Enumerated sparse extraction of important surgical planning features for mandibular reconstruction

Kazuki Nagai[1], *Student Member*, *IEEE*, Megumi Nakao[1], *Member*, *IEEE*, Nobuhiro Ueda[2], Yuichiro Imai[3], Tadaaki Kirita[2], and Tetsuya Matsuda[1], *Member*, *IEEE*

***Abstract*— Because implicit medical knowledge and experience are used to perform medical treatment, such decisions must be clarified when systematizing surgical procedures. We propose an algorithm that extracts low-dimensional features that are important for determining the number of fibular segments in mandibular reconstruction using the enumeration of Lasso solutions (eLasso). To perform the multi-class classification, we extend the eLasso using an importance evaluation criterion that quantifies the contribution of the extracted features. Experiment results show that the extracted 7-dimensional feature set has the same estimation performance as the set using all 49-dimensional features.**

I. INTRODUCTION

Clinical treatment is performed by surgeons and medical staff using a large amount of medical knowledge and experience. For instance, in mandibular reconstruction surgery, in which the patient's own fibula is transplanted [1] [2], the number of fibular segments and the placement in the mandible must be determined. Moreover, various reconstruction plans can be considered depending on the variety in the shape of the mandibles. The quality such plans is difficult to guarantee because manual planning largely depends on the success of past cases and because the planning policy depends on the subjectivity and experience of surgeons. Hence, the systematization of these surgical procedures and a more objective and low-cost surgical planning method are required. In recent years, an automatic planning system for mandibular reconstruction was proposed [3]. However, designing indicators based on the insights of researchers is time consuming and lacks objectivity [4]. However, the high-dimensional deep features that are automatically extracted from medical images are not based on clinically used anatomical features, and it is not possible to understand them from an expert's viewpoint. To obtain a reliable surgical plan, it is necessary to use uniquely calculated features that can be interpreted by humans.

This study focuses on analysis methods for extracting essential anatomical knowledge from surgical planning databases that have been manually annotated by surgeons for past surgeries. In the last decade, sparse modeling, which can extract essential information using the sparseness that is inherently universal in data, has attracted attention. Kawasaki et al. showed that patient-specific surgical plans can be estimated from past plans using sparse modeling [5]. One of the most common feature selection methods is the least absolute shrinkage and selection operator (Lasso) [6]. In the Lasso, an important feature set is obtained. However, this approach can overlook possibly relevant features not selected by the Lasso. Recently, the enumeration of Lasso solutions (eLasso) for finding multiple models using different feature sets has been proposed [7]. However, this algorithm is intended for the binary classification and cannot classify multiple classes.

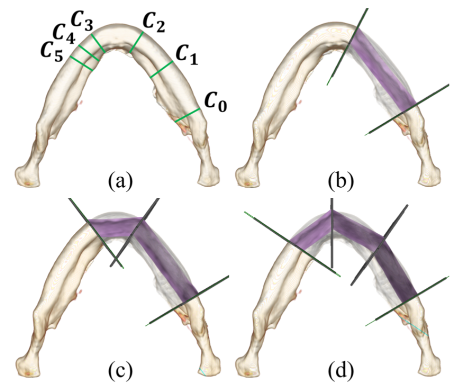


Fig. 1. Cutting planes and examples of reconstruction plans: (a) cutting planes and cases using (b) one, (c) two, and (d) three fibular segments.

In this study, we hypothesized that surgical planning, specifically the number of fibular segments used in mandibular reconstruction, can be determined by a low-dimensional feature set T_0. When the Lasso is used, only one feature set is extracted. Therefore, this study extends the concept of eLasso for selecting surgical procedures in fibular-segment based mandibular reconstruction. Our aim is to enumerate multiple sets of low-dimensional features that are important for determining the number of fibular segments. In this paper, we verify the effectiveness of the proposed method using 232 reconstruction plans given by an oral surgeon.

II. METHODS

A. Surgical planning datasets

In this study, we use 3D-CT datasets obtained from 29 patients [3] [8]. The mandible region was extracted from head and neck images, and the fibular segments were extracted from foot images. Then, to evaluate the differences in the resection areas, as shown in Fig. 1(a), we defined six cutting planes. Each cutting plane was defined based

[1]K. Nagai, M. Nakao and T. Matsuda are with Graduate School of informatics, Kyoto University, Kyoto, Japan (e-mail: knagai@sys.i.kyoto-u.ac.jp).
[2]N. Ueda and T. Kirita are with the Department of Oral and Maxillofacial Surgery, Nara Medical University, Nara, Japan.
[3]Y. Imai is with the Department of Oral and Maxillofacial Surgery, Rakuwakai Otowa Hospital, Kyoto, Japan.

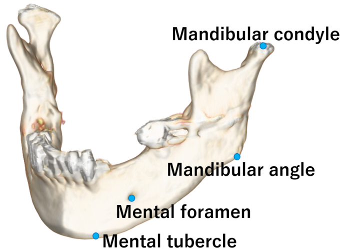


Fig. 2. Anatomical feature points of the mandible. Here, 49-dimensional features were defined using combinations of eight feature points and two cutting points on the left and right.

on anatomical distinctions, specifically C_0: the mandibular ramus, C_1: the midpoint of the chin and C_0, C_2: the symmetry point about the midline of C_3, C_3: the midpoint of the chin and C_5, C_4: the midpoint of C_3 and C_5, and C_5: the mental foramen. Using these cutting planes, we defined eight resection areas for each dataset, specifically (C_0, C_2), (C_0, C_3), (C_0, C_4), (C_0, C_5), (C_1, C_2), (C_1, C_3), (C_1, C_4), and (C_1, C_5). Assuming that the number of fibular segments was between one and three, the appropriate number of fibular segments and their placements were given by an oral surgeon for these 232 cases. Figures 1(b), (c), and (d) show examples of such reconstruction plans.

In this experiment, features were defined with an anatomical name or medical terms that have been used in the clinical literature for the past 10 years. To eliminate the dependence on the coordinate system, each feature is a distance or an angle. As shown in Fig. 2, we defined eight points as the anatomical feature points, specifically the left and right mandibular condyle, mandibular angle, mental foramen, and mental tubercle. The following 49-dimensional features were defined from these feature points as F1-F49:

- the distance from the chin to the line of the mandibular condyle: 1 dimension;
- the anterior angle of the mandible: 1 dimension;
- the posterior angle of the mandible: 2 dimensions;
- the angle based on the mandibular condyle and the mandibular angle: 1 dimension;
- the distance between anatomical feature points: 28 dimensions;
- the distance between the left and right cutting points and anatomical feature points: 16 dimensions.

The distance from the chin to the line of the mandibular condyle was defined as the distance between the forward-most point of the bottom edge of the mandible and the line connecting the left and right mandibular condyles. The anterior angle of mandible was defined as the angle formed by the line connecting the left and right mandibular angles and the mental tubercles. The posterior angle of the mandible was defined as the angle formed by the line connecting the mandibular angle and the mandibular condyle and the line connecting the mandibular angle and the mental tubercle. The angle based on the mandibular condyle and the mandibular angle was defined as the angle formed by the lines connecting the left and right mandibular condyles and mandibular angles. To incorporate more information about the mandibular morphology as features, the distance of all combinations of anatomical feature points (${}_8\mathrm{C}_2 = 28$ dimensions) was used. To represent cutting planes by relative position, the distance between two cutting points on the left and right and the anatomical feature points ($2 \times 8 = 16$ dimensions) was used. Then, we normalized each feature between 0 and 1 because the values of different unit systems were used as features.

B. Outline of the methods

In this study, based on the hypothesis that the number of fibular segments used in mandibular reconstruction can be determined by a low-dimensional feature set T_0, we extended the concept of eLasso [7] and applied it to mandibular reconstruction planning. Using a dataset of plans created by an oral surgeon as the training data, we performed the multi-class classification of one, two, and three fibular segment plans. To extend eLasso from binary classification to multi-class classification, we introduce the importance evaluation criterion, which scores the effect of each feature on the classification based on the weight of the feature calculated by the Lasso.

Figure 3 shows the processing flow of the proposed method. First, for case j, 49-dimensional features are calculated and a feature vector $\boldsymbol{d}^{(j)} = [d_1, d_2, \cdots, d_{49}]^{\mathrm{T}}$ is created. Furthermore, feature matrix $\boldsymbol{D}_f = [\boldsymbol{d}^{(1)}\boldsymbol{d}^{(2)} \cdots \boldsymbol{d}^{(232)}]^{\mathrm{T}}$ is created by concatenating the feature vectors. Next, to perform the multi-class classification of the number of fibular segments, we use the one-versus-rest [9]. When performing three-class classification with one, two, and three fibular segments, case j is labeled using a three-element vector $\boldsymbol{y}^{(j)} = (y_1, y_2, y_3)$ as follows.

- Class 1 (one fibular segment): $\boldsymbol{y}^{(j)} = (1, 0, 0)$
- Class 2 (two fibular segments): $\boldsymbol{y}^{(j)} = (0, 1, 0)$
- Class 3 (three fibular segments): $\boldsymbol{y}^{(j)} = (0, 0, 1)$

Each case is labeled, and then correct answer labels $\boldsymbol{y}_i = [{y_i}^{(1)}{y_i}^{(2)} \cdots {y_i}^{(232)}]^{\mathrm{T}}$ $(i = 1, 2, 3)$ are created by arranging every y_i. The correct answer labels $\boldsymbol{y}_i$ and feature matrix $\boldsymbol{D}_f$ are used as input data. Binary classification of y_i using the Lasso yields regression coefficient $\boldsymbol{\beta}_i$. In this study, feature set T is extracted from these using the importance evaluation criterion E described later. When the Lasso is applied once, only one feature set can be obtained. Therefore, feature matrix $\boldsymbol{D}_f$ is updated based on the extracted T, and the Lasso is performed again. This makes it possible to enumerate multiple versions of T.

C. Feature extraction using eLasso

First, we explain how we perform the binary classification of y_i. We represent the correct answer labels $\boldsymbol{y}_i$ by a weighted sparse linear combination, where $\boldsymbol{\beta}_i = [{\beta_1}^{(i)}{\beta_2}^{(i)} \cdots {\beta_n}^{(i)}]^{\mathrm{T}}$ $(i = 1, 2, 3)$(n : the number of features) denotes the weight vectors, which are calculated by minimizing the following objective function.

$$L(\boldsymbol{\beta}_i) = ||\boldsymbol{y}_i - \boldsymbol{D}_f\boldsymbol{\beta}_i||_2^2 + \lambda||\boldsymbol{\beta}_i||_1 \tag{1}$$

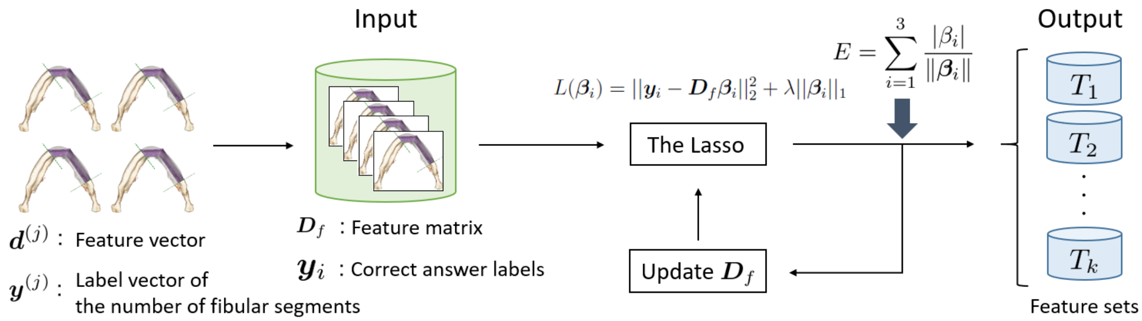


Fig. 3. Outline of the proposed eLasso for extracting important features. The Lasso uses feature matrix $\boldsymbol{D}_f$ and correct answer labels $\boldsymbol{y}_i$ to extract feature set T. Then, feature matrix $\boldsymbol{D}_f$ is updated based on the extracted T, and the Lasso is performed again.

where λ is a parameter to control the sparsity and is called a normalization coefficient. In this study, we used ADMM to calculate $\boldsymbol{\beta}_i^*$ [10]. The inner product of $\boldsymbol{\beta}_i^*$ and $\boldsymbol{d}^{(j)}$ is calculated, and the class corresponding to the largest value $\hat{y}_i^{(j)}$ is the estimated number of fibular segments. For example, the calculation for case j yields $\hat{\boldsymbol{y}}^{(j)}=(0.1, 0.7, 0.3)$, and the estimated number of fibular segments is two because of the maximum $\hat{y}_2^{(j)}$.

In the proposed algorithm, by introducing eLasso, multiple feature sets including features that could not be obtained by the Lasso are enumerated. In the method by Hara et al., feature sets are extracted based on the weight of each feature calculated by the Lasso. However, in the multi-class classification targeted by this study, as shown in the examples in Figs. 4(a), (b), and (c), three weight vectors $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, and $\boldsymbol{\beta}_3$ with different scales were obtained, so the weights of each feature cannot be compared. Therefore, the importance evaluation criterion E is defined by the following equation to quantify the effect of each feature on the multi-class classification of the number of fibular segments.

$$E=\sum_{i=1}^{3}\frac{|\beta_i|}{||\boldsymbol{\beta}_i||} \tag{2}$$

All features are evaluated by Eq. (2). Figure 4(d) shows an example of the importance evaluation criterion E calculated from Figs. 4(a), (b), and (c). It is assumed that a feature with a larger E is a more important feature in the multi-class classification. Thus, features can be compared. Using the importance evaluation criterion E as a guideline for feature extraction extends eLasso to multi-class classification, and a framework for enumerating the feature sets with the top-N importance evaluation criterion is provided below.

STEP 1 For feature matrix $\boldsymbol{D}_f$, weight vector $\boldsymbol{\beta}_i$ and the sum Q of the three objective function values $L(\boldsymbol{\beta}_i)$ are obtained from Eq. (1).

STEP 2 All features are evaluated by Eq. (2), and the feature set T containing the features with the top-N values of E is obtained. Then, $(T, Q, \boldsymbol{D}_f)$ is retained as a solution candidate.

STEP 3 Of the retained solution candidates, the one with the smallest sum of the three objective function values is output as the kth solution.

STEP 4 The following operation is performed for all features t of the output solution in STEP 3.

a) The column corresponding to feature t is removed from feature matrix $\boldsymbol{D}_f$, and a new feature matrix $\boldsymbol{D}_f$ is created.
b) For feature matrix $\boldsymbol{D}_f$, weight vector $\boldsymbol{\beta}_i$ and the sum Q of the three objective function values $L(\boldsymbol{\beta}_i)$ are obtained from Eq. (1).
c) All features are evaluated by Eq. (2), and the feature set T containing the features with the top-N values of E is obtained. Then $(T, Q, \boldsymbol{D}_f)$ is retained as a solution candidate.

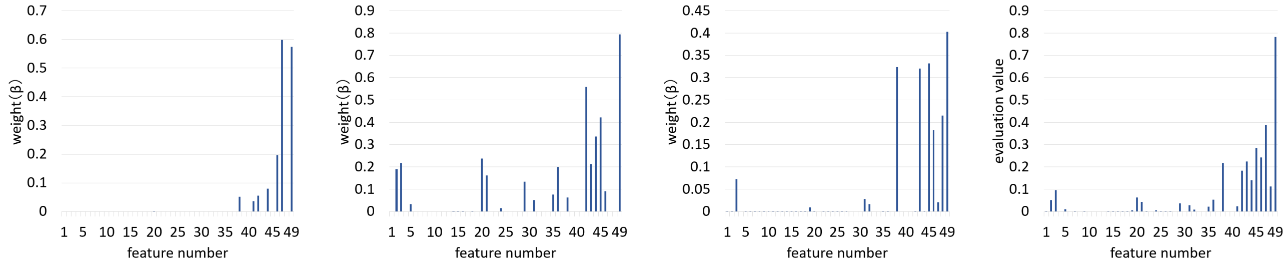
STEP 5 Repeat STEP 3 and STEP 4.

In the proposed method, the Lasso and feature extraction are performed with a new feature matrix that is updated by removing the information of the feature extracted by the Lasso from the feature matrix. By repeating this operation, several important feature sets can be output.

Finally, we describe the estimation method using the obtained N-dimensional feature set T. Creating a new feature matrix $\hat{\boldsymbol{D}}_f$ that contains only the extracted N-dimensional feature set, we calculate $\hat{\boldsymbol{\beta}}_i$ by performing the Lasso again. The number of fibular segments is estimated by calculating the inner product of these weight vectors and the feature vector of the reconstruction plan. The performance of feature set T is estimated by the proportion of estimation results that match the number of fibular segments determined by an oral surgeon, that is, by the accuracy rate.

III. EXPERIMENTS

The purpose of this experiment was to use the proposed method to obtain multiple low-dimensional feature sets with high estimation performance for classifying the number of fibular segments. To evaluate performance, the number of fibular segments was estimated using only the N-dimensional feature sets extracted according to E using the conventional (Lasso-based extraction) method and the proposed (eLasso-based extraction) method. The accuracy rates were then obtained. In the conventional method, the Lasso was applied once, and the top-N feature set was obtained based on the importance evaluation criterion E. Then, the Lasso was performed again using only the obtained feature



(a) (b) (c) (d)

Fig. 4. Examples of weight vectors extracted by the Lasso and the importance evaluation criterion: (a) weight vector $\boldsymbol{\beta}_1$, (b) weight vector $\boldsymbol{\beta}_2$, (c) weight vector $\boldsymbol{\beta}_3$, and (d) importance evaluation criterion E.

TABLE I

ACCURACY RATES FOR EXTRACTED FEATURE SETS OBTAINED BY THE CONVENTIONAL AND PROPOSED METHODS

N	Conventional method	Proposed method
1	58.62%	60.34%(2 sets)
2	45.69%	75.86%(1 sets)
3	64.66%	79.31%(2 sets)
4	79.31%	83.62%(3 sets)
5	79.31%	84.48%(7 sets)
6	82.33%	85.78%(6 sets)
7	82.33%	87.50%(1 sets)

set, and the accuracy rate was calculated. In the proposed method, N-dimensional features were enumerated, but the number of the output solutions was $10,000$. Thus, multiple feature sets were obtained, but the feature set with the highest accuracy rate was selected. This feature set was compared with the feature set obtained by the conventional method. The number of feature dimensions to be extracted was varied as $N = 1$, 2, 3, 4, 5, 6, and 7. For the normalization coefficient, the value with the highest accuracy rate after several trials was used. Specifically, $\lambda = 1.0$ was used for extraction and $\lambda = 10^{-5}$ was used for estimation. We note that the accuracy rate was 87.50% when the estimation was performed using all 49 dimensions.

The accuracy rates obtained by the proposed method and the conventional method for various numbers of extracted feature set dimensions are shown in Table I. For the proposed method, the numbers of feature sets that yielded the highest accuracy rate are also shown. For each value of N, the proposed method obtained feature sets with higher estimation performance than those obtained by the Lasso. In addition, the 7-dimensional feature set obtained by the proposed method achieved the same accuracy rate as that obtained using all 49 dimensions.

IV. CONCLUSIONS

This paper proposed an algorithm based on eLasso for extracting multiple sets of low-dimensional features that are important for determining the number of fibular segments in the mandibular reconstruction. The experiments using a dataset of plans specified by an oral surgeon showed that the extracted 7-dimensional feature set obtained by this method had the same estimation performance as one in which all 49-dimensional features defined using traditional anatomical features were used.

ACKNOWLEDGMENT

This research was supported by JSPS Grant-in-Aid for Scientific Research (B) (19H04484) and Grant-in-Aid for Challenging Exploratory Research (18K19918). We thank Kimberly Moravec, PhD, from Edanz Group (www.edanzediting.com/ac) for editing a draft of this manuscript.

REFERENCES

[1] A. F. Flemming, M. D. Brough, N. D. Evans, H. R. Grant, M. Harris, D. R. James, M. Lawlor, and I. M. Laws, "Mandibular Reconstruction Using Vascularised Fibula," *British Journal of Plastic Surgery*, Vol. 43, No. 4, pp. 403–409, Jul 1990.

[2] J. S. Brown, C. Barry, M. Ho, R. Shaw, "A New Classification for Mandibular Defects after Oncological Resection," *The Lancet Oncology*, Vol. 17, No. 1, pp. e23–e30, Jan 2016.

[3] M. Nakao, S. Aso, Y. Imai, N. Ueda, T. Hatanaka, M. Shiba, T. Kirita, and T. Matsuda, "Automated Planning with Multivariate Shape Descriptors for Fibular Transfer in Mandibular Reconstruction," *IEEE Transactions on Biomedical Engineering*, Vol. 64, No. 8, pp. 1772–1785, Aug 2017.

[4] M. Nakao, M. Hosokawa, Y. Imai, N. Ueda, T. Hatanaka, T. Kirita, and T. Matsuda, "Volumetric Fibular Transfer Planning with Shape-based Indicators in Mandibular Reconstruction," *IEEE Journal of Biomedical and Informatics*, Vol. 19, No. 2, pp. 581–589, Mar 2015.

[5] R. Kawasaki, M. Nakao, Y. Imai, N. Ueda, T. Hatanaka, M. Shiba, T. Kirita, and T. Matsuda, "Sparse Shape Model for Fibular Transfer Planning in Mandibular Reconstruction," in *38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2016, pp. 2508–2511.

[6] R. Tibshirani, "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Vol. 58, No. 1, pp. 267–288, 1996.

[7] S. Hara and T. Maehara, "Enumerate Lasso Solutions for Feature Selection," in *31st AAAI Conference on Artificial Intelligence*, 2017.

[8] M. Nakao, S. Aso, Y. Imai, N. Ueda, T. Hatanaka, M. Shiba, T. Kirita, and T. Matsuda, "Statistical Analysis of Interactive Surgical Planning Using Shape Descriptors in Mandibular Reconstruction with Fibular Segments," PLoS One, Vol. 11, No. 9, pp. e0161524, Sep 2016.

[9] J. H. Xu, "Constructing a fast algorithm for multi-label classification with support vector data description," in *IEEE Conference on Granular Computing (GrC2010)*, San Jose, CA, USA, 2010, pp. 817–821.

[10] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers," *Foundations and Trends in Machine Learning*, Vol. 3, No. 1, pp. 1–122, 2011.