Automated clustering of the three-dimensional mandibular canal course using unsupervised machine learning method

Young Hyun Kim

The Graduate School

Yonsei University

Department of Dentistry

Automated clustering of the three-dimensional mandibular canal course using unsupervised machine learning method

Directed by Professor Sang-Sun Han, D.D.S., Ph.D.

The Doctoral Dissertation submitted to the Department of Dentistry, and the Graduate School of Yonsei University in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Dental Science

Young Hyun Kim

February 2021

This certifies that the Doctoral Dissertation of
Young Hyun Kim is approved.

Thesis Supervisor: **Prof. Sang-Sun Han**

Thesis Committee Member: **Prof. Kug-Jin Jeon**

Thesis Committee Member: **Prof. Chena Lee**

Thesis Committee Member: **Prof. Yoon Joo Choi**

Thesis Committee Member: **Prof. Hoi-In Jung**

The Graduate School

Yonsei University

February 2021

감사의 글

뒤늦게 시작된 대학원 과정은 ‘하나님께서 예비하셨다.’라는 말이 아니면 설명할 수 없는 선물과 같은 시간이었습니다. 그저 하루하루 최선을 다 해 살아내는 것도 벅찬 저에게, 조건 없는 사랑과 은혜를 부어 주시는 살아 계신 하나님께 모든 영광 돌려 드립니다.

부족한 저를 한 명의 연구자가 되기까지 지도해주시고 오롯이 연구에 집중할 수 있도록 크고 작은 상황마다 배려해주신 한상선 지도교수님께 마음 깊이 감사드립니다. 교수님의 제자가 된 것이 제 인생의 큰 행운이었습니다. 앞서가신 길을 따라가면서 교수님께 힘이 될 수 있는 제자가 되겠습니다. 놀라운 세심함으로 연구의 완성도를 높여 주신 전국진 교수님, 제 멘탈이 흔들릴 때 마다 따듯한 조언과 함께 연구 방향을 고민해 주시는 이채나 교수님, 판독 업무로 바쁘신 와중에도 연구와 육아까지 해내시는 슈퍼워킹맘 최윤주 교수님, 통계 분석이 막힐 때 예의없이 불쑥 드리는 질의 메일과 카톡에도 명쾌한 답변과 응원의 말씀을 잊지 않으시던 정회인 교수님. 바쁘신 와중에 제 학위논문 심사를 맡아 주셔서 감사드립니다.

올해 꼭 졸업하게 도와주겠다며 부족한 시간을 쪼개어 연구를 도와준 아리쌤, 어리지만 배울 점이 많은 연구 파트너이자 랩실의 귀염둥이 은규쌤, 입학한지 한 학기도 되지 않아 완벽한 적응력을 보여주고 있는 한승쌤, 급박한 요청에 늘 웃으며 최선을 다해주는 혜언쌤, 연구만큼 복잡한 각종 행정 업무를 완벽히 처리해주는 은주쌤 모두 고마워요!

더불어, 회사 잘 다니던 딸의 갑작스러운 대학원 입학부터 졸업을 앞둔 지금까지도 든든한 지원과 변함없는 신뢰를 보여주시는 아빠, 엄마! 두 분의 살아가시는 모습 자체로 제 인생의 본이 되어 주셔서 감사합니다. 당신들의 딸로 태어나서 행복해요. 사랑하고 존경합니다. 생신도 잘 못 챙기는 실수 많은 맏며느리에게 괜찮다, 고생한다 응원해 주시고, 1년의 시간 동안 황혼 육아로 고생하시는 아버님, 어머님! 죄송하고 감사합니다. 저만큼 좋은 시댁 만난 며느리는 없을 거예요! 날 위해 매일 기도하며 영적 부족함을 채워주는 형부와 언니, 고등학교 선생님으로 내 예상보다 훨씬 더 멋지게 자라준 남동생 영민이, 다정다감하고 따듯한 마음을 가진 홍조 도련님, 부족한 조카를 항상 자랑스러워 해주시는 이모. 모두의 응원 덕분에 길고 긴 학위 과정을 마무리할 수 있었던 것 같습니다. 특별히, 내년이면 20년 지기가 되는 사랑하는 민정, 희선, 윤진아! 힘들고 어려운 순간마다 너희가 주는 긍정 에너지가 아니었다면 내가 견뎌내지 못했을 거야.

마지막으로, 내 평생의 연구 파트너이자 조언자인 사랑하는 홍선 오빠. 성장을 위해 끊임없이 공부하고 노력하는 모습에 많은 도전을 받아요. 우리 지금처럼 서로에게 좋은 영향을 주며 살아가요! 그리고 로미야! 한창 엄마가 필요할 시기에 떨어져 지내면서도 잘 자라주어 고마워. 엄마의 엄마가 그랬던 것처럼, 엄마도 네가 닮아가고 싶은 사람이 될 수 있도록 주어진 환경과 자리에서 최선을 다할게. 사랑해, 아주 많이.

2020년 12월 저자 씀

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract

Automated clustering of the three-dimensional mandibular canal course using unsupervised machine learning method

Young Hyun Kim

Department of Dentistry
The Graduate School, Yonsei University

(Directed by Professor Sang-Sun Han, D.D.S., Ph.D.)

Purpose: The aims of this study were to apply cluster analysis, unsupervised machine learning method, for three-dimensional (3D) course classification of mandibular canal (MC), and to visualize the standard MC courses presented by cluster analysis in the Korean population.

Methods: A total of 429 cone beam computed tomography images acquired at Yonsei University Dental Hospital were used. Four sites were selected for the measurement of the MC course and two vertical and two horizontal parameters were measured per site. Cluster analysis was carried out as follows: parameter selection, parameter normalization, cluster tendency evaluation, optimal number of clusters determination, and k-means cluster analysis.

Results: Three types of the 3D MC course were derived as cluster 1, 2 and 3 by cluster

analysis, and statistically significant mean differences were shown among clusters. Cluster 1 showed a smooth line running towards the lingual side in the axial view and a steep slope in the sagittal view. Cluster 2 ran in an almost straight line closest to the lingual and inferior border of mandible. Cluster 3 showed the pathway with a bent buccally in the axial view and an increasing slope in the sagittal view in the posterior area. Cluster 1 had the least distribution (26.0%). In this cluster, females accounted for 59.2% and males, 40.8%. On the other hand, cluster 2 showed the widest distribution of driving courses (42.1%), and males were more widely distributed (57.1%) than the female group (42.9%). Cluster 3 comprised similar ratio of male and female cases and accounted for 31.9% of the total distribution. For all three clusters, distributions of the right and left sides did not show a statistically significant difference.

Conclusions: The 3D MC courses were automatically classified as three types through cluster analysis. Cluster analysis enables the unbiased classification of the anatomical structures by reducing observer variability and can present representative standard information for each classified group.

Keywords: Cluster Analysis, Cone-Beam Computed Tomography, Inferior Alveolar Nerve, Mandibular Canal, Unsupervised Machine Learning

Automated clustering of the three-dimensional mandibular canal course using unsupervised machine learning method

Young Hyun Kim

Department of Dentistry
The Graduate School, Yonsei University

(Directed by Professor Sang-Sun Han, D.D.S., Ph.D.)

I. INTRODUCTION

The inferior alveolar nerve (IAN) is the largest branch of the mandibular nerve which is divided from trigeminal nerve.[1] As the bilateral intraosseous structure, the mandibular canal (MC) contains the neurovascular bundle, which is made up of the IAN, inferior alveolar artery and inferior alveolar vein.[2] In general, the MC refers to the course that starts from the mandibular foramen, located in the middle of the mandibular ramus, and ends at the buccal side of the mental foramen located below the 2nd premolar.[3,4]

Since the inferior alveolar neurovascular bundle plays an important role in supplying the sensation and blood to the mandible, it is recommended to assess the MC location accurately before conducting dental operations.[5-7] However, the IAN is still reported as the most commonly damaged structure during various dental procedures, such as third molar extraction, implant surgery, local anesthesia injection, or endodontic treatment.[8-10]

Temporary or permanent complications caused by a damaged IAN can significantly reduce the patient's quality of life both physically and psychologically.[11,12] Therefore, better understanding of the anatomical features of the MC is one of the key factors for a successful dental treatment.

So far, classification of the MC course using dry skulls or radiographs was tried in numerous studies with notable limitations.[3,13-17] MC course analysis using only a limited number of skulls may not reflect sufficient anatomical variations of the MC courses, and the absence of relevant information such as sex, age, and disease history makes it difficult to obtain accurate results.[17,18] With the development of medical imaging equipment, researchers began to classify the MC courses through visual assessment of panoramic or cone-beam computed tomography (CBCT) radiographs using more cases than used in cadaver dissection studies.[14,15,17] Accordingly, two to four types of the MC courses have been reported from the previous studies. This discrepancy in the reported number of the MC courses may be due to the limitation of the visual assessment, where differences in interpretation between researchers are inevitable when categorizing anatomical structures.[19] Moreover, in most studies, horizontal position[20,21] and vertical position[13-15,17] of the MC course were analyzed respectively. As a result, none of these studies classified the MC courses considering their location in three-dimensional (3D) perspective by simultaneously analyzing horizontal and vertical positions. Hence, to overcome this shortcoming, it is necessary to introduce a new approach that enables objective classification of the 3D anatomical structure without observer bias.

Cluster analysis, one of unsupervised machine learning techniques for data mining, is a useful method to categorize large data into clusters based on the similarity or distance (dissimilarity) among individual objects.[22] This statistical approach is widely applied in various fields such as marketing, clothing design, healthcare, and even deep learning.[23-26] Two types of clustering method, hierarchical clustering[27] and k-means (non-hierarchical) cluster analysis[28], have been mainly used in medical research.[23,29,30] The clustering process continues until the distance between objects assigned to one cluster is minimized and the distance between each separated cluster is maximized, in order to ensure that the classified cluster will ultimately have different characteristics.

Therefore, the aims of this study were to classify 3D MC course using cluster analysis, unsupervised machine learning method, and to visualize the standard course of each classified MC course type in Korean population.

Ⅱ. MATERIALS AND METHODS

1. Study approval

This study was approved by the Institutional Review Board (IRB) of Yonsei Dental Hospital (No.2-2019-0073) and performed in accordance with relevant guidelines and ethical regulations. The IRB of Yonsei Dental Hospital granted a waiver of informed consent form due to the retrospective nature of this study.

2. Data preparation

From July 2019 to July 2020, a total of 429 CBCT images were collected from the picture archiving and communication system of Yonsei University Dental Hospital. All images were acquired by RAYSCAN Alpha Plus device (Ray Co. Ltd, Hwaseong-si, Korea) using the following acquisition settings; field of view, 16 x 10 cm; voxel size, 0.23 mm; 90 kVp; 12 mA; and exposure time, 14s. The images were exported in anonymous Digital Image Communication in Medicine (DICOM) format.

The exclusion criteria are as follows: i) subjects who were under 18 years old, ii) subjects who had pathologic lesions, or underwent orthognathic surgery in the mandible, iii) completely edentulous subjects, and iv) unclear CBCT images with blurring or artifacts.

3. Mandibular canal course measurement

The 3D MC course was measured to be expressed as parameters through the following process:

i) DICOM images were imported to 3D software (Ondemand3D; Cybermed Inc., Seoul, Korea) and reoriented. The mandibular occlusal plane (MnOP) was used as a reference plane for reorientation. The MnOP was defined in two directions; horizontally it connected the edge of the central incisor and the mesiobuccal cusp tip of the first molar, and transversely it connected the mesiobuccal cusp of both first molars. The horizontal and transverse MnOPs were aligned in parallel with the floor of the images in sagittal and coronal views, respectively (Figure 1A and B).

ii) The cross-sectional images were reconstructed automatically at the axial level where root furcation of the first molar existed (Figure 1C).

iii) Four sites (S1, S2, S3 and S4) at the interval of 10 mm were selected from the center of the mental foramen (S0) (Figure 2A).

iv) In each site, four parameters showing the location of the MC were measured in millimeters (Figure 2B). Two vertical parameters, upper height (UH) and lower height (LH), and two horizontal parameters, buccal width (BW) and lingual width (LW), were measured. The parameters were defined in Table 1.

To evaluate the reliability of the measured parameters, 30 CBCT images were randomly selected and measured by two observers. The inter-class correlation coefficients (ICCs) were conducted with 95% confidence interval (CI) (Table 2).

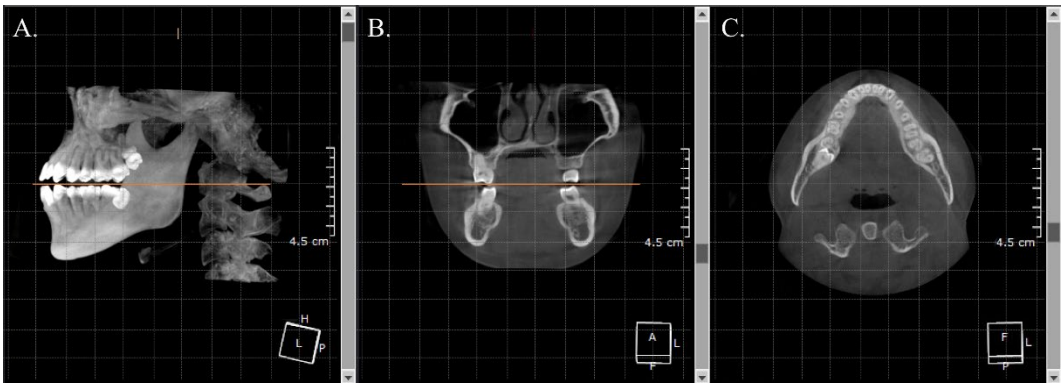

Figure 1. CBCT reorientation using image analysis software

(A) Reoriented sagittal view with the horizontal occlusal plane (orange line)

(B) Reoriented coronal view with the transverse occlusal plane (orange line)

(C) Axial view used as a reference for cross-sectional image reconstruction

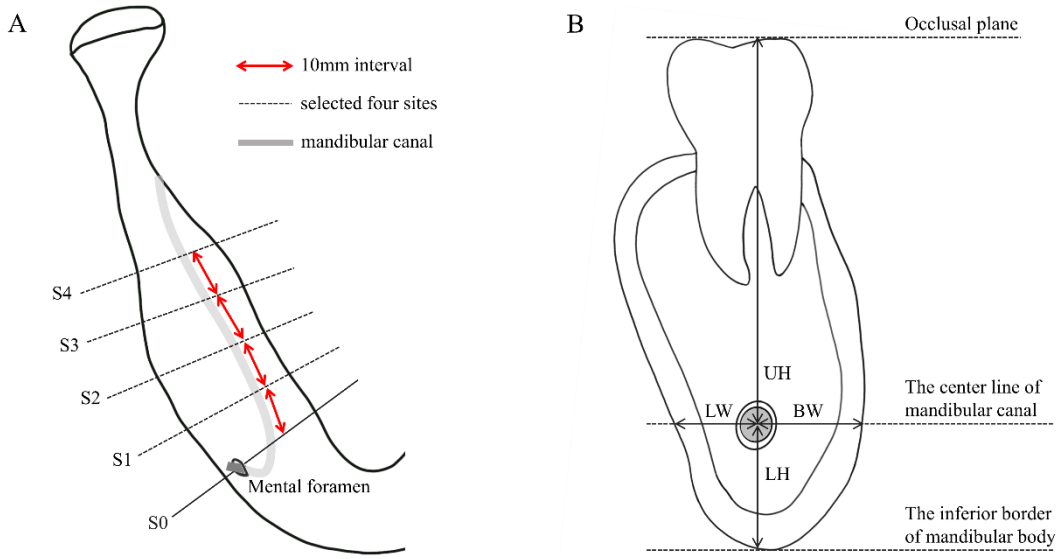

Figure 2. Two schematic diagrams illustrating the four selected sites and measured parameters

(A) Selection of four sites in a CBCT image. S0 is the baseline that penetrates the center of mental foramen. Four cross-sectional images (S1, S2, S3 and S4) were selected at intervals of 10mm from S0.
(B) Four measurement parameters - upper height (UH), lower height (LH), lingual width (LW), and buccal width (BW) – used to assess the location of the mandibular canal. The mandibular canal is represented in gray.

Table 1. Definitions and abbreviations of the parameters

Parameters	Abb.	Definition
Upper height	UH	vertical distance from the center of the MC to the occlusal plane
Lower height	LH	vertical distance from the center of the MC to a horizontal line passing through the inferior border of the mandible
Buccal width	BW	horizontal distance from the center of the MC to the buccal border of the mandible
Lingual width	LW	horizontal distance from the center of the MC to the lingual border of the mandible

Abb, abbreviations; MC, mandibular canal.

Table 2. Inter-class reliability of the measured parameters

Slices	ICC value	95% CI
S1	0.999	0.999-0.999
S2	0.995	0.994-0.996
S3	0.995	0.994-0.996
S4	0.949	0.937-0.959
Total	0.983	0.981-0.984

ICC, intra-class correlation coefficient; CI, confidence interval; S1 to S4, the cross-sectional slice that measured parameters.

4. Clustering analysis to classify mandibular canal course

Figure 3 shows the overall workflow of cluster analysis.

In the first stage, parameter selection and normalization of parameters were conducted. Since the location of the MC course can be specified in the form of coordinates by two parameters, it was not necessary to use all four parameters measured at each site. Therefore, one vertical and one horizontal parameter, LH and BW, were selected. Subsequently, the selected parameters measured in millimeters were converted into ratio values for normalization using the formula as follows:

$$\text{LH ratio (\%)} = \frac{LH}{LH + UH} \times 100 \quad (1)$$

$$\text{BW ratio (\%)} = \frac{BW}{LW + BW} \times 100 \quad (2)$$

LH and BW ratio show the location of MC in terms of its proximity to the occlusal plane and lingual border. Increase in the LH and BW ratio imply that the MC position is closer to the occlusal plane and lingual border respectively. In the following step, cluster tendency of the parameters was assessed with Hopkin's statistics[31] and Visual Assessment of cluster Tendency (VAT) algorithm[32] to determine the propriety of performing cluster analysis on the normalized parameters. The Euclidean distance, which is the most widely used method for ratio data, was applied to calculate the dissimilarity between pairs of parameters.[33]

Once clusters are found in the data set, it is necessary to determine the optimal number of clusters. Dendrogram, the output of a hierarchical clustering algorithm, is a tree diagram showing the relationship between parameters. At the beginning stage of clustering, all

parameters are treated as separate clusters, and the process of merging two closest parameters into one single cluster is repeated until they all merge into one cluster. Since dendrogram shows the distance between clusters on the vertical axis, the observer can estimate the optimal number of clusters. In this study, dendrogram was generated using the Ward linkage method, which minimizes the total within-cluster variance.[34] To objectively verify the number of clusters, computing method, NbClust, was then applied. NbClust is a useful package that proposes the best number of clusters by computing all combinations of clustering methods and distance measures.[35] In the final step, given the optimal number of cluster (k) determined by dendrogram and NbClust, k-means clustering algorithm was conducted to the data set in order to identify the center of the cluster to show the standardized courses of the MC.

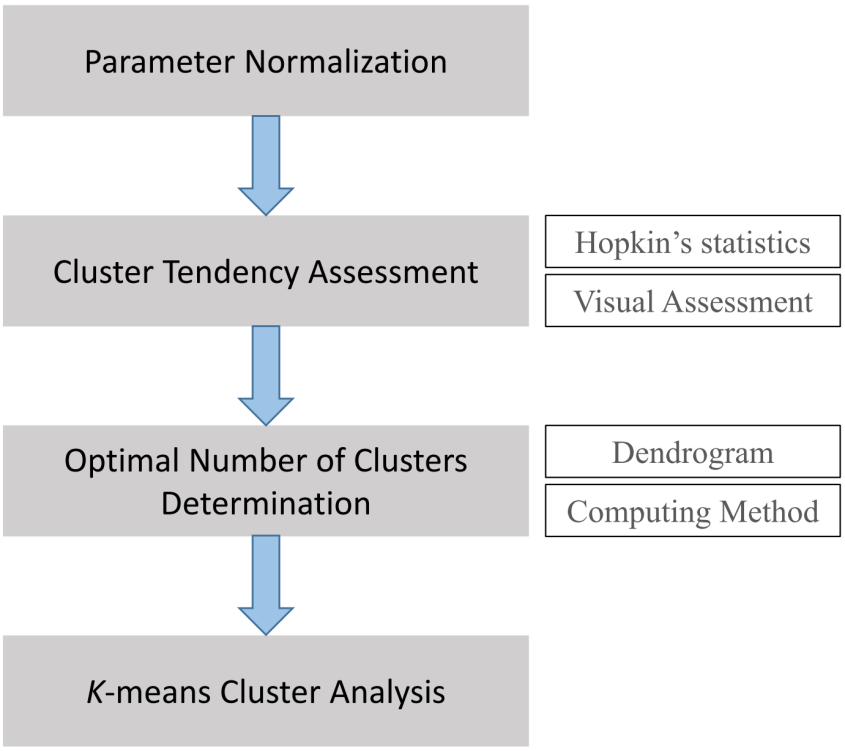

Figure 3. The overall workflow of cluster analysis

5. Statistical analysis

The measured parameters in millimeters were analyzed using descriptive statistics including mean, median, standard deviation, minimum, maximum, and inter quartile range (IQR). Independent t-test and paired t-test were performed to compare the differences in MC course according to sex and sides. IBM SPSS statistics 25.0 software (SPSS Inc., Chicago, IL, USA) was used for statistical analysis with 5% significance level ($P < 0.05$).

All codes for cluster analysis were written in R programming language with various R packages such as clustertend[36], factoextra[37], NbClust[35], and cluster.[38] The codes were run in RStudio (v1.3.1073)[39] based on R for window program (v4.0.2).[40] The equality of proportion of MC courses was analyzed according to sex and sides using Z-test with 5% significance level ($P < 0.05$). To assess the equality of variances for the cluster groups, Levene's test was applied. One-way ANOVA with Bonferroni post-hoc test was used to determine whether the homogeneity of the variables was satisfied, otherwise One-way ANOVA with Welch correction and Games-Howell post-hoc test were conducted ($P < 0.05$).

Ⅲ. RESULTS

1. General characteristics of the subjects

A total of 429 subjects were included in this study. 216 were male (50.3%), and 213 were female (49.7%). The mean age was 38.2 years, and the age ranged from 18 to 81 years. A total of 858 hemi-mandibles, 429 (50.0%) pairs of the right and left mandibles, were analyzed. The age distributions were as follows: 18 – 19 years, 63 (14.7%); 20 – 29 years, 103 (24.0%); 30 – 39 years, 74 (17.2%); 40 – 49 years, 64 (14.9%); and 50 – 59 years, 64 (14.9%); over 60 years, 61 (14.2%) (Table 3).

Table 3. General characteristics of the subjects

Subjects	N (%)
Sex	
Male	216 (50.3)
Female	213 (49.7)
Age subgroups	
18 – 19 years	63 (14.7)
20 – 29 years	103 (24.0)
30 – 39 years	74 (17.2)
40 – 49 years	64 (14.9)
50 – 59 years	64 (14.9)
> 60 years	61 (14.2)

2. Mandibular canal course comparison according to sex and sides

The mean, median, standard deviation (SD), minimum, maximum, and inter quartile range (IQR) of the measured parameters are presented in Table 4. Vertical parameters showed larger SD and IQR values than horizontal parameters, and both values tended to increase toward the posterior cross-sectional images.

Figure 4 presents parameter distributions in box and scatter plots according to sex. All measured parameters except LW of S3 and BW of S4 showed statistically significant differences ($P < 0.05$). Majority of the parameters for female were smaller than those for male, but S1 LW (3.85 for male, 4.03 for female) and S2 LW (3.58 for male, 3.83 for female) were exceptions.

According to Figure 5, vertical parameters of the left and right MC location showed similar mean values without statistical differences ($P < 0.05$). In the horizontal parameters, the left LW was slightly longer than the right LW, and the left BW was shorter than the right BW.

Table 4. Descriptive statistics of the measured parameters

Parameters		Mean	Median	SD	Min	Max	IQR
Vertical parameters (mm)							
S1	LH	9.47	9.35	1.7	3.23	19.37	2.2
	UH	28.54	28.35	3.03	15.81	39.79	3.71
S2	LH	9.93	9.89	1.94	2.94	15.81	2.75
	UH	24.93	24.85	3.23	8.39	34.34	4.25
S3	LH	11.18	11.12	2.37	3.56	23.02	3.19
	UH	19.77	19.78	3.63	6.87	30.31	4.75
S4	LH	15.13	15.15	3.18	1.35	23.74	4.31
	UH	11.79	12.19	4.34	0	26.12	5.66
Horizontal parameters (mm)							
S1	LW	3.96	3.85	1.26	0.76	8.07	1.78
	BW	7.69	7.63	1.27	4.25	11.94	1.65
S2	LW	3.71	3.6	1.15	1.02	9.59	1.64
	BW	8.69	8.62	1.49	4.69	28.89	1.78
S3	LW	4.02	3.92	1.2	0.65	12.2	1.53
	BW	8.98	8.92	1.66	4.36	23.98	2.08
S4	LW	3.82	3.71	1.51	0.33	20.1	1.77
	BW	6.9	6.76	1.77	3.02	16.72	2.29

S1 to S4 represent four sites used in the measurement of the mandibular canal location.

SD, standard deviation; Min, minimum; Max, maximum, IQR, inter quartile range; LH, lower height; UH, upper height; LW, lingual width; BW, buccal width.

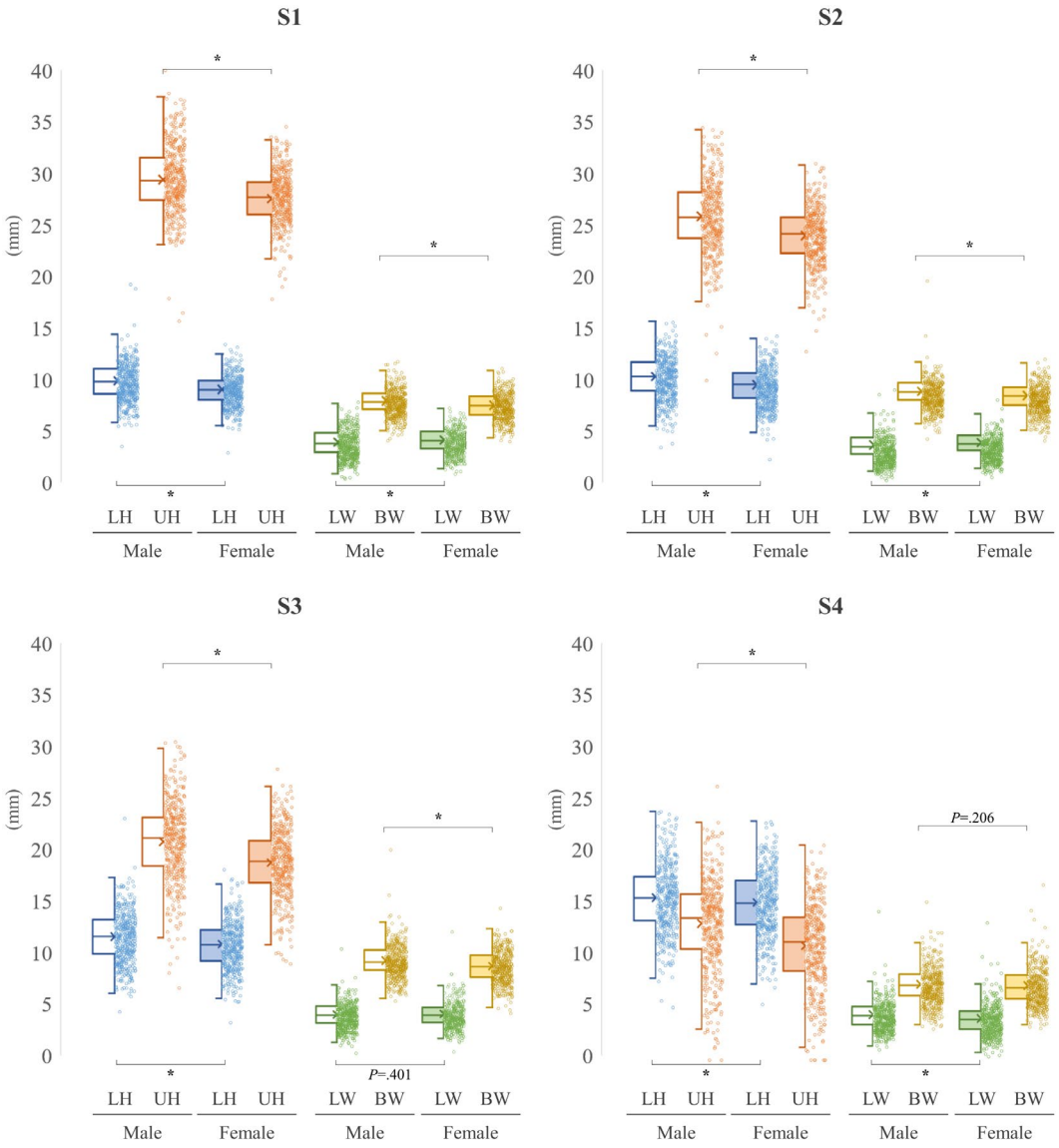

Figure 4. Box and scatter plots for measured parameters at each site according to sex

Line inside the box indicates the median; X, the mean; box, the interquartile range. LH, lower height; UH, upper height; LW, lingual width; BW, buccal width. *P*-value by Independent t-test at α=0.05. *, $P < .05$.

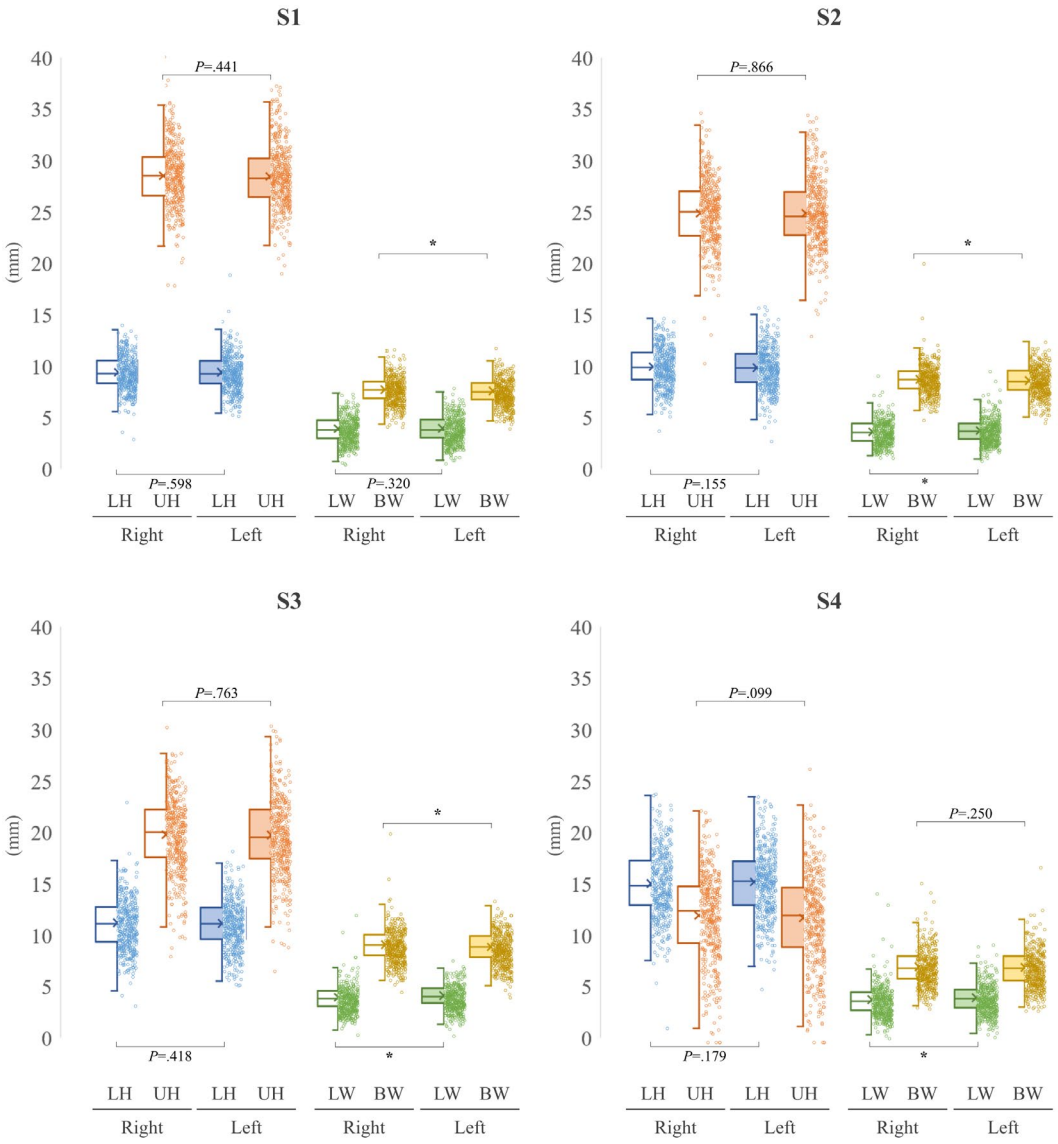

Figure 5. Box and scatter plots for measured parameters at each site according to right and left sides

Line inside the box indicates the median; X, the mean; box, the interquartile range. LH, lower height; UH, upper height; LW, lingual width; BW, buccal width. *P*-value by Independent t-test at α=0.05. *, $P < .05$.

3. Cluster analysis of 3D mandibular canal course

The results of cluster tendency in normalized parameters were presented by Hopkin's statistics value and dissimilarity matrix. Calculated Hopkin's statistics for the MC parameters was 0.2096047, far below the H threshold of 0.50. Figure 6 displays the dissimilarity matrix that obtained from VAT algorithm. Color is the indicator of the degree of dissimilarity between the parameters. Low dissimilarity is represented by orange color and high dissimilarity by purple. Cluster structures of the parameter data are noticeable from the orange areas of the matrix.

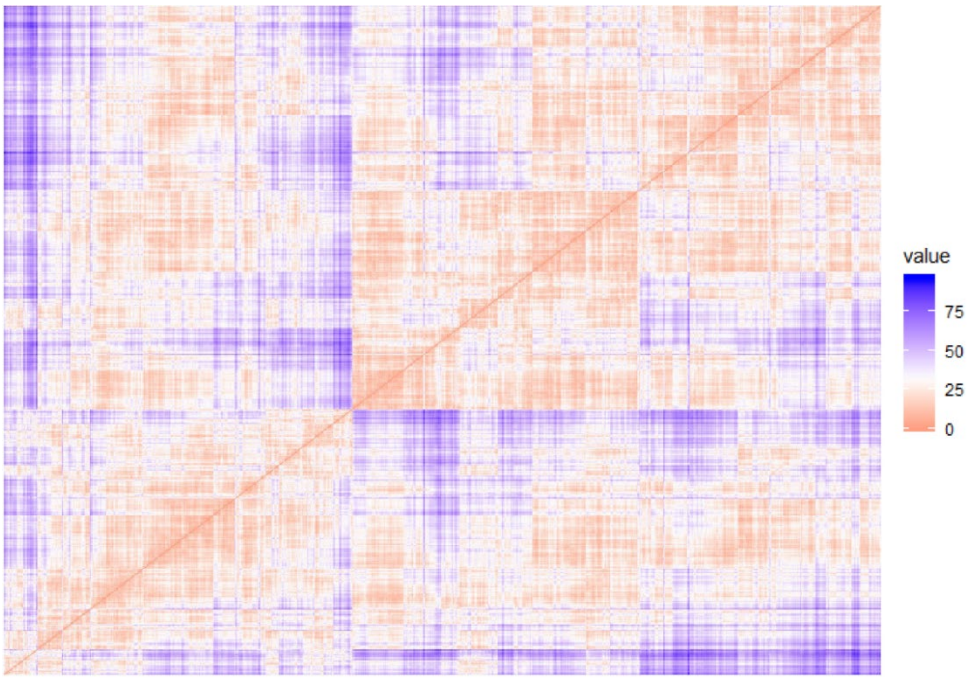

Figure 6. Dissimilarity matrix visualizing the clustering tendency of the parameters
Orange shows low dissimilarity level and purple shows high dissimilarity level.

The optimal number of clusters was determined by dendrogram and computing method. Figure 7 shows a dendrogram that was generated as the result of hierarchical clustering conducted for the classification 3D course of MC. Considering the variation of the MC course pattern, the range of the optimal number of clusters can be estimated from two clusters (400 height level) to five clusters (200 height level).

In Figure 8A, 11 indices suggest three as the optimal number of clusters. In addition, the second differences D index value was the highest for the estimated number of clusters (*k*) at three (Figure 8B).

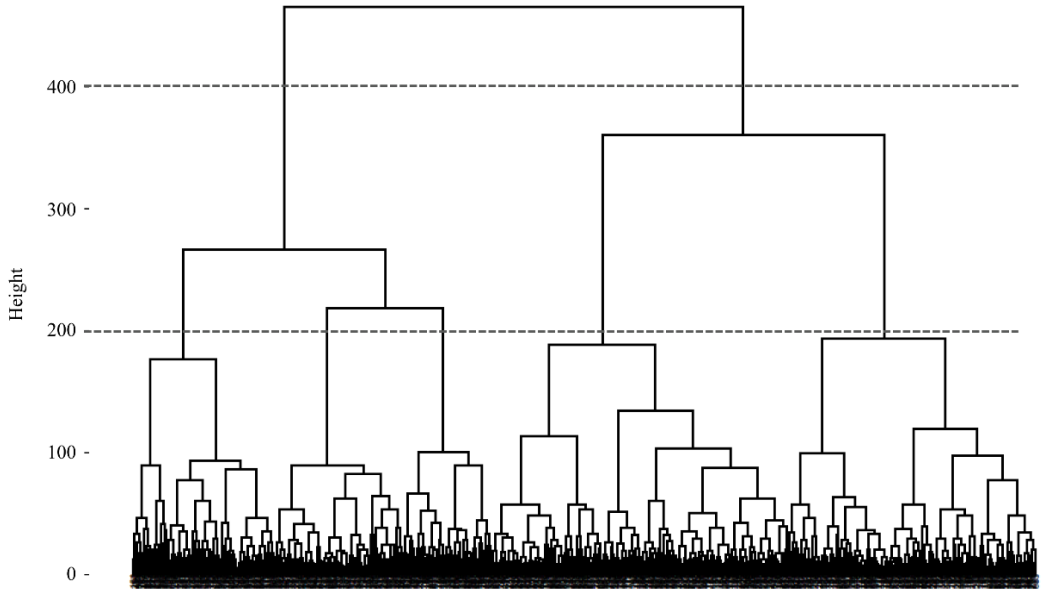

Figure 7. The cluster dendrogram generated as the result of hierarchical clustering of the three-dimensional mandibular canal course parameters

The vertical axis represents the distance between clusters and the horizontal axis represents individual variables. The dotted lines show an example of determining the optimal number of clusters. The clusters are classified into 2 and 5 groups at height 400 and 200 respectively.

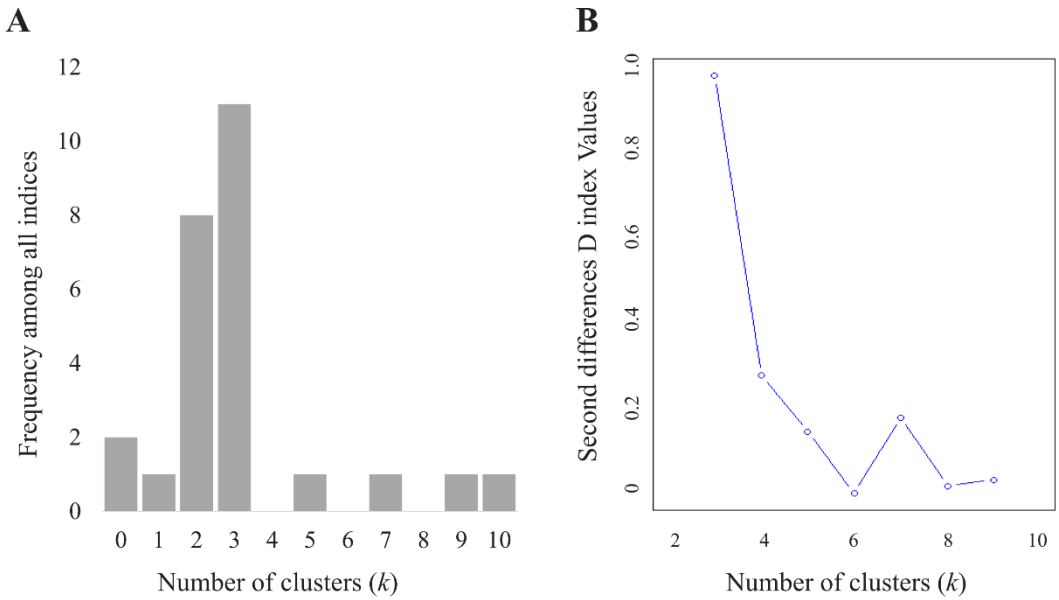

Figure 8. Determination of the optimal number of clusters using NbClust

(A) Frequency distribution of NbClust indices for the number of clusters

(B) D index graph for determining the number of clusters

Based on the result of *k*-means cluster analysis, the mean and standard deviation of three clusters of the 3D MC course were presented on Table 5. There were statistically significant differences among the means of the three clusters both in vertical and horizontal parameters. Cluster 1 showed the most rapid vertical change, ranging from 27.9 ± 4.5% in S1 to 74.0 ± 9.9% in S4. Conversely, cluster 2 showed the lowest vertical increase rate from S1 (23.5 ± 3.8%) to S4 (48.6 ± 7.7%). On the other hand, cluster 3 showed the largest distribution range in horizontal view, from 53.9 ± 8.1% to 65.9 ± 6.3%, whereas cluster 1 had only 4.6% variation range.

Figure 9 shows individual MC courses and one standard MC course, in a broader line, of each cluster. Cluster 1 showed a smooth line running towards the lingual side in the axial view (Figure 9A) and a steep slope in the sagittal view (Figure 9D). Cluster 2 ran in an almost straight line closest to the lingual and inferior border of mandible (Figure 9B and E); it showed the least change in the driving course compared to other MC course types. Cluster 3 showed the pathway with a bent buccally in the axial view and an increasing slope in the sagittal view in the posterior area (Figure 9C and F); this type showed a larger change in the posterior area. Illustrations in figure 10A and B show the superimposed courses of all three clusters, making it easy to compare relative positions of the 3D MC courses of different clusters.

Distributions of the mandibular canal courses according to clusters, age groups, sex, and right and left sides are displayed in Figure 11. Figure 11A and C shows the distributions of classified clusters and sex within clusters. Cluster 1 had the least distribution (26.0%). In

this cluster, females accounted for 59.2% and males, 40.8%. On the other hand, cluster 2 showed the widest distribution of driving courses (42.1%), and males were more widely distributed (57.1%) than the female group (42.9%). Cluster 3 comprised similar ratio of male and female cases and accounted for 31.9% of the total distribution. For all three clusters, distributions of the right and left sides did not show a statistically significant difference (Figure 11D). Distribution of cluster was inversely proportional to age of subject, and cluster 1 showed the highest distribution (41.8%) in the age of 60 or older (Figure 11B).

Table 5. Mean and standard deviation of the three-dimensional mandibular canal courses according to cluster type

	Cluster 1 (N = 223)	Cluster 2 (N = 361)	Cluster 3 (N = 274)	*P-value*
Vertical parameters (%)				
S1	27.9 ± 4.5 [a]	23.5 ± 3.8 [b]	24.5 ± 3.9 [c]	0.001†
S2	33.4 ± 5.0 [a]	26.1 ± 4.4 [b]	27.9 ± 4.2 [c]	0.001†
S3	45.0 ± 6.1 [a]	32.2 ± 5.4 [b]	34.7 ± 4.9 [c]	0.001§
S4	74.0 ± 9.9 [a]	48.6 ± 7.7 [b]	54.3 ± 8.4 [c]	0.001§
Horizontal parameters (%)				
S1	65.2 ± 8.3 [a]	70.0 ± 7.7 [b]	62.3 ± 8.0 [c]	0.001†
S2	69.2 ± 7.0 [a]	74.4 ± 6.0 [b]	65.9 ± 6.3 [c]	0.001†
S3	69.8 ± 7.0 [a]	73.5 ± 5.5 [b]	62.9 ± 5.8 [c]	0.001§
S4	68.9 ± 10.3 [a]	69.6 ± 7.6 [a]	53.9 ± 8.1 [b]	0.001§

Values are presented as mean ± standard deviation in ratio.

† *P*-value by One-way ANOVA at α=0.05,

§ *P*-value by One-way ANOVA with Welch correction at α=0.05.

S1 to S4 represent four sites measured vertical and horizontal parameters.

[a-c] Superscripts on the same row indicate significant statistical differences (Bonferroni post-hoc test for S1 and S2, and Games-Howell post-hoc test for S3 and S4)

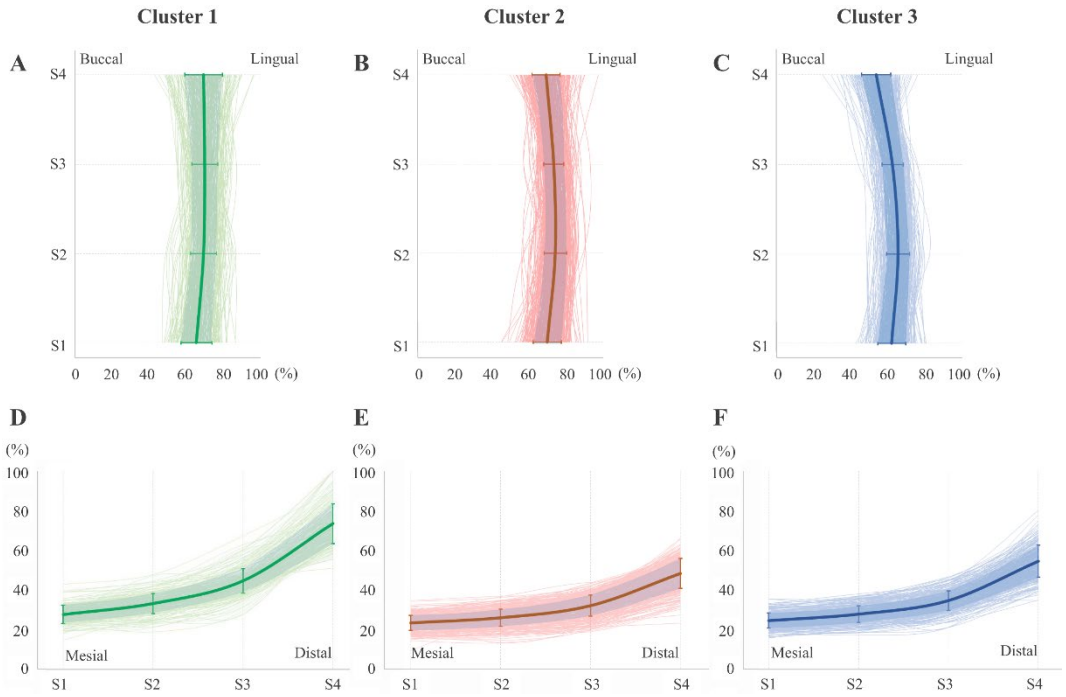

Figure 9. Three types of the mandibular canal 3D course in axial (first row) and sagittal view (second row)

The green, red, and blue lines display cluster 1 (A and D), 2 (B and E), and 3 (C and F), respectively. Thin lines represent each individual mandibular canal course, and the thick lines represent the standard mandibular canal 3D courses. The opaque area around the thick line shows standard deviation.

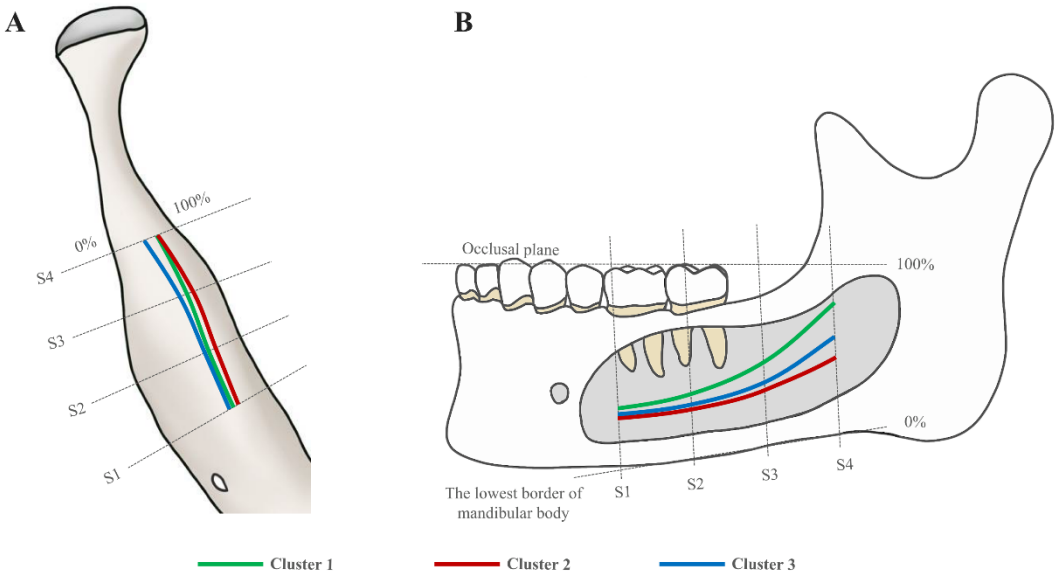

Figure 10. Illustrations of three types of the three-dimensional mandibular canal course in axial (A) and sagittal view (B)

Cluster 1, green line, shows traveling towards the lingual side with a steep vertical slope.
Cluster 2, red line, shows the closest to the lingual side and lowest vertical slope.
Cluster 3, blue line, shows a buccal curvature with gradually increasing vertical slope in the posterior area.

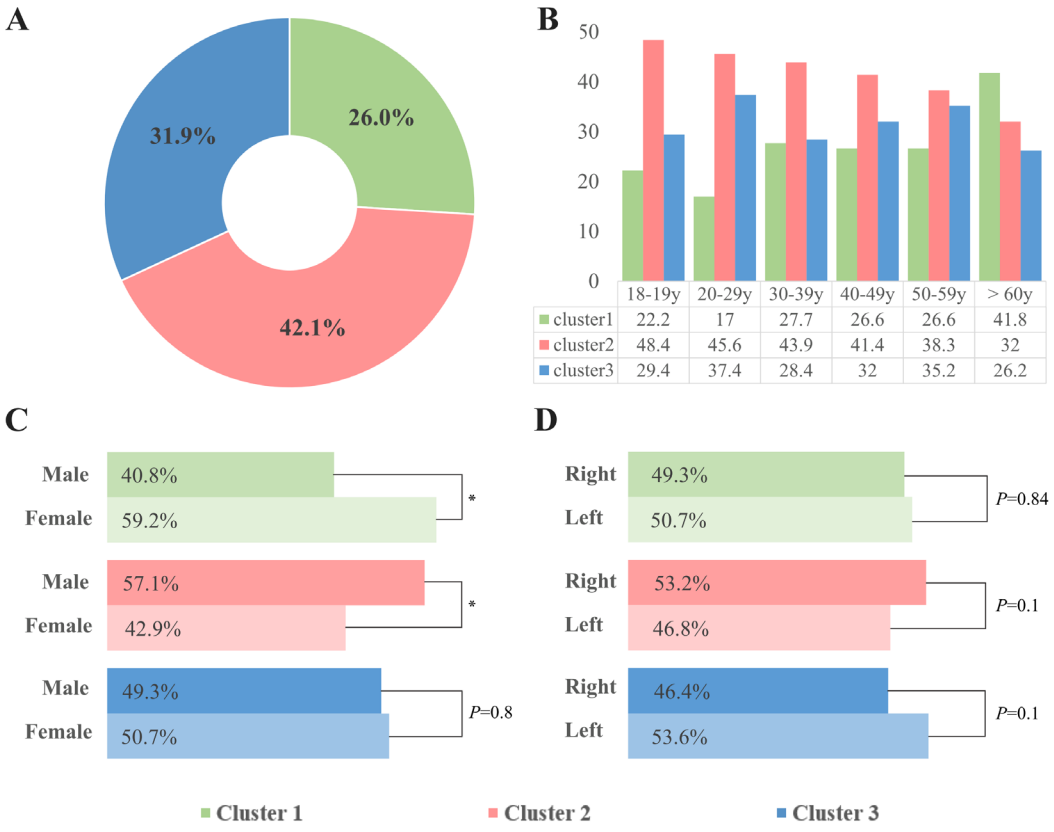

Figure 11. Distribution of the three-dimensional mandibular canal course by cluster

(A) Distributions of three clusters

(B) Distribution of three clusters for each age group

(C) Sex distribution of each cluster (*P*-value by Z-test at α=0.05. *, $P < .05$.)

(D) Right and left distribution of each cluster

Ⅳ. DISCUSSION

Since standardized information of anatomical structure facilitates communication among clinicians,[41] many studies have been conducted to provide reliable classification. Existing classifications of the anatomical structures were commonly developed through visual assessment of cadaveric dissections, and has been used in practice to date.[42] However, the inevitable variability in interpretation among experts proves that the use of previous classifications as standard information may be limited. In medical field, cluster analysis has been widely utilized to objectively distinguish the given data such as gene and patient information,[30,43] but less applications are conducted for the anatomical structure classification. In this study, cluster analysis was applied to classify 3D MC courses and to present standard types of MC courses.

To apply cluster analysis in anatomical structure classification, it is necessary to measure the anatomical structure and parameterize it in a reproducible way. In this study, measurement of parameters was taken at four different sites of the MC at 10mm interval to cover the widest possible range of the MC. To the next, only the minimum number of parameters that can represent the MC location were selected, and finally, normalization of parameters was carried out so that they were not affected by the size of the mandible that varies depending on the person.[44] This process is worth considering not only for MC courses, but for researchers who want to try cluster analysis with measurement parameters of other anatomical structures.

Based on the results of current study, three representative 3D MC course types were

derived from the cluster method. In 2004, Worthington classified the vertical MC courses into three categories: an anteriorly gentle and posteriorly progressive rising curve, a steep ascent, and a catenary-like canal.[13] Many researchers have used this classification or modified it to present the MC course and its distribution.[17,45-47] In the sagittal view of the 3D MC courses visualized in this study, a steep vertical slope (cluster 1) and the straight path closest to the inferior border (cluster 2) were presented similarly with the Worthington classification, but a catenary-like pathway was not found. This phenomenon is the same for the axial view of the MC course. The individual MC courses clustered in the current study differed from the known horizontal MC courses.[21,47,48] Discrepancy between the results of previous studies and this study may arise from different methodologies. Most of the research were confined to two-dimensional (2D) analysis of the MC course, viewed from either axial or sagittal view. In other words, the classification result was presented after considering the vertical and horizontal location of MC, respectively. On the other hand, in the present study, both horizontal and vertical parameters of the MC course were considered together by cluster algorithm firstly; and then for each cluster, 2D axial and sagittal images of the MC courses were obtained. To accurately classify the anatomical structure, an analysis taking into account the 3D information is required.

Interestingly, if the data of this study were intently divided into a larger number of clusters, the MC course would be grouped into more diverse clusters, and the MC courses reported in previous studies (e.g., catenary-like courses) may also be found among them. However, in the current study, the optimal number of clusters was established as three using

dendrogram and computing method to derive objective classification. Although statistical representativeness is reduced, if clinically necessary, observers can classify MC courses in detail with the k-means cluster algorithm.

Considering the distribution trend of the MC courses according to age group, cluster 1 showed the highest value in the group over 60 years old, and cluster 2 gradually decreased as age increased. This may be due to age-related factors such as tooth loss and attrition that cause a decrease in MnOP levels.[49-51] Levine et al.[52] reported that age and race were related to the MC location, so further study of the MC courses by race using cluster analysis is needed. As a result of analyzing the left and right distribution, no significant difference was found in accordance with the previous study, which reported that right and left side were not factors affecting the location of the MC.[53]

Expert reliability is one of the major issues to be addressed in anatomy structure classification task.[54,55] Although ensuring expert qualifications through training is widely carried out to improve reliability, even so, subjective evaluation among observers is not completely overcome.[56] Given the high variability in the anatomical structure, expert disagreement on cases where grouping is ambiguous can affect the classification results. In this study, the cluster algorithm was designed to divide groups according to the similarities between all given variables, so it automatically classified debatable MC course cases without observer variability and present representative results. This unsupervised machine learning technique can be applied to analysis of various anatomical information that has been mainly performed by experts, and the presented standard information can be used in

divergent fields such as dental treatment planning, forensic usage, and model generation for reconstructive surgery.

The limitation of this study was that it did not include areas near the mental and mandibular foramen. Regarding that the anterior loop region is accompanied by high anatomical variation,[57] it is challenging to measure the area in a standardized way in all patients. However, even if the number of additional parameters was updated as new areas were included, cluster analysis would make it easy to re-classify MC course types, considering the similarity of all given variables.

V. CONCLUSION

The 3D MC courses were automatically classified as three types through cluster analysis. The first MC course type, cluster 1, can be defined as the traveling towards the lingual side with a steep vertical slope. The second cluster can be defined as the closest to the lingual side and lowest vertical slope. The third type of MC course can be defined as a buccal curvature with gradually increasing vertical slope in the posterior area. Each classified MC course type showed a driving path that was statistically distinct from other clusters. In addition, this study demonstrates the potential of cluster analysis for the analysis of parameters measured in medical imaging. Cluster analysis enables the unbiased classification of the anatomical structures by reducing observer variability and can present representative standard information for each classified group.

REFERENCES

1. Joo W, Yoshioka F, Funaki T, Mizokami K, Rhoton AL, Jr. Microsurgical anatomy of the trigeminal nerve. *Clin Anat* 2014; **27**: 61-88.
2. Matani JD, Kheur MG, Kheur SM, Jambhekar SS. The Anatomic Inter Rel ationship of the Neurovascular Structures Within the Inferior Alveolar Cana l: A Cadaveric and Histological Study. *J Maxillofac Oral Surg* 2014; **13**: 4 99-502.
3. Carter RB, Keen EN. The intramandibular course of the inferior alveolar n erve. *Journal of anatomy* 1971; **108**: 433-440.
4. Alhassani AA, AlGhamdi AS. Inferior alveolar nerve injury in implant den tistry: diagnosis, causes, prevention, and management. *J Oral Implantol* 201 0; **36**: 401-407.
5. Yu SK, Lee MH, Jeon YH, Chung YY, Kim HJ. Anatomical configuration of the inferior alveolar neurovascular bundle: a histomorphometric analysis. *Surg Radiol Anat* 2016; **38**: 195-201.
6. Juodzbalys G, Wang H-L, Sabalys G. Injury of the Inferior Alveolar Nerve during Implant Placement: a Literature Review. *Journal of oral & maxillo facial research* 2011; **2**: e1-e1.
7. Bertram F, Bertram S, Rudisch A, Emshoff R. Assessment of Location of the Mandibular Canal: Correlation Between Panoramic and Cone Beam Co mputed Tomography Measurements. *Int J Prosthodont* 2018; **31**: 129–134.

8. Tay AB, Zuniga JR. Clinical characteristics of trigeminal nerve injury referrals to a university centre. *Int J Oral Maxillofac Surg* 2007; **36**: 922-927.

9. Morse DR. Endodontic-related inferior alveolar nerve and mental foramen paresthesia. *Compend Contin Educ Dent* 1997; **18**: 963-968, 970-963, 976-968 passim; quiz 998.

10. Pogrel MA. Permanent nerve damage from inferior alveolar nerve blocks: a current update. *J Calif Dent Assoc* 2012; **40**: 795-797.

11. Abarca M, van Steenberghe D, Malevez C, De Ridder J, Jacobs R. Neurosensory disturbances after immediate loading of implants in the anterior mandible: an initial questionnaire approach followed by a psychophysical assessment. *Clin Oral Investig* 2006; **10**: 269-277.

12. Renton T. Prevention of iatrogenic inferior alveolar nerve injuries in relation to dental procedures. *Dent Update* 2010; **37**: 350-352, 354-356, 358-360 passim.

13. Worthington P. Injury to the inferior alveolar nerve during implant placement: a formula for protection of the patient and clinician. *Int J Oral Maxillofac Implants* 2004; **19**: 731-734.

14. Jung YH, Cho BH. Radiographic evaluation of the course and visibility of the mandibular canal. *Imaging Sci Dent* 2014; **44**: 273-278.

15. Liu T, Xia B, Gu Z. Inferior alveolar canal course: a radiographic study. *Clin Oral Implants Res* 2009; **20**: 1212-1218.

16. Starkie C, Stewart D. The Intra-Mandibular Course of the Inferior Dental Nerve. *Journal of anatomy* 1931; **65**: 319-323.

17. Mirbeigi S, Kazemipoor M, Khojastepour L. Evaluation of the Course of t he Inferior Alveolar Canal: The First CBCT Study in an Iranian Populatio n. *Pol J Radiol* 2016; **81**: 338-341.

18. Zhang LZ, Meng SS, He DM, Fu YZ, Liu T, Wang FY, et al. Three-Dim ensional Measurement and Cluster Analysis for Determining the Size Rang es of Chinese Temporomandibular Joint Replacement Prosthesis. *Medicine (Baltimore)* 2016; **95**: e2897.

19. van Riel SJ, Sánchez CI, Bankier AA, Naidich DP, Verschakelen J, Scholte n ET, et al. Observer Variability for Classification of Pulmonary Nodules o n Low-Dose CT Images and Its Effect on Nodule Management. *Radiology* 2015; **277**: 863-871.

20. Safaee A, Mirbeigi S, Ezoddini F, Khojastepour L, Navab-Azam A. Buccol ingual course of the inferior alveolar canal in different mental foramen loc ations: A cone beam computed tomography study of an Iranian population. *Int J Appl Basic Med Res* 2016; **6**: 262-266.

21. Pyun JH, Lim YJ, Kim MJ, Ahn SJ, Kim J. Position of the mental foram en on panoramic radiographs and its relation to the horizontal course of th e mandibular canal: a computed tomographic analysis. *Clin Oral Implants Res* 2013; **24**: 890-895.

22. Romesburg HC. *Cluster analysis for researchers*. [Raleigh]: Lulu Press; 2007.

23. Hyun S, Kaewprag P, Cooper C, Hixon B, Moffatt-Bruce S. Exploration of critical care data by using unsupervised machine learning. *Comput Methods Programs Biomed* 2020; **194**: 105507.

24. Hosseini SMS, Maleki A, Gholamian MR. Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty. *Expert Systems with Applications* 2010; **37**: 5259-5264.

25. Kim CS, Park MR. A study on the appearance care behaviors, clothing selection behaviors and clothing design preference of 20-30's Korean men by the level of grooming. *Fashion & Textile Research Journal* 2014; **16**: 245-254.

26. Tuan TM. A cooperative semi-supervised fuzzy clustering framework for dental X-ray image segmentation. *Expert Systems with Applications* 2016; **46**: 380-393.

27. Lance G, Williams WT. A General Theory of Classificatory Sorting Strategies: 1. Hierarchical Systems. *Comput. J.* 1967; **9**: 373-380.

28. MacQueen J. Some methods for classification and analysis of multivariate observations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*. Berkeley, Calif.: University of California Press; 1967. pp. 281-297.

29. Dilts D, Khamalah J, Plotkin A. Using cluster analysis for medical resource decision making. *Med Decis Making* 1995; **15**: 333-347.

30. Liao M, Li Y, Kianifard F, Obi E, Arcona S. Cluster analysis and its application to healthcare claims data: a study of end-stage renal disease patients who initiated hemodialysis. *BMC Nephrol* 2016; **17**: 25.

31. Hopkins B, Skellam JG. A new method for determining the type of distribution of plant individuals. *Annals of Botany* 1954; **18**: 213-227.

32. Bezdek JC, Hathaway RJ. VAT: A tool for visual assessment of (cluster) tendency. *Proceedings of the 2002 International Joint Conference on Neural Networks. IJCNN'02 (Cat. No. 02CH37290)*: IEEE; 2002. pp. 2225-2230.

33. Afifi AA, May S, Clark V. *Practical multivariate analysis*. Boca Raton: Taylor & Francis; 2012.

34. Ward Jr JH. Hierarchical grouping to optimize an objective function. *Journal of the American statistical association* 1963; **58**: 236-244.

35. Charrad M, Ghazzali N, Boiteau V, Niknafs A. NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set. *2014* 2014; **61**: 36.

36. YiLan L, RuTong Z. clustertend: Check the Clustering Tendency. *R package version* 2015; **1**.

37. Kassambara A, Mundt F. Factoextra: extract and visualize the results of multivariate data analyses. *R package version* 2017; **1**: 337-354.

38. Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K. Cluster: Cluste r analysis basics and extensions. R package v. 2.0. 5. 2016.

39. RStudio Team. RStudio: Integrated Development for R. PBC, Boston, MA: RStudio; 2019.

40. R Core Team. R: A language and environment for statistical computing. In: RStudio, ed. Vienna, Austria: R Foundation for Statistical Computing; 201 9.

41. Kim M, Boyle SL, Fernandez A, Matsumoto ED, Pace KT, Anidjar M, et al. Development of a novel classification system for anatomical variants of the puboprostatic ligaments with expert validation. *Canadian Urological Ass ociation journal = Journal de l'Association des urologues du Canada* 2014; **8**: 432-436.

42. Aziz MA, McKenzie JC, Wilson JS, Cowie RJ, Ayeni SA, Dunn BK. The human cadaver in the age of biomedical informatics. *Anat Rec* 2002; **269**: 20-32.

43. Sharan R, Elkon R, Shamir R. Cluster Analysis and Its Applications to Ge ne Expression Data. Berlin, Heidelberg: Springer Berlin Heidelberg; 2002. pp. 83-108.

44. Franklin D, Cardini A, Flavel A, Kuliukas A. Estimation of sex from cran ial measurements in a Western Australian population. *Forensic Sci Int* 201 3; **229**: 158.e151-158.

45. Vieira CL, Veloso S, Lopes FF. Location of the course of the mandibular canal, anterior loop and accessory mental foramen through cone-beam computed tomography. *Surg Radiol Anat* 2018; **40**: 1411-1417.

46. Abdallah Edrees M, Moustafa Attia A, Abd Elsattar M, Fahmy Gobran H, Ismail Ahmed A. Course and topographic relationships of mandibular canal: A cone beam computed tomography study. *Int J Dentistry Oral Sci* 2017; **4**: 444-449.

47. Ozturk A, Potluri A, Vieira AR. Position and course of the mandibular canal in skulls. *Oral Surg Oral Med Oral Pathol Oral Radiol* 2012; **113**: 453-458.

48. Shokry SM, Alshaib SA, Al Mohaimeed ZZ, Ghanimah F, Altyebe MM, Alenezi MA, et al. Assessment of the Inferior Alveolar Nerve Canal Course Among Saudis by Cone Beam Computed Tomography (Pilot Study). *J Maxillofac Oral Surg* 2019; **18**: 452-458.

49. Van't Spijker A, Rodriguez JM, Kreulen CM, Bronkhorst EM, Bartlett DW, Creugers NH. Prevalence of tooth wear in adults. *Int J Prosthodont* 2009; **22**: 35-42.

50. Kumar M, Verma R, Bansal M, Singh S, Rehan S, Kumar V, et al. To Evaluate the Severity, Distribution of Occlusal Tooth Wear and its Correlation with Bite Force in Young North Indian Adults. *The open dentistry journal* 2018; **12**: 735-741.

51. Assis E, Aguiar F, Pereira R, Velo M, Lima D, Giorgi M. Re-Establishment of an Occlusal Vertical Dimension: A Case Report. *J Dent Health Oral Disord Ther* 2018; **9**: 00336.

52. Levine MH, Goddard AL, Dodson TB. Inferior alveolar nerve canal position: a clinical and radiographic study. *J Oral Maxillofac Surg* 2007; **65**: 470-474.

53. Al-Siweedi SYA, Nambiar P, Shanmuhasuntharam P, Ngeow WC. Gaining Surgical Access for Repositioning the Inferior Alveolar Neurovascular Bundle. *The Scientific World Journal* 2014; **2014**: 719243.

54. Audigé L, Bhandari M, Kellam J. How reliable are reliability studies of fracture classifications? A systematic review of their methodologies. *Acta Orthop Scand* 2004; **75**: 184-194.

55. Villarreal R, Wrobel BB, Macias-Valle LF, Davis GE, Prihoda TJ, Luong AU, et al. International assessment of inter- and intrarater reliability of the International Frontal Sinus Anatomy Classification system. *Int Forum Allergy Rhinol* 2019; **9**: 39-45.

56. Buijze GA, Guitton TG, van Dijk CN, Ring D, Science of Variation G. Training improves interobserver reliability for the diagnosis of scaphoid fracture displacement. *Clinical orthopaedics and related research* 2012; **470**: 2029-2034.

57. Prakash O, Srivastava PK, Jyoti B, Mushtaq R, Vyas T, Usha P. Radiograp

hic Evaluation of Anterior Loop of Inferior Alveolar Nerve: A Cone-Beam Computer Tomography Study. *Niger J Surg* 2018; **24**: 90-94.

Abstract (in Korean)

비지도 머신러닝 방법을 이용한 3차원 하악 신경관 주행 경로의 자동 군집화

< 지도교수 한 상 선 >

연세대학교 대학원 치의학과

김 영 현

연구목적: 본 연구의 목적은 하악 신경관의 3차원 주행 경로 분류를 위해 비지도 머신 러닝 기법인 군집 분석을 적용하고, 군집 분석을 통해 제시된 한국인의 표준 하악 신경관 주행 경로를 시각화하는 것이다.

연구대상 및 방법: 연세대학교 치과 대학병원에서 촬영된 총 429개의 치과용 콘빔시티 영상이 이용되었다. 하악 신경관의 주행의 측정을 위해 총 4개의 단면 영상이 선택되었고, 각 부위 별 2개의 수직 및 2개의 수평 변수를 획득하

였다. 군집 분석은 변수 선택, 변수 표준화, 군집 경향성 평가, 최적 군집 수 결정, k-mean 군집 분석의 순서로 진행되었다.

연구결과: 군집 분석에 의해 세 가지 유형의 하악 신경관 주행 경로가 도출되었으며 군집 간에 통계적으로 유의한 평균 차이를 보였다. 군집 1은 횡단면 뷰에서 설측 방향으로 주행하고, 시상면 뷰에서 가장 급격한 경사를 보이는 유형이다. 군집 2는 하악 하연 및 설측 경계에 가장 가깝게 거의 직선으로 주행하는 유형이다. 군집 3은 후방 영역에서 협측 구부러짐과 경사 증가가 나타나는 유형이다. 군집 1유형은 26.0%로 가장 적게 분포하였고, 해당 유형에서 여성이 59.2%, 남성이 40.8%로 나타났다. 반면 군집 2유형은 42.1%로 가장 많은 분포를 보였으며, 여성(42.9%)에 비해 남성(57.1%)에서 더 많이 분포하였다. 군집 3은 전체의 31.9%가 분포하였으며 성별 비율은 유사하였다. 3개 군집 모두 좌우측에 대해 통계적으로 유의한 분포차는 없었다.

결론: 군집 분석을 통해 하악 신경관의 3차원 주행은 세가지 유형으로 자동 분류되었다. 군집 분석은 관찰자 변동성을 감소시켜 해부학적 구조의 편향되지 않은 분류를 가능하게 하며, 각 분류 그룹에 대한 대표적인 표준 정보를 제공하는데 사용될 수 있다.

핵심어: 군집 분석, 콘빔전산화단층촬영, 하치조신경, 하악 신경관, 비지도 머신러닝