(L6, L7)에서 그 차이가 크게 나타났다. 반면, DS1은 전반적으로 MS와 유사한 측정값을 보여 통계적으로 유의하지 않은 경우가 많았으며, 비교적 정확하고 일관된 측정 경향을 보였다.

전체 14개 치아 모두에서 소프트웨어 간 근원심 폭경의 유의한 차이가 나타났으며 ($p < 0.05$), 이러한 차이는 주로 소구치 및 대구치에서 두드러지게 나타났다. 전치부의 경우에는 유의한 차이가 확인되었지만, 소프트웨어 간 측정값의 차이가 상대적으로 작았다. 네 가지 소프트웨어의 근원심 폭경 측정 값은 box plot으로 시각화 하여 평균값과 범위를 함께 제시하였다(Figure 5). 특히 DS2와 LS는 대부분의 치아에서 평균값이 높게 나타나는 경향이 뚜렷하게 관찰되었다.

추가로, 각 소프트웨어의 반복 계측에 대한 재현성을 평가한 결과, 모든 프로그램에서 높은 수준의 일치도를 나타냈다. 평균 ICC는 DS1 0.938, DS2 0.951, LS 0.963으로, 세 소프트웨어 모두 높은 반복 측정 신뢰도를 보였다.

**Table 4.** Mesiodistal widths of maxillary teeth (U1–U7) by software (Unit: mm)

Tooth	Software	Mean ± SD	Min–Max	$\chi^2(df)$	p-value	Post hoc comparison
U1	MS	8.46 ± 0.47	7.92–9.38	31.40(3)	< 0.001*	MS < DS2
	DS1	8.49 ± 0.47	7.95–9.30			> DS1 = LS
	DS2	8.74 ± 0.50	8.15–9.65			
	LS	8.76 ± 0.49	8.14–9.68			
U2	MS	6.94 ± 0.50	5.61–7.63	31.48(3)	< 0.001*	MS < DS2
	DS1	6.99 ± 0.48	5.80–7.85			> DS1 < LS
	DS2	7.17 ± 0.44	6.25–8.15			
	LS	7.30 ± 0.45	6.25–8.22			
U3	MS	7.95 ± 0.37	7.46–8.69	36.22(3)	< 0.001*	MS < DS2
	DS1	7.82 ± 0.37	7.25–8.50			= LS > DS1
	DS2	8.16 ± 0.36	7.60–8.60			
	LS	8.18 ± 0.40	7.90–8.85			
U4	MS	7.45 ± 0.36	6.86–8.11	29.64(3)	< 0.001*	MS < DS2
	DS1	7.24 ± 0.40	6.80–8.10			> DS1 < LS
	DS2	7.71 ± 0.34	7.30–8.30			
	LS	7.63 ± 0.38	7.14–8.26			
U5	MS	6.83 ± 0.43	6.02–7.92	32.88(3)	< 0.001*	MS < DS2
	DS1	6.68 ± 0.42	6.10–7.25			> DS1 = LS
	DS2	7.18 ± 0.37	6.60–8.05			
	LS	7.18 ± 0.38	6.73–8.10			
U6	MS	10.45 ± 0.57	9.66–11.83	34.17(3)	< 0.001*	MS < DS2
	DS1	10.52 ± 0.49	9.85–11.95			= LS > DS1
	DS2	11.11 ± 0.60	10.05–12.40			
	LS	11.11 ± 0.52	10.28–12.25			
U7	MS	9.89 ± 0.40	9.26–10.96	25.65(3)	< 0.001*	MS < DS2
	DS1	9.58 ± 0.43	9.05–10.60			> DS1 < LS
	DS2	10.45 ± 0.46	9.60–11.60			
	LS	10.16 ± 0.42	9.57–10.94			

Note. Friedman test indicated significant differences in mesiodistal width among the four software groups for each tooth. Bonferroni-adjusted Wilcoxon signed-rank test was used for post-hoc pairwise comparisons. Post-hoc comparisons summarize statistically significant pairwise differences using symbolic notation (e.g., MS < DS2 = LS > DS1).

* significance at the 0.05 level.

Tooth abbreviations: U1, upper central incisor; U2, upper lateral incisor; U3, upper canine; U4–U5, upper premolars; U6–U7, upper molars.

Table 5. Mesiodistal widths of mandibular teeth (L1–L7) by software (Unit: mm)

Tooth	Software	Mean ± SD	Min–Max	$\chi^2(df)$	p-value	Post hoc comparison
L1	MS	5.40 ± 0.37	4.86–6.05	23.94(3)	< 0.001*	MS < DS2
	DS1	5.41 ± 0.34	4.90–6.00			= LS > DS1
	DS2	5.56 ± 0.29	5.15–6.20			
	LS	5.60 ± 0.29	5.24–6.17			
L2	MS	6.09 ± 0.40	5.21–6.75	25.18(3)	< 0.001*	MS < DS2
	DS1	6.11 ± 0.37	5.25–6.80			> DS1 = LS
	DS2	6.18 ± 0.33	5.35–6.65			
	LS	6.34 ± 0.34	5.62–6.94			
L3	MS	6.77 ± 0.32	5.90–7.27	35.09(3)	< 0.001*	MS < DS2
	DS1	6.83 ± 0.30	6.15–7.45			= LS > DS1
	DS2	7.07 ± 0.32	6.50–7.55			
	LS	7.20 ± 0.25	6.87–7.78			
L4	MS	7.35 ± 0.32	6.85–7.80	33.80(3)	< 0.001*	MS < DS2
	DS1	7.34 ± 0.37	6.95–8.10			= LS > DS1
	DS2	7.60 ± 0.34	7.20–8.30			
	LS	7.53 ± 0.35	7.20–8.24			
L5	MS	7.17 ± 0.42	6.57–7.77	28.50(3)	< 0.001*	MS < DS2
	DS1	7.25 ± 0.29	6.85–7.50			> DS1 = LS
	DS2	7.61 ± 0.30	7.10–8.00			
	LS	7.51 ± 0.36	7.02–7.89			
L6	MS	11.01 ± 0.56	9.90–11.57	29.75(3)	< 0.001*	MS < DS2
	DS1	11.19 ± 0.52	10.20–11.95			= LS > DS1
	DS2	11.52 ± 0.48	10.95–12.40			
	LS	11.45 ± 0.48	10.97–12.31			
L7	MS	10.39 ± 0.60	9.83–11.47	24.91(3)	< 0.001*	MS < DS2
	DS1	10.60 ± 0.57	9.05–11.55			> DS1 < LS
	DS2	11.03 ± 0.62	9.60–11.95			
	LS	10.66 ± 0.58	9.57–11.70			

Note. Friedman test indicated significant differences in mesiodistal width among the four software groups for each tooth. Bonferroni-adjusted Wilcoxon signed-rank test was used for post-hoc pairwise comparisons. Post-hoc comparisons summarize statistically significant pairwise differences using symbolic notation (e.g., MS < DS2 = LS > DS1).

* significance at the 0.05 level.

Tooth abbreviations: upper molars; L1, lower central incisor; L2, lower lateral incisor; L3, lower canine; L4–L5, lower premolars; L6–L7, lower molars.

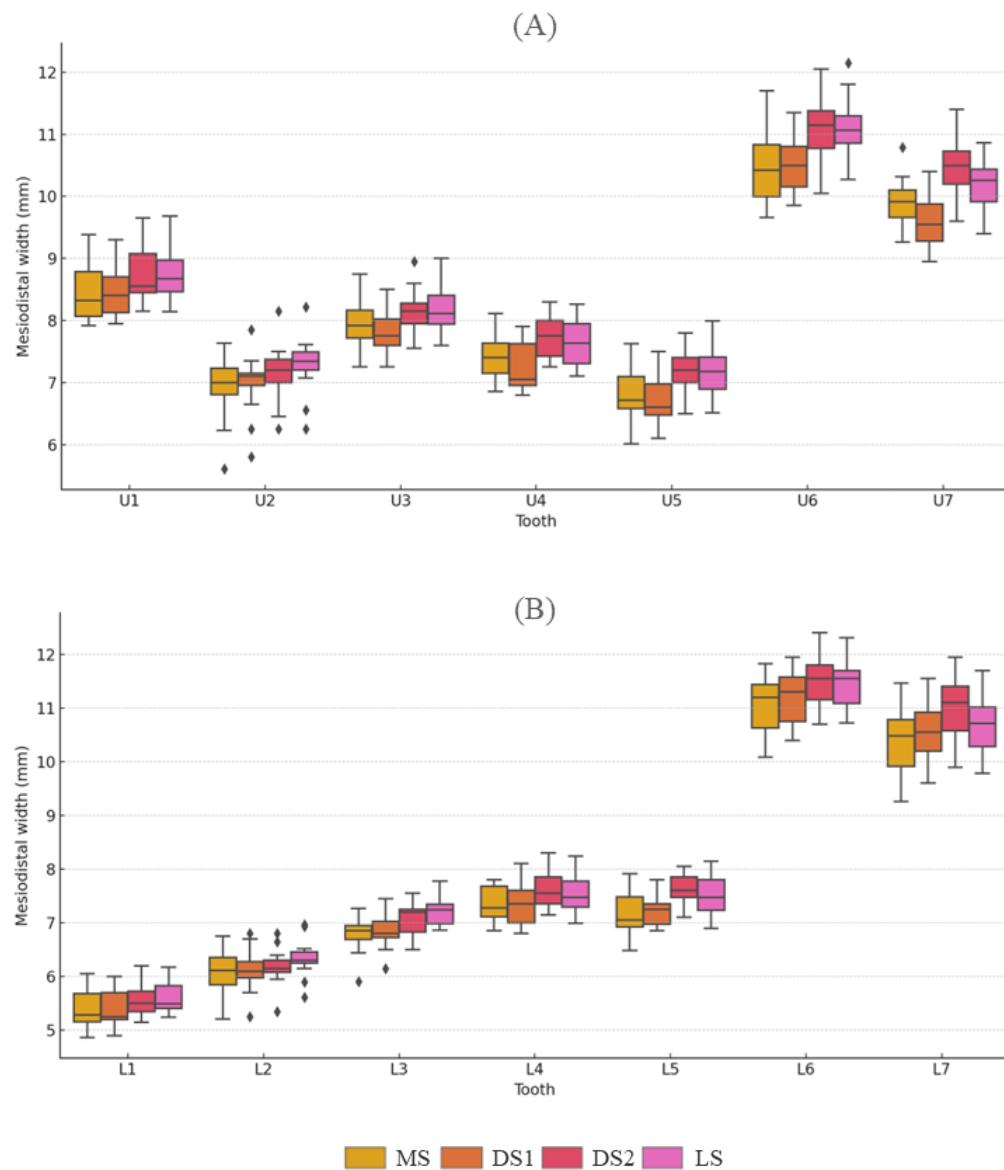


Figure 5. Comparison of mesiodistal width by tooth and software.

(A) Maxillary teeth (U1-U7) (B) Mandibular teeth (L1-L7)

Boxplots represent distribution of measurements per tooth. Outliers are shown as separate points and do not imply statistical significance.

3.4.2. 수동 분할군(MS)과 비교

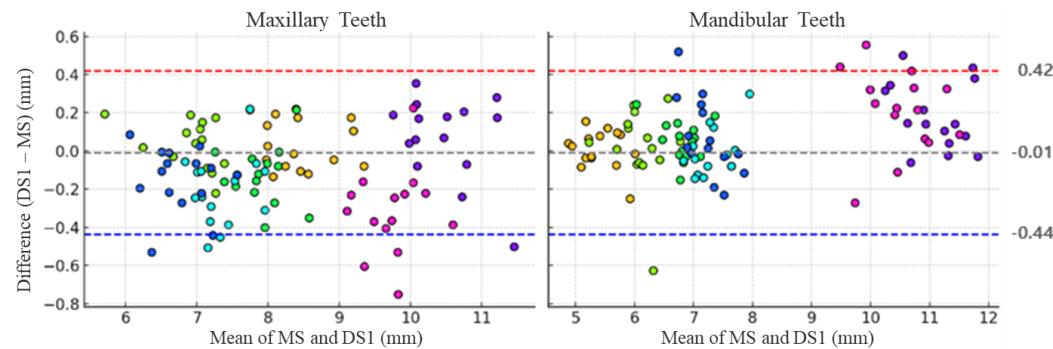
수동 분할군(MS)과 세 자동 분할 소프트웨어 간 근원심 폭경 측정값의 일치도를 평가하기 위해 Bland–Altman 분석을 실시하였다(Figure 6). 각 그래프는 MS값과의 평균 차이(mean difference) 및 95% 신뢰구간(limits of agreement, LoA)을 나타내며, 상악과 하악을 각각 분석하였다.

DS1은 평균 차이 -0.01 mm , LoA $-0.44\text{ to }+0.42\text{ mm}$ 로, MS와 거의 동일한 측정값을 보였으며 분포도 중앙선 주변에 고르게 분포하였다(Figure 6A). 드물게 LoA 범위를 초과하는 극단값이 관찰되었으나, 전체적으로 높은 일관성과 정확도를 나타냈다.

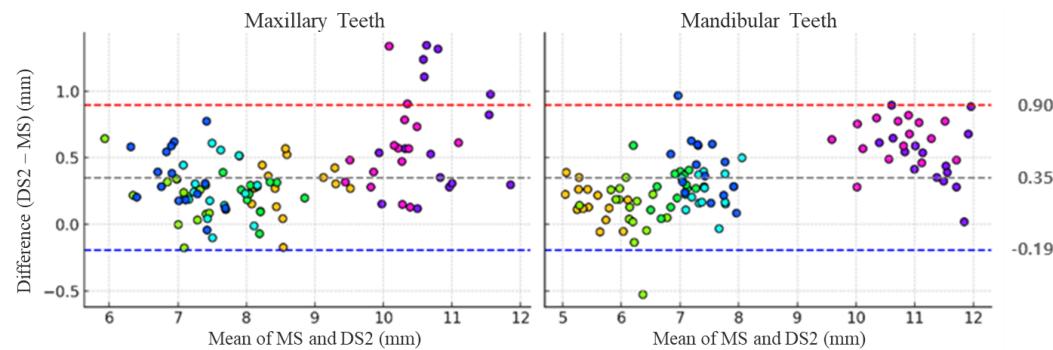
DS2는 평균차이 $+0.35\text{ mm}$, LoA $-0.19\text{ to }+0.90\text{ mm}$ 로, MS보다 전반적으로 큰 측정값을 산출하였으며, 측정값이 기준선 위쪽으로 편향되어 있고, LoA 범위를 초과한 점들이 다수 관찰되었다(Figure 6B). 또한 분포 폭이 넓어 일관성이 낮은 경향이 나타났다.

LS는 경우 평균 차이 $+0.32\text{ mm}$, LoA $-0.19\text{ to }+0.82\text{ mm}$ 로 DS2와 유사하게 MS보다 큰 측정값을 보였다. 그러나 측정값의 분포는 비교적 균등하게 나타났고, LoA범위를 초과한 값은 일부 치아에서만 관찰되었다(Figure 6C). 극단값의 빈도는 적었으나, 측정 값의 변동 범위가 커 일관성은 다소 떨어지는 경향을 보였다.

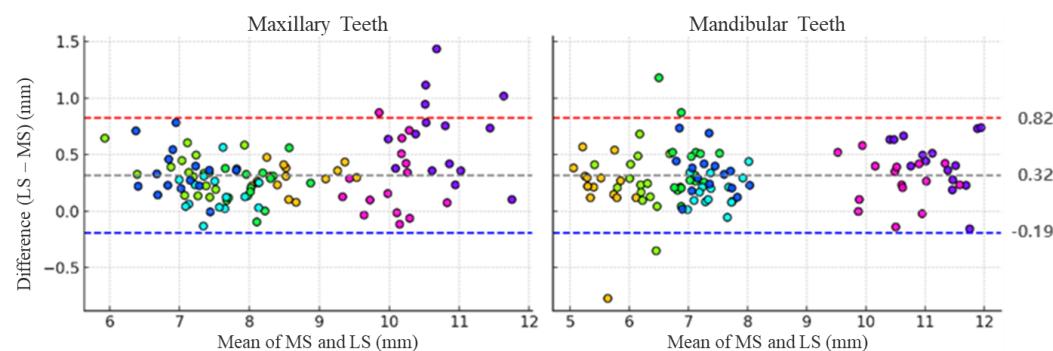
(A) MS vs DS1



(B) MS vs DS2



(C) MS vs LS



● #1 ● #2 ● #3 ● #4 ● #5 ● #6 ● #7

Figure 6. Bland-Altman plot comparing MS with each automatic tooth segmentation software.

(A) MS vs DS1: Mean difference = -0.01 mm, LoA = [-0.44, 0.42] mm

(B) MS vs DS2: Mean difference = +0.35 mm, LoA = [-0.19, 0.90] mm

(C) MS vs LS: Mean difference = +0.32 mm, LoA = [-0.19, 0.82] mm

Each plot presents the agreement between the MS and the target software in terms of mesiodistal width measurements. The X-axis represents the mean of the two methods, and the Y-axis shows their difference (automatic tooth segmentation software – MS). Colored dots indicate individual teeth (Tooth #1–#7). Dashed lines represent the mean difference (black), upper LoA (+1.96 SD, red), and lower LoA (-1.96 SD, blue).

3.5. Bolton 비율

각 소프트웨어에서 분할된 치아의 근원심 폭경 측정값을 바탕으로 전치부 치열 (Table 6) 및 전체 치열(Table 7)에 대한 Bolton 비율을 산출하였다. 전치부 비율은 MS($78.18 \pm 1.97\%$)를 기준으로 DS1 $78.73 \pm 1.63\%$, DS2 $78.17 \pm 1.97\%$, LS $79.04 \pm 1.72\%$ 로 측정되었다. Friedman 검정 결과, 네 소프트웨어 간의 전치부 비율 차이는 통계적으로 유의하지 않았다($p = 0.989$).

전체 비율 또한 MS $91.08 \pm 1.35\%$, DS1 $92.43 \pm 1.48\%$, DS2 $90.95 \pm 1.58\%$, LS는 $90.98 \pm 1.28\%$ 로 나타났으며, 이 또한 통계적으로 유의한 차이는 확인되지 않았다($p = 0.948$).

하지만 수치상으로 각 소프트웨어의 치아 계측 경향에 따른 상대적 변화는 존재하였다. 전치부에서는 LS가 하악 전치를 상대적으로 크게 측정하여 가장 높은 비율(79.04 %)을 보였으며(Table 6), 전체에서는 DS1이 상악 치아를 작게, 하악 치아를 크게 측정한 경향으로 인해 가장 높은 전체 비율(92.4%)을 보였다. 반면 DS2와 LS는 상·하악 치아를 모두 MS보다 크게 측정하여 전체 비율은 수동 분할군과 유사한 수준으로 측정되었다(Table 7).

Table 6. Anterior Bolton ratios and tooth size sums by software

Measurement	MS	DS1	DS2	LS	p-value
Anterior ratio (%)	78.18 ± 1.38	78.73 ± 1.63	78.17 ± 1.97	79.04 ± 1.72	= 0.989 (ns)
Maxillary sum (U1–U3, mm)	46.72 ± 2.17	46.61 ± 2.24	48.14 ± 2.19	48.47 ± 2.23	—
Mandibular sum (L1–L3, mm)	36.52 ± 1.82	36.69 ± 1.71	37.61 ± 1.48	38.29 ± 1.33	—

Table 7. Overall Bolton ratios and tooth size sums by software

Measurement	MS	DS1	DS2	LS	p-value
Overall ratio (%)	91.08 ± 1.35	92.43 ± 1.48	90.95 ± 1.58	90.98 ± 1.28	= 0.948 (ns)
Maxillary sum (U1–U6, mm)	96.17 ± 3.51	95.49 ± 3.84	100.15 ± 3.56	100.30 ± 3.61	—
Mandibular sum (L1–L6, mm)	87.59 ± 3.35	88.24 ± 3.29	91.06 ± 2.94	91.25 ± 3.24	—

Note. Friedman test revealed no statistically significant differences among the software groups for both anterior and overall Bolton ratios ($p > 0.05$). Statistical tests were not conducted for component sums. All values are presented as mean \pm standard deviation.

3.6. 초기 배열 정확도

초기 배열 정확도는 각 소프트웨어의 자동 배열 기능(arch line generation)을 이용하여 상·하악 전치부 치아의 배열을 수행한 후, Little's Irregularity Index를 통해 평가하였다. 해당 지표는 전치부 치아 간 접촉점 거리의 합을 기준으로 하며, 값이 낮을수록 배열이 잘 되어 있음을 의미한다.

본 연구에서는 MeshLab(Visual Computing Lab, ISTI-CNR, Italy) 프로그램을 이용하여 3차원 모델 상에서 각 치아의 접촉점 간 선형 거리(linear distance)를 측정하였다. 자동 배열은 분할 과정 이후 산출된 mesial-distal point를 기반으로 하며, 기준 치아의 Landmark를 연결한 spline 곡선을 따라 각 치아의 중심축(MD axis)을 정렬하는 방식으로 수행된다. 수동 분할군(MS)은 10년 이상의 경력을 가진 치과기공사가 수작업으로 배열한 모델을 기반으로 설정하였다.

분석 결과, MS는 상악 0.68 ± 1.30 mm로 가장 낮은 Irregularity Index를 기록하였다. 자동화 소프트웨어 중에서는 DS1이 상악 3.06 ± 2.66 mm, 하악 2.85 ± 1.65 mm로 가장 낮은 값을 보여 상대적으로 우수한 자동 배열 성능을 보였다. 반면, DS2는 상악에서 6.02 ± 2.46 mm로 가장 높은 불규칙성을 보였으며, LS는 하악에서 4.81 ± 1.99 mm로 가장 큰 편차를 나타냈다. Friedman 검정 결과, 상악($p = 0.003$)과 하악($p < 0.001$) 모두에서 통계적으로 유의한 차이가 확인되었으며, Bonferroni 보정 사후 분석 결과, 상악에서는 $MS < DS2 = LS > DS1$, 하악에서는 $MS < DS1 < DS2 < LS$ 의 순서로 정렬 정확도의 유의한 차이가 나타났다(Table 8).

이러한 결과는 초기 배열 정확도가 각 소프트웨어의 자동 배열 알고리즘 성능 뿐 아니라, 치아 분할 단계의 정확도에도 영향을 받는다는 점을 나타낸다. 특히 자동으로 산출되는 기준점의 정확도가 낮을 경우, 아치라인 생성과 배열에도 오류가 발생할 수 있다. DS1은 이러한 점에서 비교적 안정적인 성능을 보였으며, 반면 DS2와 LS는 불규칙성과 편차가 크게 나타나 알고리즘 개선이 필요할 수 있음을 보여준다.

**Table 8.** Comparison of Irregularity Index by arch and software (Unit: mm)

Software	Mx (Mean \pm SD)	Mn (Mean \pm SD)	χ^2 (Mx)	p (Mx)	χ^2 (Mn)	p (Mn)
MS	0.68 \pm 1.30	0.52 \pm 1.38	14.03	= 0.003*	18.16	< 0.001*
DS1	3.06 \pm 2.66	2.85 \pm 1.65				
DS2	6.02 \pm 2.46	4.08 \pm 1.66				
LS	4.63 \pm 3.46	4.81 \pm 1.99				

Note. The Irregularity Index represents the degree of anterior alignment irregularity for each arch. All values are presented as mean \pm standard deviation. The Friedman test was used to evaluate differences among software programs. Bonferroni-adjusted Wilcoxon signed-rank test was used for post-hoc pairwise comparisons where significant differences were found.

* significance at the 0.05 level.

Abbreviations: Mx = Maxillary teeth (U1–U7); Mn = Mandibular teeth (L1–L7).

4. 고찰

본 연구는 세 가지 상업용 CAD 소프트웨어(DS1, DS2, LS)를 활용하여 디지털 치아 모델 상의 자동 치아 분할과 배열 결과를 정량적으로 평가하고, 수동 분할군(MS)과 비교하여 자동화 소프트웨어의 치아 분할 방식과 배열 기능의 정확도 및 효율성을 평가하였다.

분할 성공률은 DS1 과 DS2 가 각각 98.4%, 97.4%로 높았던 반면, LS 는 73.6%로 낮아 사용자 개입이 필요한 수준으로 판단되었다. 분할 소요 시간은 DS1 이 평균 35.9 초로 가장 짧았으며, DS2(62.0 초), LS(73.5 초) 순으로 길게 나타났다. 이러한 차이는 소프트웨어별 분할 방식의 차이에 기인한 것으로 보이며, DS1 은 경계 기반(edge-based) 분할 방식을 사용해 비교적 빠르고 단순한 경계 설정이 가능했던 반면, DS2 와 LS 는 곡률 기반(curvature flow) 알고리즘을 통해 치아 외형을 정밀하게 재현하므로 처리 시간이 길어지는 경향을 보였다. 특히 LS 는 각 치아에 대해 두 개의 기준점을 수동으로 지정해야 하므로 전체 분할 과정에서 시간이 가장 많이 소요되었다.

근원심 폭경 비교 결과, DS2 와 LS 는 대부분의 치아에서 MS 보다 크게 측정되었으며, 특히 하악 대구치(L6, L7)에서 그 차이가 두드러졌다. 반면, DS1 은 전반적으로 MS 와 유사한 값을 보여 상대적으로 정확하고 일관된 경향을 나타냈다. 이러한 경향은 주로 소구치와 대구치에서 뚜렷하게 나타난 반면, 전치부에서는 소프트웨어간 측정값의 차이가 상대적으로 작았다. 유사하게, 최근 연구에서도 AI 기반 자동 치아 분할 방식이 수동 방식보다 치아 크기를 더 크게 측정하는 경향이 보고된 바 있다(Yu et al., 2023). 이러한 차이는 소프트웨어별 계측 방식의 구조적 차이에 기인한 것으로 판단된다. LS 는 치아 분할 이전에 사용자가 지정한 landmark 를 기반으로 생성된 mesial-distal point 를 기반으로 계측을 수행하기 때문에, 인접면에서 최대 풍옹부(contact point)를 정확하게 지정하기 어려운 경우 과소 또는 과대 측정될 수 있다. 반면 DS1 과 DS2 는 치아 분할 후 정규화(normalization)된 모델 상에서 근원심 폭경을 산출하므로, 정렬 기준(예: 치축

방향, bounding box 기준)에 따라 측정값의 편차가 발생할 수 있다. 이러한 계측 방식의 차이는 동일 치아를 대상으로 하더라도 측정값 간 오차를 유발할 수 있으며, 이는 디지털 셋업 모델의 계측 방식과 관련하여 보고된 바 있다(Im et al., 2022; Xu et al., 2019). 또한, 디지털 셋업 모델은 석고 모형과 달리 표면 기반 계측 방식으로 인해 경계 식별의 한계가 존재하여 미세한 계측의 정확도가 떨어질 수 있으며, 작업 효율성은 높지만 계측 정확도에서는 석고 모형보다 낮은 신뢰도를 보일 수 있다는 점이 지적되었으며, 두 방식 간 비교 검토의 중요성을 강조하였다(Fleming et al., 2011; Hildebrand et al., 2008; Okunami et al., 2007).

이러한 점을 고려할 때, 디지털 모델 분석 결과를 임상에 적용하기 위해서는 이러한 한계점을 충분히 인지하고 해석할 필요가 있다. 특히, 소프트웨어별 분할 알고리즘, 석고 모형과 디지털 모델 간 재현 방식의 차이, 근원심 폭경 산출 방식, 그리고 치아의 해부학적 형태 차이는 특정 치아 부위에서 측정값 편차를 유발할 수 있는 요인으로 작용하기 때문이다.

Bolton 비율 분석 결과, 소프트웨어 간 절대값 차이는 존재했으나 통계적으로 유의한 차이는 없었다($p > 0.05$). 전치부에서는 LS 가 상악 전치를 상대적으로 크게 측정하여 가장 높은 전치부 비율(79.04%)을 나타냈으며, 전체 치열에서는 DS1 이 상악 치아를 작게, 하악 치아를 크게 측정함으로써 가장 높은 전체 비율(92.43%)을 보였다. 반면 DS2 와 LS 는 상·하악 치아 모두를 크게 측정했지만, 비율은 MS 와 유사하게 유지되었다. 이러한 결과는 단순한 비율만으로는 임상 적용 가능성을 평가하기 어렵고, 치아 크기의 절대값과 상·하악 간 상대적 관계를 함께 고려해야 함을 시사한다. 특히 치아가 과소 또는 과대 계측된 경우, 비율이 수동군(MS)과 유사하더라도 치아 크기 분석이나 실제 배열에서는 정확도가 떨어질 수 있다. 이와 관련하여 (Cho et al., 2010)은 치아 크기와 비율의 정밀한 평가가 치료 예측의 정확도를 높이는 데 중요하다고 강조하였으며, 본 연구 결과 역시 비율이 수동군에 근접하더라도 치아 크기의 절대값 차이에 따라 임상 적용에 영향을 미칠 수 있음을 보여준다.



초기 배열 정확도에서는 DS1 이 상악 3.06 mm, 하악 2.85 mm 로 가장 낮은 값을 나타내 우수한 배열 성능을 보였고, DS2는 상악에서 6.02 mm, LS는 하악에서 4.81 mm 로 각각 가장 큰 오차를 나타냈다. 이는 배열 정확도가 단순히 자동 배열 알고리즘의 성능뿐 아니라, 분할 단계에서 산출된 기준점(mesial-distal point)의 정확도에도 영향을 받는다는 점을 보여준다.

본 연구는 다음과 같은 한계를 가진다. 첫째, 분석에 사용된 디지털 모델은 결손치, 중증 총생 등 특이 조건이 없는 케이스만 포함되어, 다양한 임상 상황에서 결과가 다를 수 있다. 둘째, 각 소프트웨어의 세부 알고리즘이 비공개되어 있어 기술적 차이를 명확히 분석하기 어렵다. 셋째, 소프트웨어 버전 및 업데이트에 따라 결과가 달라질 수 있어 일관성 확보에 제약이 있을 수 있다.

이러한 한계에도 불구하고, 본 연구는 실제 임상 데이터를 기반으로 자동 치아 분할 및 배열 기능에 대한 정량적 비교를 시도하였으며, 성공률, 분할 시간, 치아 크기 측정값, Bolton 비율, 배열 정확도 등을 포함한 분석을 통해 각 소프트웨어의 특성을 파악할 수 있었다.

자동 치아 분할에서는 일부 소프트웨어에서 낮은 성공률이나 작업 간 편차가 관찰되었으며, 자동 배열 결과 또한 인접치 간 충돌이나 교합 간섭 등 임상 조건을 충분히 반영하지 못하는 한계가 존재하였다. 따라서 자동 치아 배열의 결과를 임상적으로 활용하기 위해서는 교정의사의 임상적 판단과 수동 조정이 병행되어야 하며, 분할 정밀도 향상도 함께 고려되어야 한다.



5. 결론

본 연구는 DS1(Onyxceph), DS2(OneOrtho), LS(Autolign) 세 가지 CAD 소프트웨어의 자동 치아 분할 및 배열 성능을 비교·평가하였다. 주요 결과는 다음과 같다.

1. 분할 성공률은 DS1(98.3%)과 DS2(97.4%)가 높았으며, LS 는 73.6%로 낮은 결과를 나타냈다.
2. 분할 소요 시간은 DS1 이 35.9 초로 가장 짧았고, DS2(62.0 초), LS(73.5 초) 순으로 나타났다.
3. 근원심 폭경 측정에서는 DS1 이 MS 과 가장 유사하게 측정되었으며, DS2 와 LS 는 대부분의 치아에서 MS 보다 크게 측정되었고, 하악 대구치(L6, L7)에서 차이가 두드러졌다.
4. Bolton 비율은 네 소프트웨어간 유의한 차이는 없었으나, LS 는 전치부 비율(79.06%), DS1 은 전체 비율(92.43%)이 가장 높았다.
5. 초기 배열 정확도는 DS1 이 상·하악 모두에서 가장 낮은 Irregularity Index 를 나타냈으며, DS2(상악)와 LS(하악)는 상대적으로 높은 값을 나타냈다.

본 연구를 통해 소프트웨어의 자동 분할 및 배열 성능에는 차이가 있음을 확인하였으며, 임상 적용 시 각 소프트웨어별 특성을 이해하고 선택적으로 활용하는 것이 중요함을 보여주었다. DS1은 분할 성공률, 처리 속도, 분할 및 배열 정확도에서 상대적으로 우수한 성능을 보인 반면, LS는 일부 치아에서 계측 오차와 작업 시간의 변동 폭이 크게 나타났다. 이러한 결과는 자동화 소프트웨어를 임상에 활용할 때 교정의사의 임상적 판단이 병행되어야 하며, 향후 알고리즘의 정밀도 개선이 필요함을 시사한다. 본 연구는 디지털 셋업과 자동 치아 분할 방식을 교정 진단과 치료 계획에 적용할 때, 임상 판단과 결정에 도움이 되는 근거 자료로 활용될 수 있다.

참고 문헌

- Cho, M. Y., Choi, J. H., Lee, S. P., & Baek, S. H. (2010). Three-dimensional analysis of the tooth movement and arch dimension changes in Class I malocclusions treated with first premolar extractions: a guideline for virtual treatment planning. *Am J Orthod Dentofacial Orthop*, 138(6), 747–757.
- Fleming, P. S., Marinho, V., & Johal, A. (2011). Orthodontic measurements on digital study models compared with plaster models: a systematic review. *Orthod Craniofac Res*, 14(1), 1–16.
- Gonzalez Guzman, J. F., & Teramoto Ohara, A. (2019). Evaluation of three-dimensional printed virtual setups. *Am J Orthod Dentofacial Orthop*, 155(2), 288–295.
- Hildebrand, J. C., Palomo, J. M., Palomo, L., Sivik, M., & Hans, M. (2008). Evaluation of a software program for applying the American Board of Orthodontics objective grading system to digital casts. *Am J Orthod Dentofacial Orthop*, 133(2), 283–289.
- Hou, D., Capote, R., Bayirli, B., Chan, D. C. N., & Huang, G. (2020). The effect of digital diagnostic setups on orthodontic treatment planning. *Am J Orthod Dentofacial Orthop*, 157(4), 542–549.
- Im, J., Cha, J. Y., Lee, K. J., Yu, H. S., & Hwang, C. J. (2014). Comparison of virtual and manual tooth setups with digital and plaster models in extraction cases. *Am J Orthod Dentofacial Orthop*, 145(4), 434–442.
- Im, J., Kim, J. Y., Yu, H. S., Lee, K. J., Choi, S. H., Kim, J. H., Ahn, H. K., & Cha, J. Y. (2022). Accuracy and efficiency of automatic tooth segmentation in digital dental models using deep learning. *Sci Rep*, 12(1), 9429.
- Kesling, H. D. (1946). Coordinating the predetermined pattern and tooth positioner with conventional treatment. *Am J Orthod Oral Surg*, 32, 285–293.
- Kok, H., Acilar, A. M., & Izgi, M. S. (2019). Usage and comparison of artificial intelligence algorithms for determination of growth and development by cervical vertebrae stages in orthodontics. *Prog Orthod*, 20(1), 41.

- Lee, J. H., Yu, H. J., Kim, M. J., Kim, J. W., & Choi, J. (2020). Automated cephalometric landmark detection with confidence regions using Bayesian convolutional neural networks. *BMC Oral Health*, 20(1), 270.
- Little, R. M. (1975). The irregularity index: a quantitative score of mandibular anterior alignment. *Am J Orthod*, 68(5), 554–563.
- Macchi, A., Carrafiello, G., Cacciafesta, V., & Norcini, A. (2006). Three-dimensional digital modeling and setup. *Am J Orthod Dentofacial Orthop*, 129(5), 605–610.
- Okunami, T. R., Kusnoto, B., BeGole, E., Evans, C. A., Sadowsky, C., & Fadavi, S. (2007). Assessing the American Board of Orthodontics objective grading system: digital vs plaster dental casts. *Am J Orthod Dentofacial Orthop*, 131(1), 51–56.
- Suhail, Y., Upadhyay, M., Chhibber, A., & Kshitiz. (2020). Machine Learning for the Diagnosis of Orthodontic Extractions: A Computational Analysis Using Ensemble Learning. *Bioengineering (Basel)*, 7(2).
- Tian, S., Dai, N., Wang, C., and Lin, H. (2019). Automatic classification and segmentation of teeth on 3D dental model using hierarchical deep learning networks. *IEEE Access*, 7, 84817–84828.
- Xie, X., Wang, L., & Wang, A. (2010). Artificial neural network modeling for deciding if extractions are necessary prior to orthodontic treatment. *Angle Orthod*, 80(2), 262–266.
- Xu, X., Liu, C., & Zheng, Y. (2019). 3D Tooth Segmentation and Labeling Using Deep Convolutional Neural Networks. *IEEE Trans Vis Comput Graph*, 25(7), 2336–2348.
- Yacout, Y. M., Eid, F. Y., Tageldin, M. A., & Kassem, H. E. (2024). Evaluation of the accuracy of automated tooth segmentation of intraoral scans using artificial intelligence-based software packages. *Am J Orthod Dentofacial Orthop*, 166(3), 282–291 e281.
- Yoon, J. H., Yu, H. S., Choi, Y., Choi, T. H., Choi, S. H., & Cha, J. Y. (2018). Model Analysis of Digital Models in Moderate to Severe Crowding: In Vivo Validation



and Clinical Application. *Biomed Res Int*, 2018, 8414605.

Yu, J. H., Kim, J. H., Liu, J., Mangal, U., Ahn, H. K., & Cha, J. Y. (2023). Reliability and time-based efficiency of artificial intelligence-based automatic digital model analysis system. *Eur J Orthod*, 45(6), 712–721.

ABSTRACT

The Accuracy and Efficiency of Digital Dental Model Analysis According to Automatic Tooth Segmentation Methods

This study evaluated the accuracy and efficiency of digital setups for orthodontic diagnosis and treatment planning by comparing the digital tooth model analysis and automatic tooth alignment accuracy according to the automatic tooth segmentation methods of three commercial CAD software programs (Onyxceph [DS1], OneOrtho [DS2], and Autolign [LS]). A total of 15 pairs of maxillary and mandibular digital tooth models from clear aligner patients were analyzed to compare and evaluate segmentation success rate, segmentation time, mesiodistal width, Bolton ratio, and initial alignment accuracy with a manual segmentation group (MS).

The analysis showed that DS1 demonstrated a high segmentation success rate (98.3%) and the shortest segmentation time (35.9 seconds), indicating high efficiency and accuracy, and yielded mesiodistal width measurements comparable to the manual reference group. DS2 also showed a high segmentation success rate but exhibited overestimation in measurements for some teeth and large variability in processing time. LS presented the lowest segmentation success rate and the longest processing time among the three software programs, with a tendency to overestimate measurements for certain teeth. While there were no statistically significant differences in the Bolton ratios among the software programs, DS1 demonstrated the lowest Irregularity Index in the initial alignment accuracy analysis, indicating relatively higher alignment accuracy.

These results indicate that automatic tooth segmentation functions directly impact the diagnostic and alignment processes within digital setups, underscoring the importance of understanding and utilizing the algorithmic structures and characteristics of each software in clinical applications. This study confirmed the clinical applicability of automatic tooth



segmentation and alignment functions and may contribute to expanding the clinical reliability and utility of automatic tooth segmentation technologies through future studies involving various clinical conditions and complex cases.

Key words: automatic tooth segmentation, digital tooth model analysis, digital setup, CAD software, segmentation success rate, segmentation time, mesiodistal width, Bolton ratio, alignment accuracy