



401 N. Lindbergh Blvd.
St. Louis, MO 63141
Tel.: 314.993.1700, #546
Toll Free: 800.424.2841, #546
Fax: 800.708.1364

Send via email to: jbode@aaortho.org and cyoung@aaortho.org

AAO Foundation Final Report Form
(a/o 5/30/2022)

Please prepare a report that addresses the following:

Type of Award, e.g., Orthodontic Faculty Development Fellowship Award, Postdoctoral Fellowship Award, Biomedical Research Award, Center Award, Educational Innovation Award, Program Award, Research Aid Award

Name(s) of Principal Investigator(s) Lucia Cevidanes

Institution University of Michigan

Title of Project Decision Support System for Early Diagnosis of Osteoarthritis of the Temporomandibular Joints

Period of AAOF Support (e.g. 07-01-2021 to 06-30-2022): 07-01-20 to 6-30-22

Amount of Funding \$30,000.00

Summary/Abstract

Aim 1. We designed and implemented automated image processing and machine learning models for segmenting the temporomandibular joints. We designed a semi-supervised learning algorithm called active deep learning active contour for automatic image segmentation of large datasets. Labeled available 3D CBCT high resolution scans of 200 right and left TMJs from 100 patients were used for validation and training of algorithms for automatic segmentation. The image processing algorithms accurately initiate image "labeling" to create data for training a generic machine learning method for segmentation now available for interested clinicians and researchers.

Aim 2. We designed a feature selection method that identifies a combination of features for classification and prediction of TMJ OA progression. The prospectively collected data sets consisted of 100 condyles from 50 asymptomatic subjects and 100 condyles from 50 TMJ OA patients. Accounting for inter-features mutual information, we determined the most correlated

biomarkers for diagnosis and prediction progression. While clinical pain-related markers remain strong predictors of disease progression, the data science approaches in this AAOF BRA proposal scaled up our previous integrative indicators of TMJ OA including levels of proteins in blood and saliva as well as TMJ condyles and articular fossa radiomic markers. The final trained machine learning models were deployed in a cutting-edge clinical Decision Support System hosted in the Dental Storage for Computation and Integration, an open-source cloud-based solution for clinicians and researchers.

Detailed results and inferences:

1. If the work has been published please attach a pdf of manuscript OR Please see PDFs of the manuscripts for Aim1: one published and one currently under review, submitted in August for publication at the Journal of Medical Systems; and Aim 2 manuscript currently under review, submitted in August to SPIE Medical Imaging- Journal of Medical Imaging.

Respond to the following questions:

1. Were the original, specific aims of the proposal realized? Yes.
2. Were the results published?
 - a. If so, cite reference/s for publication/s including titles, dates, author or co-authors, journal, issue and page numbers-

Le C, Deleat-Besson R, Prieto J, Brosset S, Dumont M, Zhang W, Cevidanes L, Bianchi J, Ruellas A, Gomes L, Gurgel M, Massaro C, Aliaga-Del Castillo A, Yatabe M, Benavides E, Soki F, Al Turkestani N, Evangelista K, Goncalves J, Valladares-Neto J, Alves Garcia Silva M, Chaves C, Costa F, Garib D, Oh H, Gryak J, Styner M, Fillion-Robin JC, Paniagua B, Najarian K, Soroushmehr R. Automatic Segmentation of Mandibular Ramus and Condyles. Annu Int Conf IEEE Eng Med Biol Soc. 2021 Nov;2021:2952-2955. doi: 10.1109/EMBC46164.2021.9630727.

The other 2 manuscripts are currently under review and the PDFs are included with this report.

- b. Was AAOF support acknowledged? Yes.
 - c. If not, are there plans to publish? If not, why not? Not applicable.
3. Have the results of this proposal been presented?
 - a. If so, list titles, author or co-authors of these presentation/s, year and locations
 - * SPIE Medical Imaging, San Diego , February 19 – 23, 2023
 - * AADOCR, Portland, OR, March 15-18, 2023
 - b. Was AAOF support acknowledged? Yes
 - c. If not, are there plans to do so? If not, why not? Not applicable.

4. To what extent have you used, or how do you intend to use, AAOF funding to further your career?

The AAOF award has provided broad and exciting new frontiers for translational research in the field of TMJ health and disease. This Biomedical Research Award will lead to continued investigations to better understand bone repair and remodeling in the TMJ.

Accounting for Project; (i.e.), any leftover funds, etc. No left over funds.



HHS Public Access

Author manuscript

Annu Int Conf IEEE Eng Med Biol Soc. Author manuscript; available in PMC 2022 April 09.

Published in final edited form as:

Annu Int Conf IEEE Eng Med Biol Soc. 2021 November ; 2021: 2952–2955. doi:10.1109/EMBC46164.2021.9630727.

Automatic Segmentation of Mandibular Ramus and Condyles

Celia Le¹, Romain Deleat-Besson¹, Juan Prieto², Serge Brosset¹, Maxime Dumont¹, Winston Zhang¹, Lucia Cevitanes¹, Jonas Bianchi³, Antonio Ruellas^{1,4}, Liliane Gomes⁵, Marcela Gurgel⁶, Camila Massaro⁷, Aron Aliaga-Del Castillo⁷, Marilia Yatabe¹, Erika Benavides¹, Fabiana Soki¹, Najla Al Turkestani¹, Karine Evangelista⁸, Joao Goncalves⁵, Jose Valladares-Neto⁸, Maria Alves Garcia Silva⁸, Cauby Chaves Jr⁶, Fábio Costa⁶, Daniela Garib⁷, Heesoo Oh³, Jonathan Gryak⁹, Martin Styner², Jean-Christophe Fillion-Robin¹⁰, Beatriz Paniagua¹⁰, Kayvan Najarian⁹, Reza Soroushmehr⁹

¹School of Dentistry, University of Michigan, Ann Arbor, MI 48109

²Psychiatry Department, University of North Carolina, Chapel Hill, NC

³Department of Orthodontics, University of the Pacific, Arthur A. Dugoni, School of Dentistry, San Francisco, CA

⁴Department of Orthodontics, Federal University of Rio de Janeiro, Brazil

⁵Department of Orthodontics, São Paulo State University, Araraquara, Brazil

⁶Department of Orthodontics, Federal University of Ceara, Fortaleza, Brazil

⁷Department of Orthodontics, University of São Paulo, Bauru, Brazil

⁸Department of Orthodontics, Federal University of Goias, Goiania, Brazil

⁹Department of Computational Medicine and Bioinformatics, Univ. of Michigan, Ann Arbor

¹⁰Kitware Inc., Chapel Hill, NC

Abstract

In order to diagnose TMJ pathologies, we developed and tested a novel algorithm, MandSeg, that combines image processing and machine learning approaches for automatically segmenting the mandibular condyles and ramus. A deep neural network based on the U-Net architecture was trained for this task, using 109 cone-beam computed tomography (CBCT) scans. The ground truth label maps were manually segmented by clinicians. The U-Net takes 2D slices extracted from the 3D volumetric images. All the 3D scans were cropped depending on their size in order to keep only the mandibular region of interest. The same anatomic cropping region was used for every scan in the dataset. The scans were acquired at different centers with different resolutions. Therefore, we resized all scans to 512×512 in the pre-processing step where we also performed contrast adjustment as the original scans had low contrast. After the pre-processing, around 350 slices were extracted from each scan, and used to train the U-Net model. For the cross-validation, the dataset was divided into 10 folds. The training was performed with 60 epochs, a batch size of 8 and a learning rate of 2×10^{-5} . The average performance of the models on the test set

presented 0.95 ± 0.05 AUC, 0.93 ± 0.06 sensitivity, 0.9998 ± 0.0001 specificity, 0.9996 ± 0.0003 accuracy, and 0.91 ± 0.03 F1 score. This study findings suggest that fast and efficient CBCT image segmentation of the mandibular condyles and ramus from different clinical data sets and centers can be analyzed effectively. Future studies can now extract radiomic and imaging features as potentially relevant objective diagnostic criteria for TMJ pathologies, such as osteoarthritis (OA). The proposed segmentation will allow large datasets to be analyzed more efficiently for disease classification.

I. INTRODUCTION

Osteoarthritis (OA) is a top cause of chronic disability, and with aging, the disease progresses to considerable structural and functional alterations in the joint. If the condition is detected earlier, treatment can prevent the large joint destruction; however, there is a lack of studies focusing on the early diagnosis [1–3]. There is no cure for OA, and current treatments attempt to reduce pain and improve function by slowing disease progression. The Temporomandibular joints (TMJ) are small joints that connect the lower jaw (mandible) to the skull. After chronic low back pain, TMJ disorders (TMD) are the second most commonly occurring musculoskeletal conditions, resulting in pain and disability, with an annual cost estimated at \$4 billion [4].

The recommended Diagnostic Criteria for TMD protocol [5] include clinical and imaging diagnostic criteria for differentiating health and disease status, and recent studies have indicated the biological markers may also improve the diagnostic sensitivity and specificity [6]. However, feature extraction from Cone-Beam Computed Tomography (CBCT) images remains time consuming before this integrative model can be applied in larger scale studies.

There are some commercial or open-source tools such as ITK-SNAP [7] and 3D-Slicer [8] that clinicians use to interactively segment condyles in each individual image at a time and calculate some parameters of images. However, this process is time-consuming and challenging for clinicians due to low signal/noise ratio of the large field of view CBCT images commonly used in dentistry [9]. Therefore, our goal is to develop a method to automatically segment the mandibular ramus. More efficient and reproducible mandibular segmentation will help clinicians extract features from the mandibular condyles and ramus, analyze changes in the shape and anatomy of the condyles over time to properly diagnose the disease, as well as plan the anatomy for surgical interventions. This would facilitate the study of the TMJ OA and could help prevent the disease progression and predict the disease at early stages.

Manual, user interactive or semi-automatic methods use different imaging modalities such as magnetic resonance (MR) imaging, computed tomography (CT), cone-beam computed tomography (CBCT), ultrasonography, and conventional radiography [10–17] to segment the mandibular condyle and ramus with applications for TMJ and dentofacial treatment planning and assessment of outcomes. Up to date, automatic segmentation tools for condylar and thin bone cortical areas of the mandibular ramus have been limited to high resolution CBCT images [18] or small sample size acquired with the same scanning protocol [19]. The algorithm presented in this paper aimed to create a fully automated method to segment

the ramus and condyle out of large field CBCT scans of the head from 4 different clinical centers and scanning acquisition protocols. The dataset is presented in Section II and the different steps of the proposed method are explained in Section III. We then show the experimental results of the proposed method and compare them with condyles manually segmented by clinician experts. Finally, conclusion remarks are presented in Section IV.

II. DATASET

We used de-identified datasets from the University of Michigan, State University of Sao Paulo, Federal University of Goias and Federal University of Ceara, that consisted of 3D large field of view scans CBCT scans of the head of 109 patients. At the different clinical centers, the images were acquired with different scanners, spatial resolutions varying from 0.2 to 0.4mm³ voxels, and image acquisition protocols.

The dataset used in this study contains both patients with radiographic diagnosis of osteoarthritis and healthy condyles. The inclusion of both OA and non-OA patients in the dataset helps develop a more generalizable segmentation model across healthy and diseased patients. The images were first interactively segmented by clinicians using ITK-SNAP (3.8.0) or 3D Sheer (4.11). These segmentations were used as ground-truth to train and evaluate the performance of the proposed method.

III. Proposed method and Experimental results

The proposed method developed to segment the mandibular condyles and ramus out of CBCT scans is based on image processing and machine learning approaches that are summarized in the flowchart shown in Figure 1.

We first describe image pre-processing to deal with the quality of the images and region of interest. After that, we explain the machine learning techniques used to segment the mandibular condyles and ramus and to detect its contours out of the craniofacial structures. After identification of the mandibular condyles and ramus contours, we perform post-processing for artifact removal and improvement the segmentations quality.

A. Pre-processing

Figure 2 shows an example of a cross-sectional image from a raw large field of view CBCT scan with the mandibular ramus and condyles on each side of the image. The head large field CBCT scans were low contrast images, therefore we adjusted the contrast to improve the training of our deep learning model and help it to make a better prediction. We performed slice cropping according to the number of slices in each scan to keep only the region of interest where the condyles are in the large field of view scans. The algorithm selected the same anatomic cropping region for every 3D scan in the dataset, then split it into 2D cross-sections, and every cross-section was resized to 512 × 512 pixels to standardize the dataset. Each CBCT scan resulted in 300-400 cross-sectional images after the pre-processing, depending on the number of slices composing the scan, which variates with the acquisition protocol used.

Figure 3 shows an example of a CBCT image after pre-processing. The output of the pre-processing is used in the next step where we train our deep learning model.

B. U-Net training

We used the images obtained from the pre-processing to train a U-Net model. This network was first developed for biomedical image segmentation and later utilized in other applications, such as field boundary extraction from satellite images [20].

We split the dataset into 2 parts: 90 patients CBCT scans for training (approximately 80% of the total dataset) and 19 patients CBCT scans for testing (approximately 20% of the total dataset). We performed a 10 folds cross-validation on the training set and used the testing set to evaluate the model performances. Each fold of the cross-validation contained the cross-sectional images from 9 scans. We equally distributed the scans into the different folds according to the acquisition center, to avoid the overfitting of the model.

The models were trained during 60 epochs to ensure that the model would converge, with a batch size of 8, due to computer performance limitations, and a learning rate of 2×10^{-5} , to be able to determine with precision the most appropriate epoch. We used Tensorboard to measure and visualize the loss and accuracy of the model and selected the epoch of the model before it overfitted.

We gave the high-contrast cross-sectional images from the testing dataset to every trained model for them to predict a segmentation of the condyles for every image.

C. Post-processing

The post-processing consisted in binarizing the output images coming from the U-Net model using a threshold based on Otsu's method, resize them to their original size, and adding them to reconstitute the original 3D scan. We then calculated the volume of each component on the 3D image, and used a volumetric threshold depending on the size of the image to remove small objects (artefacts) that are not part of the condyle.

The performance of the proposed segmentation method was evaluated by comparing the output of the method to the ground truth, scans manually segmented by clinicians.

Figure 4 shows both the manual segmentation by the clinicians and the automatic segmentations output by our algorithm.

We used Area Under the Receiver Operating Characteristic Curve (AUC), F1 score, accuracy, sensitivity and specificity to quantify the precision of the models. These measurements vary from zero to one, where zero means no superposition between the two volumes, and one shows a perfect superposition between both. They were performed on the binarized 3D images resulting from the post-processing.

The results we obtained for the validation dataset and the testing dataset are summarized in the following tables.

The average measurements of the AUC, F1 Score, accuracy, sensitivity and specificity of the testing dataset for the 10 folds of the cross-validation were each above 0.9 as shown in Tables 1 and 2, which demonstrates the precision of the automatic segmentations compared to the ground truth interactive segmentations. Additionally, the standard deviations were quite low, indicating that the automatic segmentations were very consistent and generalizable to unseen patients.

We selected the trained model presenting the highest F1 score when evaluating the model on the test dataset and used it to deploy the validated algorithm as a docker container, called MandSeg, in an open-source data management system, the Data Storage for Computation and Integration (DSCI) [21], that allows clinicians and researchers to access a secure user interface to compute automated segmentations for their patients or study datasets.

IV. CONCLUSION AND FUTURE WORK

The MandSeg algorithm produces accurate automated mandibular ramus and condyles segmentation compared to the ground truth interactive segmentation. Such an efficient automatic mandibular segmentation of CBCT scans will help clinicians early diagnose and predict TMJ disease progression by extracting imaging features of the condyle scans. We expect that the fully automated mandibular ramus and condyles segmentation algorithm presented in this study will improve accuracy in the classification of degeneration in the TMJs even when using the low-resolution large field of view CBCT images that are conventionally taken for jaw surgery planning.

The current dataset is only composed of 109 scans, coming from 4 different clinical centers and the trained models utilized segmentations of only the condyles and ramus, which are the most challenging mandibular areas to segment due to the thinness of the cortical bone in those anatomic regions. Our future objectives include the addition of scans from other clinical centers, training new deep learning models with segmentations of the full mandibles, and integration of the resulting automatic segmentations with other imaging modalities such as digital dental models for clinical applications in dentistry (Figures 5 and 6).

Acknowledgments

Grant supported by NIDCR DEO24450 and 2020 AAOF BRA award

REFERENCES

- [1]. "CDC - arthritis - data and statistics." [Online]. Available: https://www.cdc.gov/arthritis/data_statistics/index.htm
- [2]. Kalladka M, Qnek S, Heir G, Eliav E, Mupparapu M, and Viswanath A, "Temporomandibular joint osteoarthritis: diagnosis and long-term conservative management: a topic review," *The Journal of Indian Prosthodontic Society*, vol. 14, no. 1, pp. 6–15, 2014. [PubMed: 24604992]
- [3]. Bianchi J, et al. "Quantitative bone imaging biomarkers to diagnose temporomandibular joint osteoarthritis." *International Journal of Oral and Maxillofacial Surgery* 50.2 (2021): 227–235. [PubMed: 32605824]
- [4]. National institute of dental and craniofacial research - facial pain statistics. Available: <https://www.nidcr.nih.gov/research/data-statistics/facial-pain>

- [5]. Schiffman E et al. Diagnostic Criteria for Temporomandibular Disorders (DC/TMD) for Clinical and Research Applications: recommendations of the International RDC/TMD Consortium Network and Orofacial Pain Special Interest Group. *J. oral facial pain headache*. 2014. Vol.1, 6–27.
- [6]. Bianchi J et al. Osteoarthritis of the Temporomandibular Joint can be diagnosed earlier using biomarkers and machine learning. *Sci. Rep*2020, no.1, 8012.
- [7]. Yushkevich PA, Gao Y, and Gerig G, “ITK SNAP: An interactive tool for semi-automatic segmentation of multi-modality biomedical images,” in 2016 38th Annual Int. Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2016, pp. 3342–3345.
- [8]. Pieper M Halle, and Kikinis R, “3D shear,” in 2004 2nd IEEE International Symposium on Biomedical Imaging: Nano to Macro (IEEE Cat No. 04EX821). IEEE. 2004. pp. 632–635.
- [9]. Cevdanes LHS, Styner MA, Proffit Image analysis WR and superimposition of 3-dimensional cone-beam computed tomography models. *American journal of orthodontics and dentofacial orthopedics*. 2006. 129(5). 611–618S. [PubMed: 16679201]
- [10]. Talmaceanu D, Lenghel LM, Bolog N, Hedesiu M, Buduru S, Rotar H, Baciut M, and Baciut G, “Imaging modalities for temporomandibular joint disorders: an update,” *Clujul Medical*, vol. 91, no. 3, p. 280. 2018. [PubMed: 30093805]
- [11]. Verhelst P, Verstraete L, Shaheen E, Shujaat S, Darche V, Jacobs R, Swennen G, and Politis C, “Three-dimensional cone beam computed tomography analysis protocols for condylar remodelling following orthognathic surgery: a systematic review,” *International Journal of Oral and Maxillofacial Surgery*. 2020; 49(2):207–217. [PubMed: 31221473]
- [12]. Smirg O, Liberia O, and Smekal Z, “Segmentation and creating 3D model of temporomandibular condyle,” in 2015 38th International Conference on Telecommunications and Signal Processing (TSP). IEEE, 2015. pp. 729–734.
- [13]. Méndez-Mani3n I, Haas OL Jr, Guijarro-Mart3n R, de Oliveira RB, Valls-Onta3n A, and Hern3ndez-Alfaro F, Semi-automated three-dimensional condylar reconstruction,” *Journal of Craniofacial Surgery*. 30(8): 2555–2559. 2019.
- [14]. Xi T, Schreurs R, Heerink WJ, Berge SJ, and Maal TJ, “A novel region-growing based semi-automatic segmentation protocol for three-dimensional condylar reconstruction using cone beam computed tomography (CBCT),” *PloS one*. vol. 9, no. 11, p. e111126, 2014. [PubMed: 25401954]
- [15]. Ma R.-h. et al. Quantitative assessment of condyle positional changes before and after orthognathic surgery based on fused 3d images from cone beam computed tomography,” *Clinical Oral Investigations*, pp. 1–10, 2020 Aug; 24(8):2663–2672. [PubMed: 31728734]
- [16]. Koç A, Sezgin 3S and Kayipmaz S, “Comparing different planimetric methods on volumetric estimations by using cone beam computed tomography,” *La Radiologia Medica*. pp. 1–8, 2020. [PubMed: 31562581]
- [17]. Rasband WS ImageJ. (U. S. National Institutes of Health, 1997–2011) <http://imagej.nih.gov/ij/>
- [18]. Brosset S et al. 3D Auto-Segmentation of Mandibular Condyles. *Annu Int Conf IEEE Eng Med Biol Soc*. 2020:1270–12
- [19]. Fan Y et al. Marker-based watershed transform method for hilly automatic mandibular segmentation from CBCT images. *Dentomaxillofac Radiol*. 2019 Feb;48(2):20180261. [PubMed: 30379569]
- [20]. Waldner F and Diakogiannis FI, 2020. Deep learning on edge: extracting field boundaries from satellite images with a convolutional neural network. *Remote Sensing of Environment*, 245, p. 11174
- [21]. DSCI. Available: <https://dsci.dent.umich.edu>

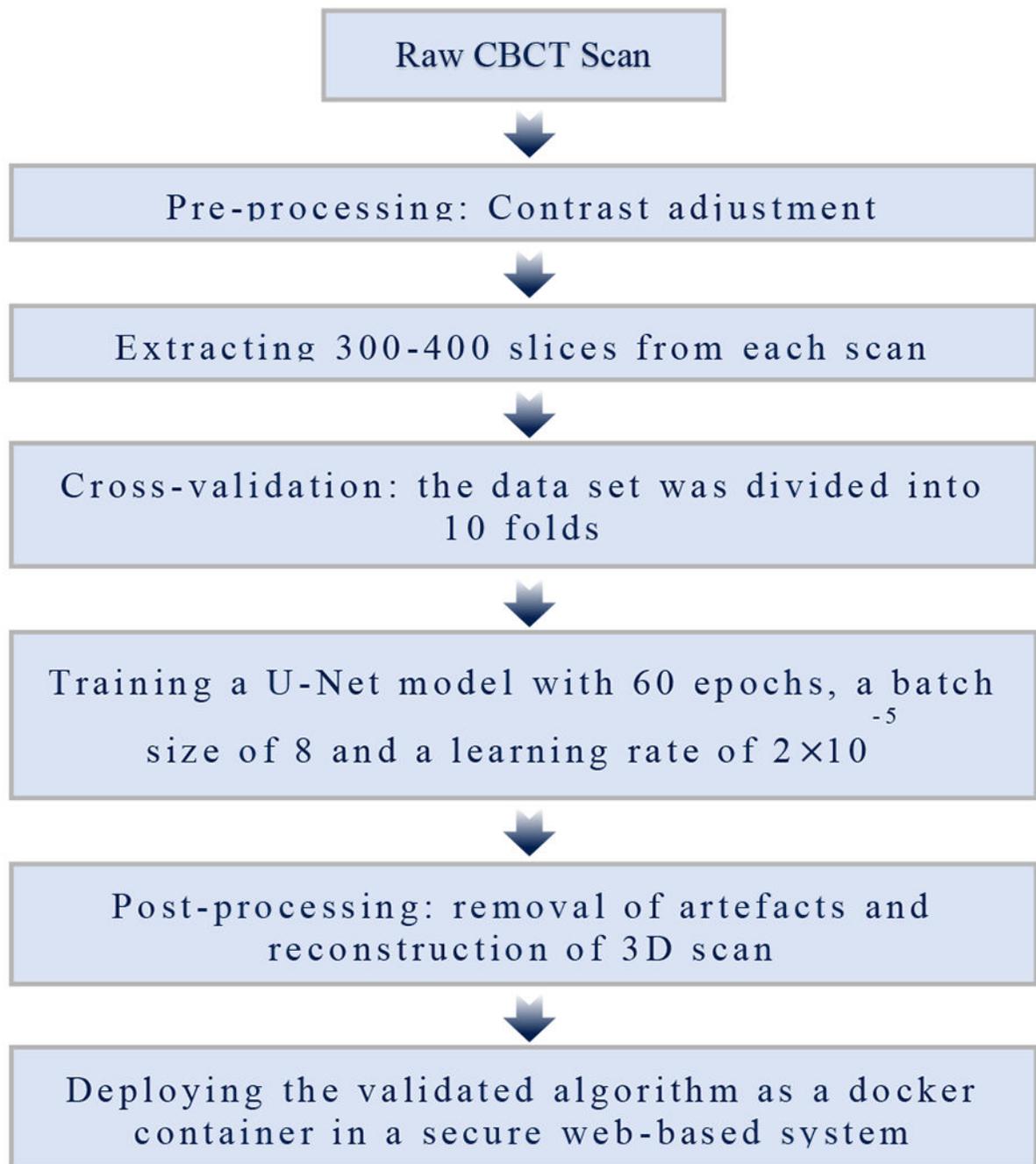


Figure 1 -
Schematic diagram of the proposed method

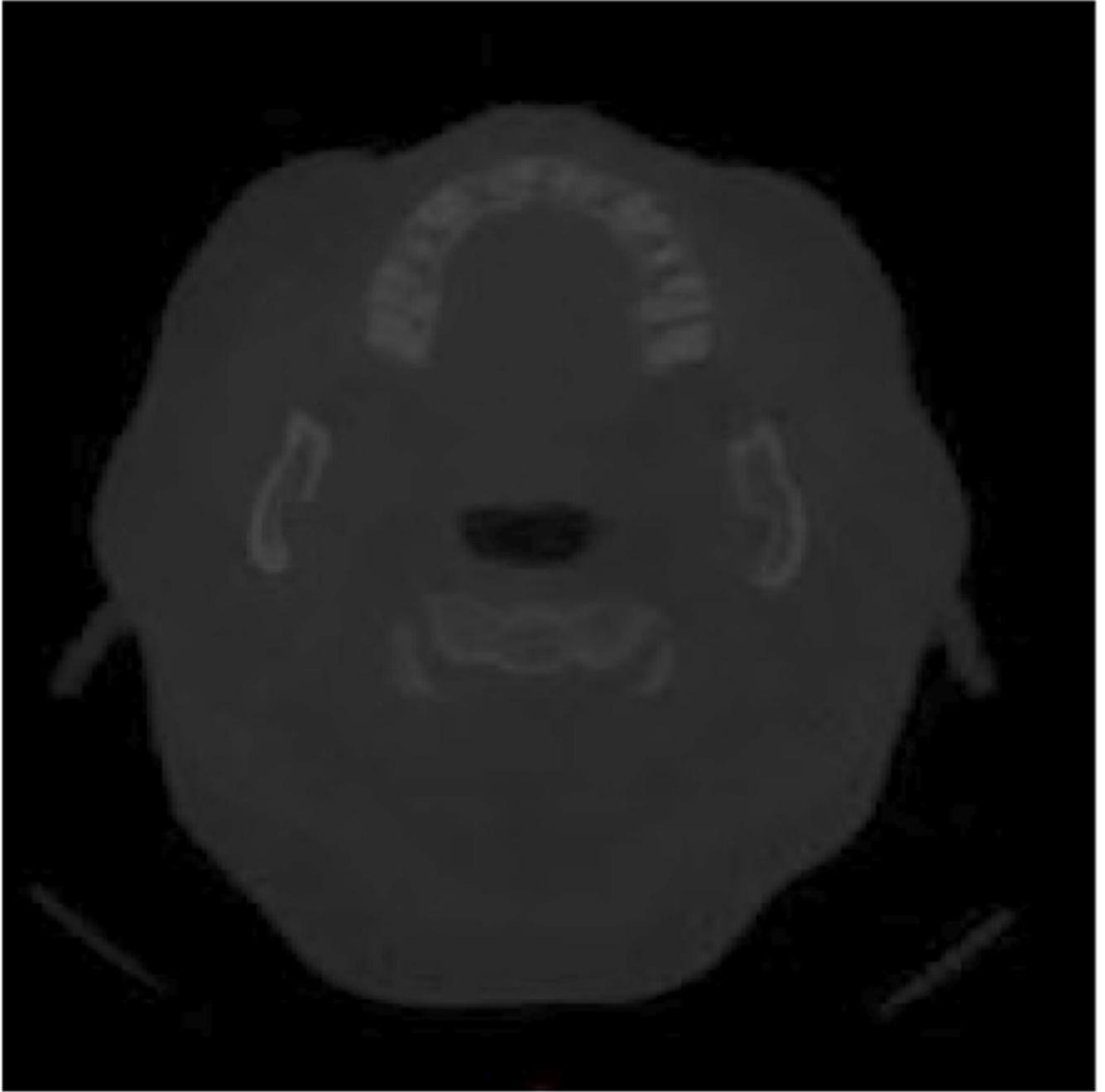


Figure 2 -
An example of one raw CBCT image

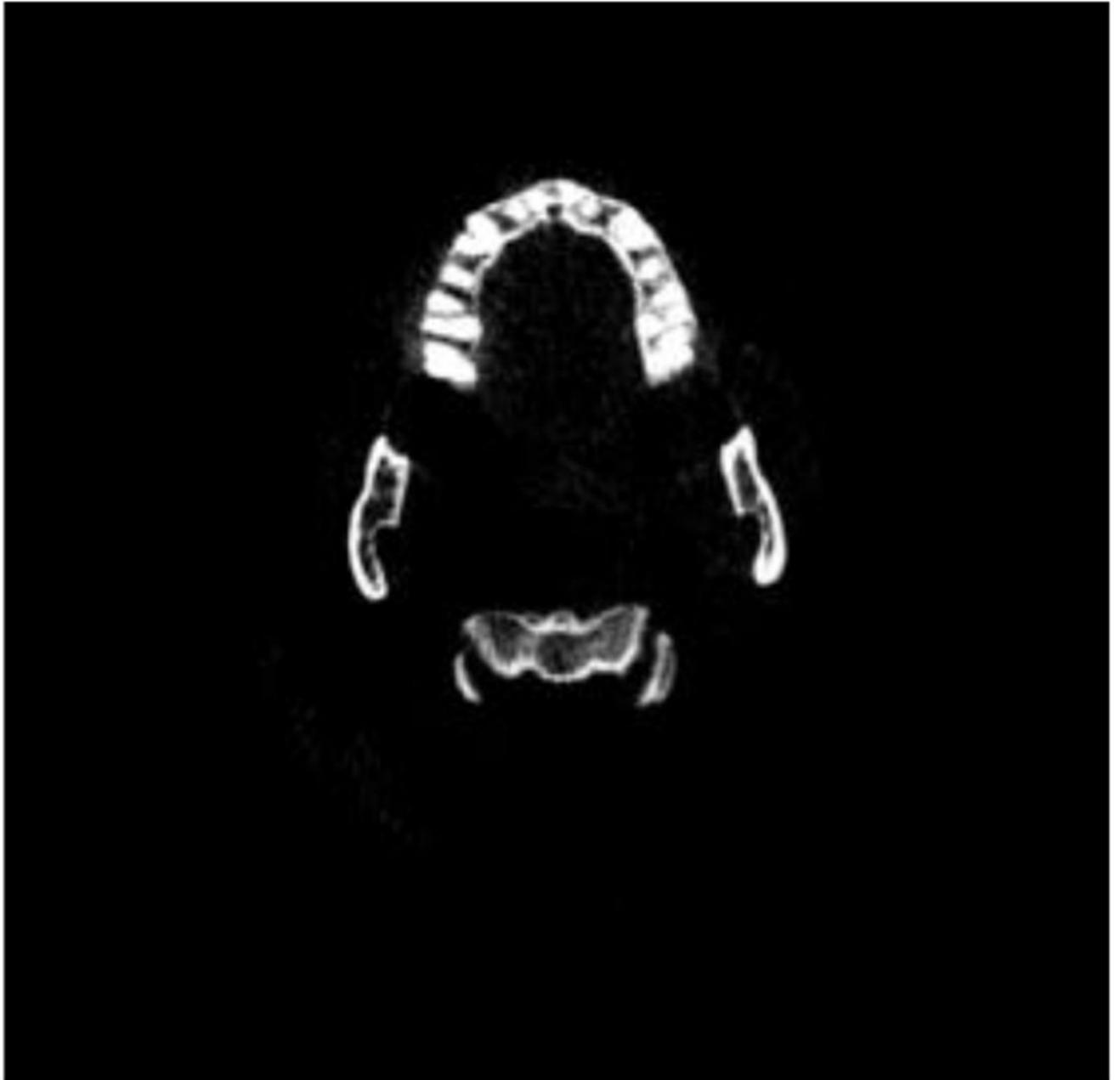


Figure 3 -
Scan after pre-processing

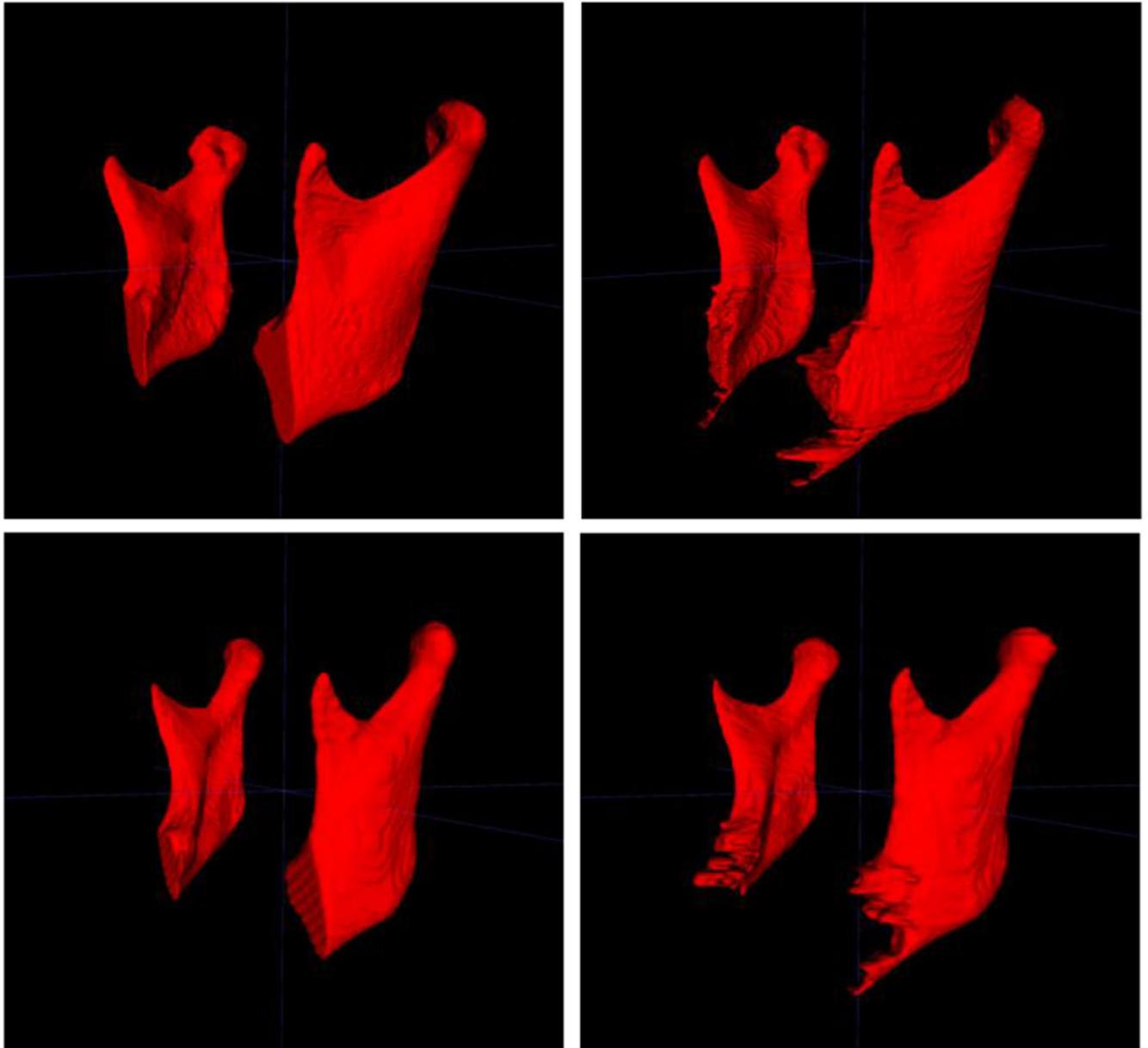


Figure 4 -
Comparison of manual segmentations (left) and automatic segmentations (right), for two different cases

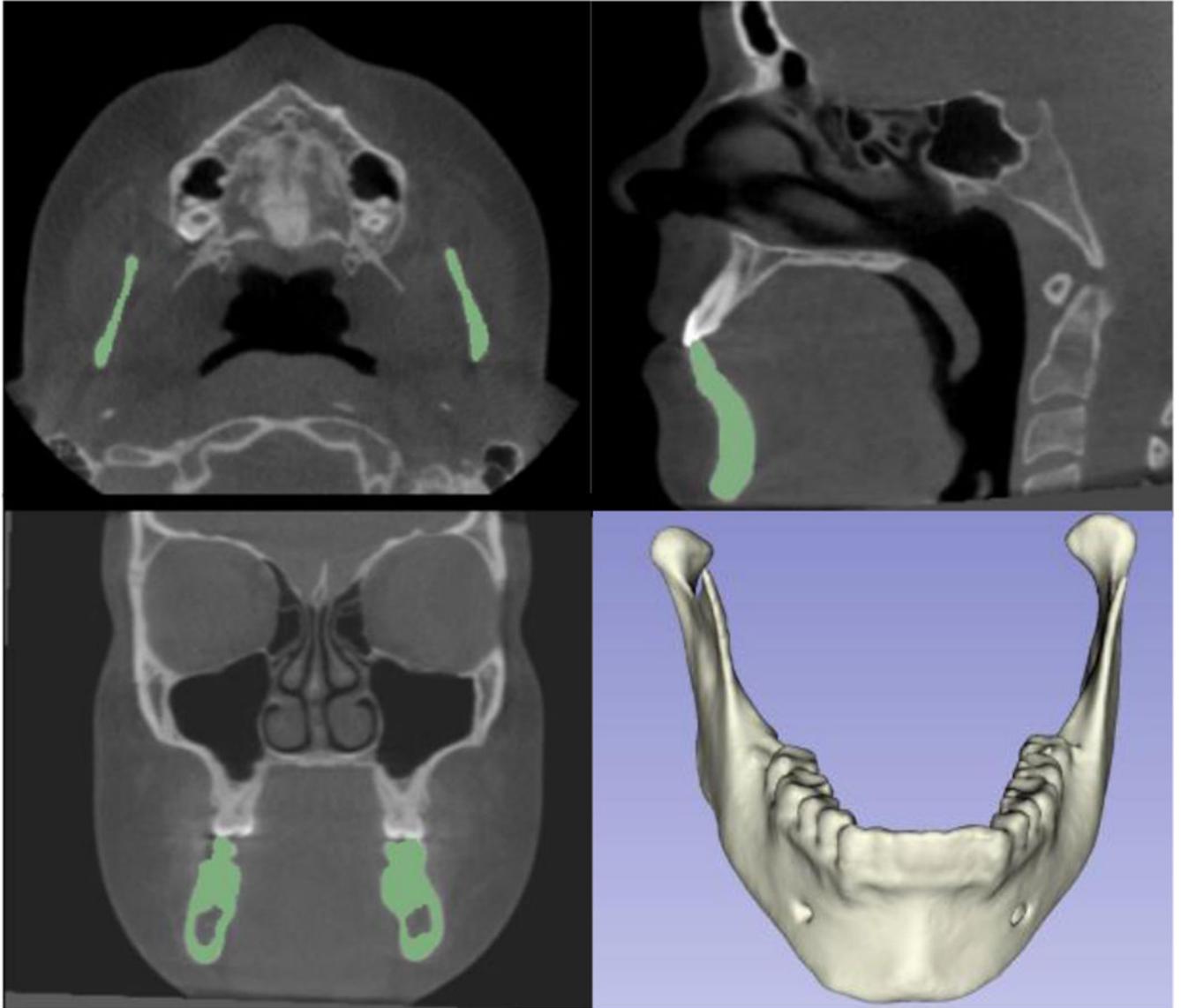


Figure 5 -
Additional training with datasets from other clinical centers for automatic segmentation of the lower jaw (full mandible)

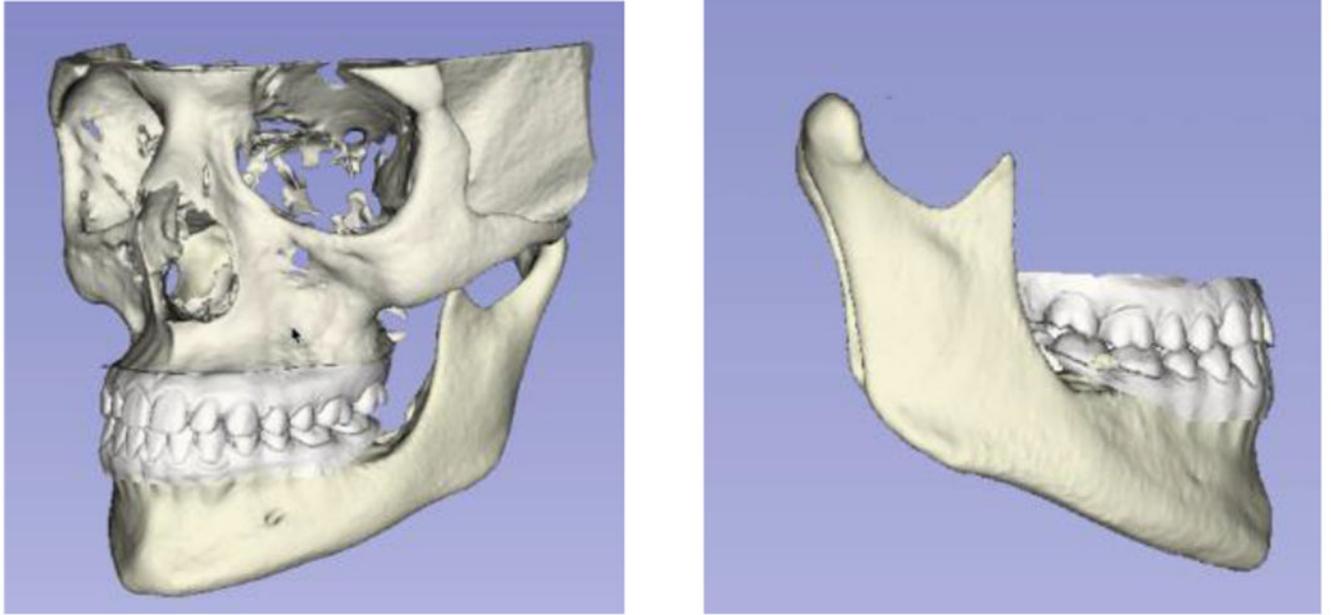


Figure 6 -
Integration of lower jaw automatic segmentation with digital dental models for decision support systems in dentistry

Table 1 -

AUC, F1 Score, accuracy, sensitivity and specificity of the validation dataset

Validation dataset	AUC	F1 score	Sensitivity	Specificity	Accuracy
Average	0.955	0.907	0.923	0.9998	0.9996
Standard deviation	0.040	0.045	0.065	0.0002	0.0003

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2 -

AUC, F1 Score, accuracy, sensitivity and specificity of the test dataset

Test dataset	AUC	F1 score	Sensitivity	Specificity	Accuracy
Average	0.954	0.915	0.926	0.9998	0.9996
Standard deviation	0.051	0.031	0.057	0.0001	0.0003

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Osteoarthritis Diagnosis Integrating Whole Joint Radiomics and Clinical Features for Robust Learning Models using Biological Privileged Information

Najla Al Turkestani^{1,2†}, Lingrui Cai^{3†}, Lucia Cevidanes¹, Jonas Bianchi⁴, Winston Zhang³, Marcela Gurgel¹, Maxime Gillot¹, Baptiste Baquero¹, Kayvan Najarian³ and Reza Soroushmehr³

¹Department of Orthodontics and Pediatric Dentistry, University of Michigan, 1011 North University Ave, Ann Arbor, 48109, Michigan, United States.

²Department of Restorative and Aesthetic Dentistry, Faculty of Dentistry, King Abdulaziz University, Jeddah, 22252, Saudi Arabia.

³Department of Computational Medicine and Bioinformatics, University of Michigan, 100 Washtenaw Ave, Ann Arbor, 48109, Michigan, United States.

⁴Department of Orthodontics, University of the Pacific, Arthur A. Dugoni School of Dentistry, 155 5th St, San Francisco, 94103, CA, United States.

Contributing authors:

†These authors contributed equally to this work.

Abstract

This paper proposes a machine learning model using privileged information (LUPI) and normalized mutual information feature selection method (NMIFS) to build a robust and accurate framework to diagnose patients with Temporomandibular Joint Osteoarthritis (TMJ OA). To build such a model, we employ clinical, quantitative imaging and additional biological markers as privileged information. We show that clinical features play a leading role in the TMJ OA diagnosis and quantitative imaging features, extracted from cone-beam computerized tomography (CBCT) scans, improve the model performance. As the proposed LUPI model employs biological data in the training phase (which boosted the model performance), this data is unnecessary for the testing stage, indicating the model can be widely used even when only clinical and imaging data are collected. The model was validated using 5-fold stratified cross-validation with hyperparameter tuning to avoid the bias of data splitting. Our method achieved an AUC, specificity and precision of 0.81, 0.79 and 0.77, respectively.

Keywords: Temporomandibular joint, Osteoarthritis, Machine learning, Feature selection, Learning using privileged information

1 Introduction

Osteoarthritis (OA) of the temporomandibular joint (TMJ) is a chronic, degenerative disease that affects articular cartilage, synovial tissue and osseous structures of the condyle, articular eminence and articular fossa [1]. It causes chronic pain, jaw dysfunction, deterioration of the quality of life and, in advanced stages, necessitates joint replacement [2, 3]. Current diagnosis of TMJ OA occurs primarily at moderate-severe stage of the disease, following the protocols of the diagnostic criteria for temporomandibular disorders (DC/TMD) [4, 5]. Although various therapeutic measures can relieve disease symptoms at these stages, to date, no treatment modality can cure or reverse degenerative changes within the joint tissues [6]. Hence, identification of diagnostic biomarkers that reflect early pathological changes of the joint is crucial for prevention of the irreversible sequelae of the disease [7].

Animal studies indicated that microstructural change of the subchondral bone was essential for the initiation and progression of OA [8]. However, no robust tools were available to assess these changes, in humans, at early stages of the disease. More recently, advancement of image processing/analysis and high-performance computing techniques allowed extracting quantitative imaging features, i.e., radiomics, which reflect subtle changes within the examined tissues [9]. Along with radiomics, the level of biochemical markers in saliva or blood samples could reflect incipient pathological changes and improve diagnosis, severity assessment and risk of progression of osteoarthritis [10, 11]. The potential of radiomics and biochemical markers has been elucidated in early detection of various diseases, including knee OA; nevertheless, their value in TMJ OA diagnosis has been scarcely investigated [8, 12–16]. Our preliminary studies [17, 18], showed a significant difference in radiomics at the condyles' subchondral bone in TMJ OA and control subjects. We also found a correlation between the resorptive/anabolic changes of the condyles and the level of several biological markers in TMJ OA subjects [19]. As it is unlikely that a single biomarker would drive or identify a complex disease such as osteoarthritis [17–20], we hypothesize that clinical symptoms, subchondral bone radiomics and biological markers are optimal integrative indicators of TMJ health status.

Analysis of large and complex datasets derived from different sources yields better understanding of the disease. However, detection of unknown patterns in big data requires the use of high-end computing solutions and advanced analytical approaches such as machine-learning algorithms [21]. Although prediction models can analyze a large amount of data, incorporating less variables into the model reduces computing resources' consumption and

prevents model overfitting [22, 23]. Therefore, using a dimensionality reduction technique to identify the optimal subset of the original features is crucial for accurate construction of prediction models [5, 24]. Another challenge for developing a predictive model for TMJ OA diagnosis is inclusion of the biochemical markers. This is due to the restricted specimens' collection, cost and limitations of protein expression measurement systems [25].

In this study, we address the need for comprehensive quantitative phenotyping of OA in the whole jaw joint. We employ a machine learning paradigm called learning using privileged information (LUPI) and train it with clinical, quantitative imaging and additional biological features as privileged information to classify TMJ OA patients. We also adopt feature selection method to remove redundant and irrelevant features from the feature space. Furthermore, we utilize features occurrence and Shapely additive explanations method to interpret the model predictions [26, 27].

2 Methods

2.1 Dataset

Our dataset consisted of 46 early-stage TMJ OA patients and 46 age and gender-matched healthy controls recruited at the University of Michigan School of Dentistry. All the diagnoses were confirmed by a TMD and orofacial pain specialist based on the DC/TMD. The clinical, biological and radiographic data described below were collected from TMJ OA and control subjects with informed consent and following the guidelines of the Institutional Review Board HUM00113199.

2.1.1 Clinical data

Clinical dataset entailed three features obtained from diagnostic tests assessed by the same investigator: 1) headaches in the last month, 2) muscle soreness in the last month, 3) vertical range of unassisted jaw opening without pain (mouth opening).

2.1.2 Biological data

Association of proteins expression with arthritis initiation and progression was investigated in a previous study [28]. In this project, using customized protein microarrays (RayBiotech, Inc. Norcross, GA), the expression level of 13 proteins was measured in the participants' saliva and serum samples. The analyzed proteins included: Angiogenin, BDNF, CXCL16, ENA-78, MMP-3, MMP-7, OPG, PAI-1, TGFb1, TIMP-1, TRANCE, VE-Cadherin and VEGF. As the protein expression of MMP3 was not detected in the saliva, it was excluded from subsequent analysis.

2.1.3 Radiological data

Using the 3D Accuitomo machine (J. Morita MFG. CORP Tokyo, Japan), cone-beam computed tomography (CBCT) scans were performed for each subject. Radiomics analysis was centered on the lateral region of the articular fossa, articular eminence and condyle, a site where greater OA bone degeneration occurs. Radiomic features were extracted using BoneTexture module in 3D-slicer software v.4.11(www.3dslicer.org) [29]. We measured 23 texture features: 5 bone morphometry features, 8 Gray Level Co-occurrence Matrix(GLCM) and 10 Grey-Level Run Length Matrix (GLRLM) features. ClusterShade and HaralickCorrelation measurements were highly variable among all participants, therefore, they were not included in the following analysis.

Joint space measurement was evaluated using 3D condylar-to-fossa distances at the anterior, anterolateral, medial, superior and posterior regions.

2.2 Statistical and machine learning approaches

In this section, we describe methods utilized for building a robust TMJOA diagnosis model (Figure 1). These methods include: 1) cross-validation and grid search, 2) feature selection and 3) learning using privileged information.

2.2.1 Cross-validation and grid search

Cross-validation is an effective approach to model hyperparameter optimization and model selection that attempts to overcome the overfitting issue. The dataset was split into 80% for training and 20% holdout for testing. The 5fold cross-validation with the same portion of data split was nested inside the 80% train dataset, and grid search was performed in each fold of data for hyperparameters tuning. The best combination of hyperparameters was picked based on the mean and standard deviation of F1 scores over the 5-fold cross-validation. The overall procedure was repeated 10 times with 10 random seeds to avoid sampling bias from data partitioning. The final evaluation scores reported in this study are the mean±standard deviation of the holdout test set performance across all 10 repetitions.

2.2.2 Feature selection

Feature selection is a common dimensional reduction technique for building a machine learning model. Increasing the number of features often results in decreasing the prediction error. However, it increases the risk of model overfitting particularly with small datasets. Here, we customized a feature selection method that takes the advantages of privileged variables and mutual information to improve the performance of the classifier.

Normalized mutual information feature selection (NMIFS) method and its modified version called called NMIFS+ was used to measure the relevance and

redundancy of features with the primary objective of high accuracy with the least possible time complexity [30]. NMIFS+ extends the NMIFS algorithm with the LUPI framework, which could take full account of the privilege features along with standard features and make feature selection from those two sets separately [31]. The NMIFS+ was applied to all the LUPI models in this study and, correspondingly, the NMIFS on non-LUPI models.

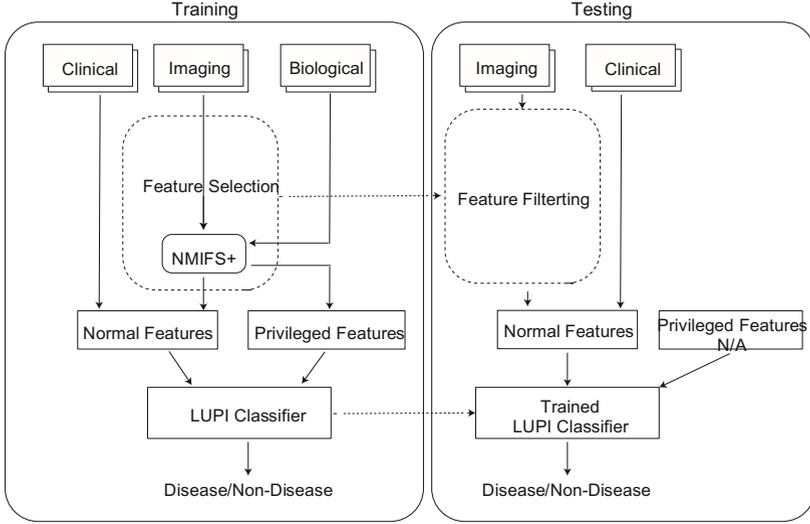


Fig. 1 Diagram of training and testing process

2.2.3 LUPI framework

The idea of learning using privileged information (LUPI) was first proposed as capturing the essence of teacher-student-based learning by Vapnik and Vashist [32]. In contrast to the existing machine learning paradigm, where the model learns and makes predictions with fixed information, the LUPI paradigm considers several specific forms of privileged information, just like a teacher who provides additional information, which can include comments, explanations, and logic to students and thus increases the learning efficiency.

In the classical binary classification model, we were given training pairs $(x_1, y_1), \dots, (x_l, y_l)$, where $x_i \in X, y_i \in \{-1, 1\}, i = 1, \dots, l$, and each pair is independently generated by some underlying distribution P_{XY} , which is unknown. The model is trained to find among a given set of functions $f(x, \alpha), \alpha \in \Lambda$, the function $y = f(x, \alpha)$ that minimizes the probability of incorrect classifications over the unknown distribution P_{XY} .

In the LUPI framework, we were given training triplets $(x_1, x^*_1, y_1), \dots, (x_l, x^*_l, y_l), x_i \in X, x^*_i \in X^*, y_i \in \{-1, 1\}, i = 1, \dots, l$, which is slightly different from the classical one. Each triplet is independently generated by some underlying distribution P_{XX^*Y} , which is still unknown. The additional privileged

information is available only for the training examples, not for the test phase. In this scenario, we can utilize X^* to improve learning performance.

There are a few implementations of LUPI models. One of them is based on random vector functional link network (RVFL) that is a randomized version of the functional link neural network [33, 34]. A kernel-based RVFL, called KRVFL+, has been proposed based on the LUPI paradigm [35]. It incorporates efficient ways to use kernel tricks for highly complicated nonlinear feature training and train RVFL networks with privileged information (Figure 2). The parameters, including weights and biases, from the input layer to the hidden layers are generated randomly from a fixed domain, and only the output weights need to be computed.

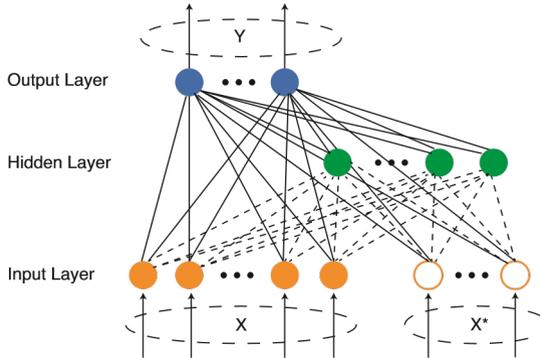


Fig. 2 The architecture of KRVFL+ network. Solid lines are output weights and dash lines stand for random weights and biases.

3 Results

3.1 LUPI and non-LUPI models

Figure 3 shows the comparison of the classification performance between LUPI and non-LUPI models. We evaluated the diagnostic potential of imaging features extracted from the articular eminence, articular fossa, condyle, and joint space measurement, as well as clinical features. Only the clinical feature sets provided discriminative models (AUC=0.723) for TMJ OA diagnosis. By introducing LUPI-based models with additional biological features, LUPI paradigm significantly enhanced the model performance on clinical (AUC=0.794), joint space measurement (AUC=0.625), and condyle (AUC=0.641) datasets.

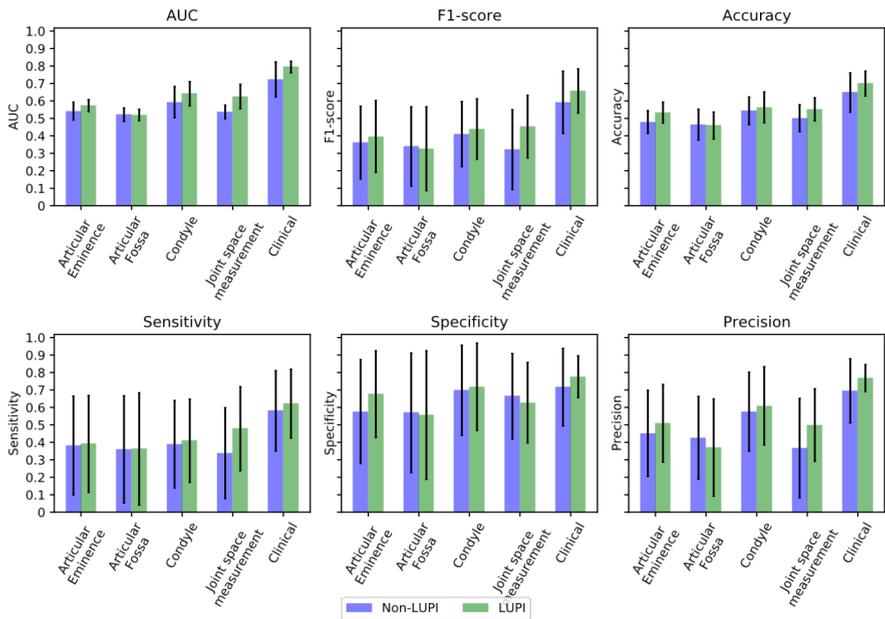


Fig. 3 Comparison of LUPI and non-LUPI models. The non-LUPI models only trained with normal features and RVFL model. The LUPI model trained with KRVFL+ and biological data as privilege information.

3.2 Feature integration comparison

Table 1 shows the classification performances with different feature integration strategies. Given that clinical features had strong discriminative power for TMJ OA diagnosis, two groups of experiments were conducted to investigate the effect of an enlarged candidate pool for feature selection. Adding more features into the clinical dataset and selecting from combined set improved the model performance markedly, i.e., the models had higher AUC scores. With an AUC=0.794, the clinical feature model achieved fairly well performance. Selecting features from a pool of condyle radiomic features together with the clinical features increased the AUC score to 0.804. The performance was even higher when feature selection was conducted on all condyle, 3D measurements and clinical datasets, AUC=0.807. Keeping all clinical criteria and applying feature selection on the remaining dataset resulted in slightly higher AUC values. The AUC scores became 0.808 and 0.809 for the condyle and condyle with additional 3D measurement features models, respectively.

Table 1 Comparison of different feature integration methods (in percentage %)

Feature Set	AUC	F1 score	Accuracy	Sensitivity	Specificity	Precision
Cl	79.4±3.4	65.7±12.7	69.9±7.2	62.2.0±19.8	77.6±12.0	76.8±7.8
(Cl+Cd)*	80.4±3.8	67.5±9.4	70.4±5.6	64.4±18.6	76.4±16.0	76.1±9.2
Cl+Cd*	80.8±4.1	64.8±11.6	69.4±6.4	60.2±19.4	78.7±13.5	76.0±9.3
(Cl+Cd JS)*	80.7±3.8	64.2±15.0	69.8±6.9	61.3±22.9	78.2±15.3	75.1±12.2
Cl+Cd JS*	80.9±3.6	66.1±12.2	70.9±6.0	62.7±19.7	79.1±13.6	77.4±9.8

Cl: Clinical; Cd: Condyle; Cd _JS: Condyle and 3D Joint Space measurements.

* indicates feature selection by NMIFS+ method.

The feature sets in parentheses have been pooled together for feature selection, otherwise it proceeded on feature set with * separately.

All the models have been trained with KRVFL+ with Biological data as privilege information.

3.3 Feature occurrence and importance

To interpret the prediction of our proposed model, we utilized feature occurrence and Shapley values. The NMIFS+ method is a measure of redundancy among features. The calculation of mutual information and redundancy highly depends on the training samples which varied from split to split. Feature occurrence means how many times a feature was selected by NMIFS+ method among the total 50 models. The more times a feature occurs, the more reliable its importance is (Figure 1A). Shapley values were used to interpret the contribution of individual features into the prediction of the trained model. Contributing features are shown in Figure 4B according to the order of the mean absolute of Shapley values across all the data, which indicate the average impact of feature on model output magnitude. Figure 4C provides further indication of Shapley values and shows the complexity of feature contribution in models. Each circle represents a feature value of one patient/control, either increases or decreases the prediction(positive value and negative value). Figure 4D combines feature importance with feature effects. Here we picked one model for visualization instead of pulling all 50 models together. Each point in the summary plot is a Shapley value for a feature and a patient/control. The order of the features on the y-axis is based on their importance. The color represents the Shapley value of the features from low to high. We divided the instances into TMJOA diseased group and Control group, displayed in different markers. Higher values of headache, LongRunHighGreyLevelRunEmphasis and muscle soreness increased the probability of assigning TMJ OA diagnosis.

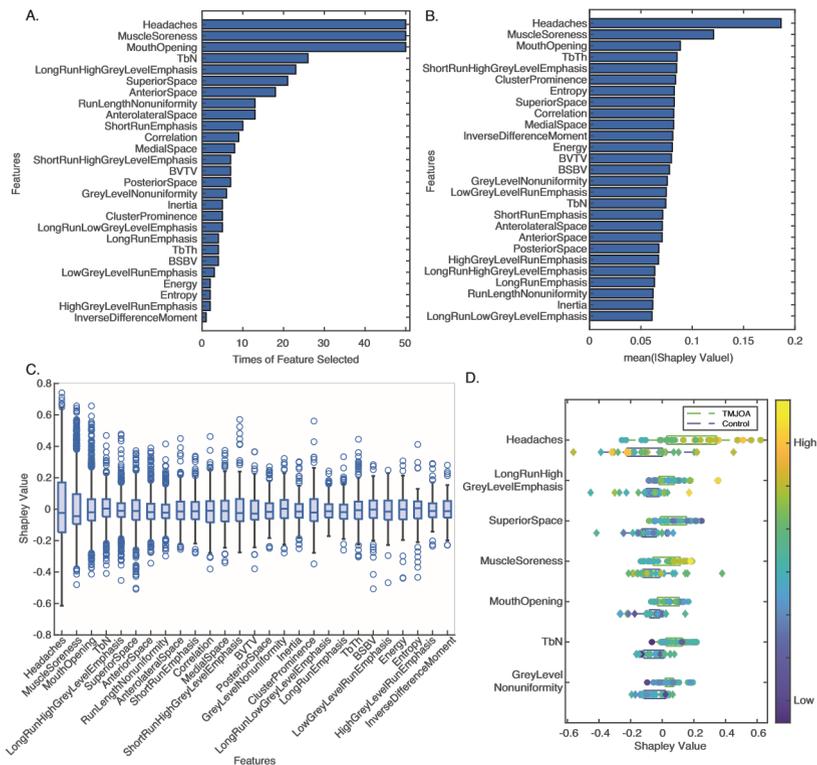


Fig. 4 A. Feature occurrence in 50 trained models using NMIFS method. B. Feature importance measured as the mean absolute Shapley values in 50 models. C. Distribution of Shapley values in each query point in the 50 models. The order of the features shown in the x-axis is based on the feature occurrence. D. Shapley summary plot for one model. The boxplots represent the distribution of TMJOA and control groups (each TMJOA patient is shown as a circle and control as a diamond). The Heatmap color bar shows the value of the feature itself from high to low (yellow to blue). Low number of Shapley value of features reduce the predicted TMJOA diseased probability, a large number of Shapley value increase the probability.

4 Discussion

This study developed an enhanced model for TMJ OA diagnosis, utilizing state-of-the-art machine learning technology and considering clinical, quantitative imaging markers, and additional biological features used only for training. This is the first study to utilize quantitative imaging markers of the whole joint: condyle, articular space, articular fossa and articular eminence. We employed feature selection to minimize feature sets and improve the model robustness. Furthermore, feature occurrence and Shapley value were assessed to reduce the black-box nature of the machine learning model, as well as improve the domain experts' confidence in the model's prediction. This study findings demonstrate excellent performance of the feature integration methods and LUPI paradigm in predicting TMJ OA status.

The Diagnostic Criteria for Temporomandibular Disorders (DC/TMD) have been the most utilized protocol for TMJ OA diagnosis. However, these criteria are dependent on subjective clinical signs/symptoms and subjective radiological interpretation of imaging features associated with irreversible bone changes [4, 5]. Early treatment and modification of the disease course requires precise diagnosis of TMJ OA at initial stages [36]. In this study, we utilized multi-source data collected from subjects at early stages of TMJ OA. We employed the LUPI paradigm and used biological features of inflammation, neuroception, bone resorption and angiogenesis as privileged information. The LUPI algorithm allowed benefiting from diagnostic information within the existing biological data and eliminated future need for biological samples' collection and analysis. Inclusion of biological data with the LUPI framework boosted our model performance, confirming the need for biological data only for model training. We developed a robust model for TMJ OA diagnosis and validated its performance using extensive evaluation metrics (Figure 1). Our model demonstrated sensitivity and specificity of 63% and 79%, respectively. These values exceeded the sensitivity and specificity, 58% and 72%, of TMJ OA diagnosis following DC/TMD protocol without imaging [4]. Honda and colleagues [37] reported that the CBCT scan's use improved the sensitivity and specificity for detecting condylar osseous defects to 80% and 90%, sequentially. Nevertheless, CBCT sensitivity is dependent on the defects' size, it is challenging to detect early alterations that are <2mm. Hence, we extracted objective, quantitative imaging features from the subchondral bones of the condyle, articular fossa and articular eminence. Using the LUPI-based model, we found that only condyle's radiomics could differentiate between healthy and diseased subjects (Table 1). In line with this observation, Massilla and Sivasubramanian [38] reported that patients with early TMJ OA had osteoarthritic bone alterations in their condyles (69.93%) more than articular fossa (10%) and articular eminence (6.6%). Interestingly, we noted that the superior 3D joint space distinguished TMJ OA subjects using LUPI-based models (AUC=.63), denoting the importance of this feature in detecting osteoarthritic changes. In fact, in another study [38], joint space narrowing was the second predominant radiographic sign observed in patients with early TMJ OA. Along with radiomics and joint space measurements, we supplemented the model with clinical signs that were measurable in both groups. Elimination of leaky variables prevents biasing the model and promotes its reliability and well generalization with new data [39].

Machine learning models are leveraged for clinical predictive modeling, where clinical values are used to predict clinical diagnosis. However, these models do not explain the basis for their prediction. This raise concerns in medical domains and challenge researchers to identify reasons behind the model outcomes [40]. Here, we facilitated the interpretability of our model by reducing the number of candidate features. In general, for a fixed sample size,

the error of designed classifier decreases and then increases as the number of feature grows. Finding an optimal number of features is crucial in terms of reducing the time to build the learning model and increasing the accuracy in the learning process. For uncorrelated features, the optimal feature size is $N-1$, where the N is the sample size. As the feature correlation increases, the optimal feature size becomes proportional to N for highly correlated features [41]. Furthermore, texture features turned out to be highly correlated in Cho's work [42]. Those further proof of the necessity of feature selection.

Using the NMIFS method, we calculated feature occurrence to identify the discriminative features of TMJ OA. Moreover, we calculated Shapley values to demonstrate how each clinical and imaging feature is contributing to the outcome/disease diagnosis in individual patients. Headache, muscle soreness and limited range of vertical mouth opening without pain were among the top features that contributed to the model prediction for TMJ OA. This aligns with the common observation of these symptoms in individuals with painful temporomandibular disorders [43]. TrabecularNumber, superior 3D joint space and LongRunHighGreyLevelRunEmphasis were the top imaging features selected in the majority of the trained models. Importantly, the amalgamation of different data-sources in this study is essential for comprehensive assessment of individuals' health. In line with our results, Liang and colleagues found significant differences of the TrabecularNumber in subjects with TMJ OA compared to healthy individuals [44]. Our findings corroborate those that indicate radiomics provide an objective assessment of the pathological changes and may overcome the subjectivity of patients-reported symptoms [45]. Previous studies have reported joint space narrowing in subjects with TMJ OA [46, 47]. Zhang et al. [48] validated the importance of detecting TMJ morphological changes using 3D measurements, showing that 2D and 3D TMJ space measurements varied significantly in CBCT scans of healthy individuals. The present study is the first to test whole joint (condylar, articular eminence and articular fossa) radiomics and incorporate 3D joint space measurements into a comprehensive diagnostic tool for TMJ OA.

5 Conclusion

Normalized mutual information feature selection method and LUPI paradigm established a robust model for TMJ OA diagnosis. The identified clinical and quantitative imaging markers can be considered a foundation for reliable detection of TMJ OA pathological alterations and are potential markers for prediction of disease progression in future longitudinal studies.

Declarations

- **Funding:** This study was funded by American Association of Orthodontists Foundation Biomedical Research Award; National Institute of Dental and Craniofacial Research, Grant/Award Number: R01DE024450.
- **Conflict of interest:** The authors declare no potential conflicts of interest with respect to the authorship and/or publication of this article.
- **Ethics approval:** This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Ethics Committee of University Michigan Institutional Review Board HUM00113199.
- **Consent to participate:** Informed consent was obtained from all individual participants included in the study.
- **Consent for publication:** Not applicable.

- **Availability of data and materials:** The data that support the findings of this study are available upon reasonable request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.
- **Code availability:** The codes are available from the authors upon reasonable request and with the permission of technology transfer office.
- **Authors' contributions:** Najla Al Turkestani and Lingrui Cai equally contributed to the paper and worked on data curation, software development, formal analysis, investigation, methodology, project administration and writing original draft. Lucia Cevidanes: conceptualization, formal analysis, investigation, methodology, funding acquisition, project administration, manuscript review editing. Jonas Bianchi and Marcela Gurgel: data curation, formal analysis for training data, investigation, manuscript review editing. Baptiste Baquero and Maxime Gillot: Manuscript review editing. Winston Zhang: Methodology and Software development. Kayvan Najarian and Reza Soroushmehr: Methodology, Software development, Writing – review editing.

References

- [1] Frequency of temporomandibular joint osteoarthritis and related symptoms in a hand osteoarthritis cohort **25**. <https://doi.org/10.1016/j.joca.2016.12.028>
- [2] Tanaka, E., Detamore, M.S., Mercuri, L.G.: Degenerative disorders of the temporomandibular joint: etiology, diagnosis, and treatment. *Journal of Dental Research* **87**(4), 296–307 (2008). <https://doi.org/10.1177/154405910808700406>
- [3] O'Connor, R.C., Fawthrop, F., Salha, R., Sidebottom, A.J.: Management of the temporomandibular joint in inflammatory arthritis: Involvement of surgical procedures. *European Journal of Rheumatology* **4**(2), 151–156 (2017). <https://doi.org/10.5152/eurjrheum.2016.035>
- [4] Schiffman, E., Ohrbach, R., Truelove, E., Look, J., Anderson, G., Goulet, J.-P., List, T., Svensson, P., Gonzalez, Y., Lobbezoo, F., Michelotti, A., Brooks, S.L., Ceusters, W., Drangsholt, M., Ettlin, D., Gaul, C., Goldberg, L.J., Haythornthwaite, J.A., Hollender, L., Jensen, R., John, M.T., De Laat, A., de Leeuw, R., Maixner, W., van der Meulen, M., Murray, G.M., Nixdorf, D.R., Palla, S., Petersson, A., Pionchon, P., Smith, B., Visscher, C.M., Zakrzewska, J., Dworkin, S.F.: Diagnostic Criteria for Temporomandibular Disorders (DC/TMD) for Clinical and Research Applications: Recommendations of the International RDC/TMD Consortium Network and Orofacial Pain Special Interest Group. *Journal of oral & facial pain and headache* **28**(1), 6–27 (2014). Accessed 2022-03-22
- [5] Jamshidi, A., Pelletier, J.-P., Martel-Pelletier, J.: Machine-learningbased patient-specific prediction models for knee osteoarthritis. *Nature Reviews Rheumatology* **15**(1), 49–60 (2019). <https://doi.org/10.1038/s41584-018-0130-5>. Number: 1 Publisher: Nature Publishing Group. Accessed 2022-05-24
- [6] Shi, J., Lee, S., Pan, H.C., Mohammad, A., Lin, A., Guo, W., Chen, E., Ahn, A., Li, J., Ting, K., Kwak, J.H.: Association of Condylar Bone Quality with TMJ Osteoarthritis. *Journal of Dental Research* **96**(8), 888– 894 (2017). <https://doi.org/10.1177/0022034517707515>. Publisher: SAGE Publications Inc. Accessed 2022-05-24
- [7] Martel-Pelletier, J., Tardif, G., Paiement, P., Pelletier, J.-P.: Common Biochemical and Magnetic Resonance Imaging Biomarkers of Early Knee Osteoarthritis and of Exercise/Training in Athletes: A Narrative Review. *Diagnostics* **11**(8), 1488 (2021). <https://doi.org/10.3390/d11081488>

[diagnostics11081488](#). Number: 8 Publisher: Multidisciplinary Digital Publishing Institute. Accessed 2022-05-24

- [8] Hu, Y., Chen, X., Wang, S., Jing, Y., Su, J.: Subchondral bone microenvironment in osteoarthritis and pain. *Bone Research* **9**(1), 1–13 (2021). <https://doi.org/10.1038/s41413-021-00147-z>. Number: 1 Publisher: Nature Publishing Group. Accessed 2022-05-24
- [9] Marias, K.: The Constantly Evolving Role of Medical Image Processing in Oncology: From Traditional Medical Image Processing to Imaging Biomarkers and Radiomics. *Journal of Imaging* **7**(8), 124 (2021). <https://doi.org/10.3390/jimaging7080124>. Accessed 2022-05-24
- [10] Munjal, A., Bapat, S., Hubbard, D., Hunter, M., Kolhe, R., Fulzele, S.: Advances in Molecular biomarker for early diagnosis of Osteoarthritis. *Biomolecular Concepts* **10**(1), 111–119 (2019). <https://doi.org/10.1515/bmc-2019-0014>. Publisher: De Gruyter. Accessed 2022-05-25
- [11] Lotz, M., Martel-Pelletier, J., Christiansen, C., Brandi, M.-L., Bruy`ere, O., Chapurlat, R., Collette, J., Cooper, C., Giacobelli, G., Kanis, J.A., Karsdal, M.A., Kraus, V., Lems, W.F., Meulenbelt, I., Pelletier, J.-P., Raynauld, J.-P., Reiter-Niesert, S., Rizzoli, R., Sandell, L.J., Van Spil, W.E., Reginster, J.-Y.: Value of biomarkers in osteoarthritis: current status and perspectives. *Annals of the Rheumatic Diseases* **72**(11), 1756–1763 (2013). <https://doi.org/10.1136/annrheumdis-2013-203726>. Accessed 2022-05-24
- [12] Mattonen, S.A., Palma, D.A., Johnson, C., Louie, A.V., Landis, M., Rodrigues, G., Chan, I., Etemad-Rezai, R., Yeung, T.P.C., Senan, S., Ward, A.D.: Detection of Local Cancer Recurrence After Stereotactic Ablative Radiation Therapy for Lung Cancer: Physician Performance Versus Radiomic Assessment. *International Journal of Radiation Oncology*Biophysics*Physics* **94**(5), 1121–1128 (2016). <https://doi.org/10.1016/j.ijrobp.2015.12.369>. Accessed 2022-05-24
- [13] Barua, S., Elhalawani, H., Volpe, S., Al Feghali, K.A., Yang, P., Ng, S.P., Elgohari, B., Granberry, R.C., Mackin, D.S., Gunn, G.B., Hutcheson, K.A., Chambers, M.S., Court, L.E., Mohamed, A.S.R., Fuller, C.D., Lai, S.Y., Rao, A.: Computed Tomography Radiomics Kinetics as Early Imaging Correlates of Osteoradionecrosis in Oropharyngeal Cancer Patients. *Frontiers in Artificial Intelligence* **4** (2021). Accessed 2022-05-24
- [14] Khorrami, M., Bera, K., Leo, P., Vaidya, P., Patil, P., Thawani, R., Velu, P., Rajiah, P., Alilou, M., Choi, H., Feldman, M.D., Gilkeson, R.C., Linden, P., Fu, P., Pass, H., Velcheti, V., Madabhushi, A.: Stable and discriminating radiomic

predictor of recurrence in early stage nonsmall cell lung cancer: Multi-site study. *Lung Cancer* **142**, 90–97 (2020). <https://doi.org/10.1016/j.lungcan.2020.02.018>. Accessed 2022-05-24

- [15] Oikonomou, A., Khalvati, F., Tyrrell, P.N., Haider, M.A., Tarique, U., Jimenez-Juan, L., Tjong, M.C., Poon, I., Eilaghi, A., Ehrlich, L., Cheung, P.: Radiomics analysis at PET/CT contributes to prognosis of recurrence and survival in lung cancer treated with stereotactic body radiotherapy. *Scientific Reports* **8**(1), 4003 (2018). <https://doi.org/10.1038/s41598-018-22357-y>. Accessed 2022-05-24
- [16] Zhang, B., Lian, Z., Zhong, L., Zhang, X., Dong, Y., Chen, Q., Zhang, L., Mo, X., Huang, W., Yang, W., Zhang, S.: Machine-learning based MRI radiomics models for early detection of radiation-induced brain injury in nasopharyngeal carcinoma. *BMC Cancer* **20**(1), 502 (2020). <https://doi.org/10.1186/s12885-020-06957-4>. Accessed 2022-05-24
- [17] Cevidanes, L., Hajati, A.-K., Paniagua, B., Lim, P., Walker, D., Palconet, G., Nackley, A., Styner, M., Ludlow, J., Zhu, H., Phillips, C.: Quantification of Condylar Resorption in TMJ Osteoarthritis. *Oral surgery, oral medicine, oral pathology, oral radiology, and endodontics* **110**(1), 110–117 (2010). <https://doi.org/10.1016/j.tripleo.2010.01.008>. Accessed 2022-05-05
- [18] Bianchi, J., Goncalves, J.R., de Oliveira Ruellas, A.C., Ashman, L.M., Vimort, J.-B., Yatabe, M., Paniagua, B., Hernandez, P., Benavides, E., Soki, F.N., Ioshida, M., Cevidanes, L.H.S.: Quantitative bone imaging biomarkers to diagnose temporomandibular joint osteoarthritis. *International Journal of Oral and Maxillofacial Surgery* **50**(2), 227–235 (2021). <https://doi.org/10.1016/j.ijom.2020.04.018>. Accessed 2022-05-24
- [19] Cevidanes, L.H.S., Walker, D., Schilling, J., Sugai, J., Giannobile, W., Paniagua, B., Benavides, E., Zhu, H., Marron, J.S., Jung, B.T., Baranowski, D., Rhodes, J., Nackley, A., Lim, P.F., Ludlow, J.B., Nguyen, T., Goncalves, J.R., Wolford, L., Kapila, S., Styner, M.: 3D osteoarthritic changes in TMJ condylar morphology correlates with specific systemic and local biomarkers of disease. *Osteoarthritis and Cartilage* **22**(10), 1657–1667 (2014). <https://doi.org/10.1016/j.joca.2014.06.014>
- [20] Deveza, L.A., Nelson, A.E., Loeser, R.F.: Phenotypes of osteoarthritis - current state and future implications. *Clinical and experimental rheumatology* **37**(Suppl 120), 64–72 (2019). Accessed 2022-05-24
- [21] Al Turkestani, N., Bianchi, J., Deleat-Besson, R., Le, C., Tengfei, L., Prieto, J.C., Gurgel, M., Ruellas, A.C.O., Massaro, C., Aliaga Del Castillo, A., Evangelista, K., Yatabe, M., Benavides, E., Soki, F., Zhang, W., Najarian, K., Gryak, J.,

- Styner, M., Fillion-Robin, J.-C., Paniagua, B., Soroushmehr, R., Cevidanes, L.H.S.: Clinical decision support systems in orthodontics: A narrative review of data science approaches. *Orthodontics & Craniofacial Research* **24**(S2), 26–36 (2021). <https://doi.org/10.1111/ocr.12492>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/ocr.12492>. Accessed 2022-05-24
- [22] Liu, H., Dougherty, E.R., Dy, J.G., Torkkola, K., Tuv, E., Peng, H., Ding, C., Long, F., Berens, M., Parsons, L., Yu, L., Zhao, Z., Forman, G.: Evolving feature selection. *IEEE Intelligent Systems* **20**(6), 64–76 (2005). <https://doi.org/10.1109/MIS.2005.105>. Conference Name: IEEE Intelligent Systems
- [23] Venkatesh, B., Anuradha, J.: A Review of Feature Selection and Its Methods. *Cybernetics and Information Technologies* **19**(1), 3–26 (2019). <https://doi.org/10.2478/cait-2019-0001>. Accessed 2022-05-24
- [24] Jalali-najafabadi, F., Stadler, M., Dand, N., Jadon, D., Soomro, M., Ho, P., Marzo-Ortega, H., Helliwell, P., Korendowych, E., Simpson, M.A., Packham, J., Smith, C.H., Barker, J.N., McHugh, N., Warren, R.B., Barton, A., Bowes, J.: Application of information theoretic feature selection and machine learning methods for the development of genetic risk prediction models. *Scientific Reports* **11**(1), 23335 (2021). <https://doi.org/10.1038/s41598-021-00854-x>. Number: 1 Publisher: Nature Publishing Group. Accessed 2022-05-24
- [25] Doetzer, A.D., Herai, R.H., Buzalaf, M.A.R., Trevilatto, P.C.: Proteomic Expression Profile in Human Temporomandibular Joint Dysfunction. *Diagnostics* **11**(4), 601 (2021). <https://doi.org/10.3390/diagnostics11040601>. Accessed 2022-05-24
- [26] Lundberg, S., Lee, S.-I.: A Unified Approach to Interpreting Model Predictions. Technical Report arXiv:1705.07874, arXiv (November 2017). <https://doi.org/10.48550/arXiv.1705.07874>. arXiv:1705.07874 [cs, stat] type: article. <http://arxiv.org/abs/1705.07874> Accessed 2022-05-25
- [27] Aas, K., Jullum, M., Løland, A.: Explaining individual predictions when features are dependent: More accurate approximations to Shapley values. Technical Report arXiv:1903.10464, arXiv (February 2020). <https://doi.org/10.48550/arXiv.1903.10464>. arXiv:1903.10464 [cs, stat] type: article. <http://arxiv.org/abs/1903.10464> Accessed 2022-05-25
- [28] Shoukri, B., Prieto, J.C., Ruellas, A., Yatabe, M., Sugai, J., Styner, M., Zhu, H., Huang, C., Paniagua, B., Aronovich, S., Ashman, L., Benavides, E., de Dumast, P., Ribera, N.T., Mirabel, C., Michoud, L., Allohaibi, Z., Ioshida, M.,

Bittencourt, L., Fattori, L., Gomes, L.R., Cevidanes, L.: Minimally Invasive Approach for Diagnosing TMJ Osteoarthritis. *Journal of Dental Research* **98**(10), 1103–1111 (2019). <https://doi.org/10.1177/0022034519865187>

- [29] Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.-C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F., Sonka, M., Buatti, J., Aylward, S., Miller, J.V., Pieper, S., Kikinis, R.: 3D Slicer as an Image Computing Platform for the Quantitative Imaging Network. *Magnetic resonance imaging* **30**(9), 1323–1341 (2012). <https://doi.org/10.1016/j.mri.2012.05.001>. Accessed 2022-05-05
- [30] Estevez, P.A., Tesmer, M., Perez, C.A., Zurada, J.M.: Normalized Mutual Information Feature Selection. *IEEE Transactions on Neural Networks* **20**(2), 189–201 (2009). <https://doi.org/10.1109/TNN.2008.2005601>. Accessed 2022-04-07
- [31] Zhang, W., Turkestani, N.A., Bianchi, J., Le, C., Deleat-Besson, R., Ruellas, A., Cevidanes, L., Yatabe, M., Goncalves, J., Benavides, E., Soki, F., Prieto, J., Paniagua, B., Gryak, J., Najarian, K., Soroushmehr, R.: Feature Selection for Privileged Modalities in Disease Classification. In: *Multimodal Learning for Clinical Decision Support: 11th International Workshop, ML-CDS 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, October 1, 2021, Proceedings*, pp. 69–80. Springer, Berlin, Heidelberg (2021). <https://doi.org/10.1007/978-3-030-89847-2> 7. <https://doi.org/10.1007/978-3-030-89847-2> Accessed 2022 - 05 - 26
- [32] Vapnik, V., Vashist, A.: A new learning paradigm: Learning using privileged information. *Neural Networks* **22**(5), 544–557 (2009). <https://doi.org/10.1016/j.neunet.2009.06.042>. Accessed 2022-04-19
- [33] Pao, Y.-H., Takefuji, Y.: Functional-link net computing: theory, system architecture, and functionalities. *Computer* **25**(5), 76–79 (1992). <https://doi.org/10.1109/2.144401>. Conference Name: Computer
- [34] Pao, Y.-H., Park, G.-H., Sobajic, D.J.: Learning and generalization characteristics of the random vector functional-link net. *Neurocomputing* **6**(2), 163–180 (1994). [https://doi.org/10.1016/0925-2312\(94\)90053-1](https://doi.org/10.1016/0925-2312(94)90053-1). Accessed 2022-05-26
- [35] Zhang, P.-B., Yang, Z.-X.: A new learning paradigm for random vector functional-link network: RVFL. *Neural Networks: The Official Journal of the International Neural Network Society* **122**, 94–105 (2020). <https://doi.org/10.1016/j.neunet.2019.09.039>

- [36] Chu, C.R., Williams, A.A., Coyle, C.H., Bowers, M.E.: Early diagnosis to enable early treatment of pre-osteoarthritis. *Arthritis research & therapy* **14**(3), 1–10 (2012)
- [37] Honda, K., Larheim, T., Maruhashi, K., Matsumoto, K., Iwai, K.: Osseous abnormalities of the mandibular condyle: diagnostic reliability of cone beam computed tomography compared with helical computed tomography based on an autopsy material. *Dentomaxillofacial Radiology* **35**(3), 152–157 (2006)
- [38] Mani, F.M., Sivasubramanian, S.S.: A study of temporomandibular joint osteoarthritis using computed tomographic imaging. *biomedical journal* **39**(3), 201–206 (2016)
- [39] Dong, Q.: Leakage prediction in machine learning models when using data from sports wearable sensors. *Computational Intelligence and Neuroscience* **2022** (2022)
- [40] Petch, J., Di, S., Nelson, W.: Opening the black box: the promise and limitations of explainable machine learning in cardiology. *Canadian Journal of Cardiology* (2021)
- [41] Hua, J., Xiong, Z., Lowey, J., Suh, E., Dougherty, E.R.: Optimal number of features as a function of sample size for various classification rules. *Bioinformatics* **21**(8), 1509–1515 (2004). <https://doi.org/10.1093/bioinformatics/bti171>. eprint: <https://academic.oup.com/bioinformatics/article-pdf/21/8/1509/691983/bti171.pdf>
- [42] Cho, Y.J., Kim, W.S., Choi, Y.H., Ha, J.Y., Lee, S., Park, S.J., Cheon, J.-E., Kang, H.J., Shin, H.Y., Kim, I.-O.: Computerized texture analysis of pulmonary nodules in pediatric patients with osteosarcoma: Differentiation of pulmonary metastases from non-metastatic nodules. *PLOS ONE* **14**(2), 1–14 (2019). <https://doi.org/10.1371/journal.pone.0211969>
- [43] Kalladka, M., Quek, S., Heir, G., Eliav, E., Mupparapu, M., Viswanath, A.: Temporomandibular joint osteoarthritis: diagnosis and long-term conservative management: a topic review. *The Journal of Indian Prosthodontic Society* **14**(1), 6–15 (2014)
- [44] Liang, X., Liu, S., Qu, X., Wang, Z., Zheng, J., Xie, X., Ma, G., Zhang, Z., Ma, X.: Evaluation of trabecular structure changes in osteoarthritis of the temporomandibular joint with cone beam computed tomography imaging. *Oral surgery, oral medicine, oral pathology and oral radiology* **124**(3), 315–322 (2017)

- [45] Patel, V.K., Naik, S.K., Naidich, D.P., Travis, W.D., Weingarten, J.A., Lazzaro, R., Gutterman, D.D., Wentowski, C., Grosu, H.B., Raoof, S.: A practical algorithmic approach to the diagnosis and management of solitary pulmonary nodules: part 1: radiologic characteristics and imaging modalities. *Chest* **143**(3), 825–839 (2013)
- [46] Alexiou, K., Stamatakis, H., Tsiklakis, K.: Evaluation of the severity of temporomandibular joint osteoarthritic changes related to age using cone beam computed tomography. *Dentomaxillofacial Radiology* **38**(3), 141–147 (2009)
- [47] Ko, c, N.: Evaluation of osteoarthritic changes in the temporomandibular joint and their correlations with age: A retrospective cbct study. *Dental and Medical Problems* **57**(1), 67–72 (2020)
- [48] Zhang, Y., Xu, X., Liu, Z.: Comparison of morphologic parameters of temporomandibular joint for asymptomatic subjects using the two-dimensional and three-dimensional measuring methods. *Journal of Healthcare Engineering* **2017** (2017)

Integrative Risk Predictors of Temporomandibular Joint Osteoarthritis Progression

Lingrui Cai,^a Najla Al Turkestani,^{b,c} Lucia Cevidanes,^c Jonas Bianchi,^d Marcela Gurgel,^c Kayvan Najarian,^a Reza Soroushmehr^a

^aDepartment of Computational Medicine and Bioinformatics, University of Michigan, 100 Washtenaw Ave, Ann Arbor, 48109, MI, USA;

^bDepartment of Orthodontics and Pediatric Dentistry, University of Michigan, 1011 North University Ave, Ann Arbor, 48109, MI, USA;

^cDepartment of Restorative and Aesthetic Dentistry, Faculty of Dentistry, King Abdulaziz University, Jeddah, 22252, Saudi Arabia

^dDepartment of Orthodontics, University of the Pacific, Arthur Dugoni School of Dentistry, 155 5th St, San Francisco, 94103, CA, United States.

ABSTRACT

In this paper we propose feature selection and machine learning approaches to identify a combination of features for risk prediction of Temporomandibular Joint (TMJ) disease progression. In a sample of 32 TMJ osteoarthritis and 38 controls, feature selection of 5 clinical comorbidities, 43 quantitative imaging, 28 biological features and was performed using Maximum Relevance Minimum Redundancy, Chi-Square and Least Absolute Shrinkage and Selection Operator (LASSO) and Recursive Feature Elimination. We compared the performance of learning using concave and convex kernels (LUCCK), support vector machine (SVM) and random forest (RF) approaches to predict disease cure/improvement or persistence/worsening. We show that the SVM model using LASSO achieves area under the curve (AUC), sensitivity and precision of 0.92 ± 0.08 , 0.85 ± 0.19 and 0.76 ± 0.18 , respectively. Baseline levels of headaches, lower back pain, restless sleep, muscle soreness, articular fossa bone surface/bone volume and trabecular separation, condylar High Gray Level Run Emphasis and Short Run High Gray Level Emphasis, saliva levels of 6Ckine, Osteoprotegerin (OPG) and Angiogenin, and serum levels of 6ckine and Brain Derived Neurotrophic Factor (BDNF) were the most frequently occurring features to predict more severe TMJ osteoarthritis prognosis.

Keywords: Temporomandibular Joint Osteoarthritis, disease progression, feature selection, machine learning

1. INTRODUCTION

Osteoarthritis (OA) of the Temporomandibular joint (TMJ) is a prevalent progressive disorder characterized by chronic joint degradation. Rapidly progressive OA may involve multiple joints [1] and severe stages require joint replacement [2]. Assessments of OA have focused on disk and cartilage degradation with no symptom or test that predict the risk of severe prognosis [3]. The bone of the mandibular condyles is located just beneath the fibrocartilage, making it particularly vulnerable to inflammatory damage and a valuable model for studying arthritic changes. It is unlikely a single marker would drive this intricate disease. No disease-modifying therapy exists.

Based on our published results [4,5] we hypothesize that patterns of clinical symptoms, TMJ bone structure and biological mediators are unrecognized indicators of the severity of progression of TMJ OA. Selecting the combination of features that optimizes the performance of machine learning/statistical models is an important task. Many feature selection methods have been proposed in literatures and here we compare filter-based method (Chi-Square), a filter algorithm called Maximum Relevance Minimum Redundancy (mRMR) that uses mutual information criteria as a measure of both relevance and redundancy of features to quantify nonlinear relationships between variables, a wrapper-based algorithm called Recursive Feature Elimination (RFE), and Least Absolute Shrinkage and Selection Operator (LASSO), a method that applies a shrinking/regularization process and feature selection. We then evaluate the performance metrics of Learning Using Concave and Convex Kernels (LUCCK), Support Vector Machine (SVM) and Random Forest (RF) for classifying the patients at risk of severe prognosis.

2. METHODS

This study followed the “Strengthening the Reporting of Observational studies in Epidemiology” (STROBE) guidelines for observational studies and was approved by the Institutional Review Board HUM00113199 from the University of Michigan and the informed consent was obtained from all participants. The longitudinal sample consisted of 32 early-stage TMJ OA patients and 38 healthy controls recruited at the University of Michigan School of Dentistry with a 2.5 ± 0.9 y follow up interval between the subjects’ assessments. All subjects were examined by a Temporomandibular Disorder (TMD) and orofacial pain specialist based on the diagnostic criteria for TMD. The clinical symptoms features entailed 5 comorbidities obtained from diagnostic questionnaire and exam by the same investigator: 1) headaches in the last month, 2) muscle soreness in the last month, 3) vertical range of unassisted jaw opening without pain (mouth opening), 4) restless sleep and 5) lower back pain. Using the 3D Accuitomo cone-beam computed tomography (CBCT, J. Morita MFG. CORP Tokyo, Japan), TMJ scans were performed for each subject to analyze the TMJ bone structure. The joint space was measured at the most superior region of the fossa. Radiomics analysis was centered on the lateral region of the articular fossa and condyle, sites where greater OA bone degeneration occurs. Twenty-one radiomic features in the articular fossa and in the condyle were extracted using the BoneTexture module in 3D Slicer software v.4.11 (www.3dslicer.org) including 5 bone morphometry features, 6 Gray Level Co-occurrence Matrix (GLCM) and 10 Grey-Level Run Length Matrix (GLRLM) features. The biologic mediators were evaluated using customized protein microarrays (RayBiotech, Inc. Norcross, GA), which determines the expression level of 14 proteins was measured in the participants’ saliva and serum samples. The analyzed proteins included: 6Ckine, Angiogenin, BDNF, CXCL16, ENA-78, MMP-3, MMP-7, OPG, PAI-1, TGFb1, TIMP-1, TRANCE, VE-Cadherin and VEGF. As MMP3 measures were below the level of detection in saliva, it was excluded from subsequent analysis. The criteria for determining health status at the follow-up evaluation was scored as 0 (healthy), 1 (improved), 2 (same) or 3 (worsened). These scores were based on the level of clinical pain related symptoms compared to baseline levels, radiographic signs of the disease (subchondral cyst, erosion, osteophyte) assessed by 2 radiologist experts and 3D degenerative morphological changes. Since the small and unbalanced sample size of 4 categories (4 worsened, 11 same, 9 improved, 8 healthy after treatment) are not sufficient for cross-validation and modeling, we binarized the follow-up evaluation score by combining healthy and improved categories as one group, and same and worsened groups as the other group. The 38 non-TMJ controls have also been included in the training phase, while we only report the evaluation scores on patients with treatment in the results section. Towards building a robust model for TMJ OA prognosis model we performed: 1) stratified cross-validation and grid search, 2) comparison of feature selection methods and 3) comparison of machine learning approaches.

2.1 Cross-validation

The dataset was shuffled randomly and stratified split into 80% for training and 20% for testing based on the severity of disease progression and diagnosis at baseline visit. We performed 5-fold cross-validation and the grid search was performed in each fold of data for hyperparameters tuning, based on the mean and standard deviation of F1 scores. The overall procedure was repeated 10 times with different random seeds for shuffling to avoid sampling bias and overfitting from data partitioning. The final evaluation scores reported in this study are the mean \pm standard deviation of the test set performance on 10 times 5-fold cross-validation.

2.2 Feature Selection

We have collected and measured 76 features from our dataset. To improve the efficiency of training, enhance accuracy of the models and reduce the complexity of models, we performed feature selection and chose a subset of features. These methods are roughly divided into three categories: 1) filter based methods which select features independently from the learning process, 2) wrapper based methods which are based on the learning procedure and perform greedy search by evaluating different combinations of features against an evaluation criterion and 3) embedded methods which integrate feature selection and training of the model. Four feature selection methods were chosen, considering the effectiveness and complexity and covering primary methodologies. 1) Maximum relevance and minimum redundancy (mRMR) method [6] is a filter based method that tends to select a feature subset based on the importance of features and least correlation among them. The relevancy is calculated by mutual information and the redundancy is implemented by Pearson correlation; 2) Chi-Squared [7] is another filter-based method which calculates the chi-squared metric between each feature and the prediction target. It tests whether the occurrences of a specific feature and a specific class are independent. Then top-ranking features with maximum chi-squared values are selected which are highly dependent on the prediction target for

modeling; 3) Recursive feature elimination [8] is a wrapper-based method which selects features by recursively eliminating the number of features. The ranking of feature weights is obtained by training a classifier on the initial feature set and then low ranked features are removed. The procedure is repeated recursively until it achieves the optimum number of features needed to assure peak performance; 4) LASSO [9] regression is an embedded method using l_1 -norm as regularizer. The objective of LASSO is to solve (1) where N and p are the number of samples and features respectively, X is the feature vector, β is the coefficient vector. An important property of LASSO norm regularizer is that it could generate an estimation of penalty with exact zero coefficients, which denotes that the corresponding feature is eliminated. The parameter λ controls the strength of the shrinkage, where the higher the value of λ , the fewer features are selected with non-zero coefficient value.

$$(1) \min_{\beta \in \mathbb{R}^p} \left\{ \frac{1}{N} \|y - X\beta\|_2^2 + \lambda \|\beta\|_1 \right\}$$

2.3 Machine learning approaches

Although there are many machine learning classifiers, in this paper, the performance of three of them was tested on our dataset including RF, SVM and LUCCK, as they learn patterns in data with different approaches. Random forest [10] is an ensemble learning method for classification that operates by combining multiple decision trees at the training time and outputs the class selected by most trees. The decision tree recursively partitions the given dataset into two groups based on a certain criterion until a predetermined stopping condition is met. However, this method is prone to overfitting, especially for decision trees when they perfectly classify the training data. The bootstrap aggregating method and randomization in the data nodes selection process prevent overfitting and improve the performance of a single decision tree. SVM is based on statistical learning theory which finds an optimal hyperplane by minimizing the norm of a vector that defines the separating hyperplanes [11]. The basic intuition of SVM is finding a hyperplane that best separates the datapoints into different classes. In real word, the data might be noisy and the presence of a few outliers can lead to overfitting and eventually misclassification. SVM can work with a hyperplane that separates most but not all datapoints, which is called soft margin, to deal with outliers and provide a generalized robust model. LUCCK [12] is a recently developed classification method that highlighted on its ability of complex data. The algorithm could use vital feature-specific information to determine the complex pattern of changes in the data with adjusted concavity or convexity of similarity function, and then adjust the importance of each feature for the classifier. SVM and RF give fixed weight for each feature across all individuals, and LUCCK, which gives dynamic weight to each feature depending on the context of the prediction target.

3. RESULTS

We implemented three machine learning models by incorporating four feature selection methods. We calculated six metrics to compare these models and summarized the results in heatmap tables. Figure 1 shows the heatmap comparison of LUCCK, SVM and RF predictive models with each feature selection method.

AUC	MRMR	Chi-Square	RFE	LASSO
LUCCK	0.82 ± 0.17	0.8 ± 0.17	0.78 ± 0.16	0.79 ± 0.17
SVM	0.89 ± 0.12	0.84 ± 0.12	0.76 ± 0.15	0.92 ± 0.08
RF	0.80 ± 0.12	0.82 ± 0.16	0.76 ± 0.14	0.78 ± 0.16

Sensitivity	MRMR	Chi-Square	RFE	LASSO
LUCCK	0.83 ± 0.23	0.83 ± 0.22	0.82 ± 0.22	0.77 ± 0.24
SVM	0.86 ± 0.18	0.77 ± 0.2	0.75 ± 0.22	0.85 ± 0.19
RF	0.62 ± 0.22	0.65 ± 0.19	0.61 ± 0.24	0.51 ± 0.24

Precision	MRMR	Chi-Square	RFE	LASSO
LUCCK	0.64 ± 0.16	0.63 ± 0.17	0.61 ± 0.15	0.65 ± 0.21
SVM	0.72 ± 0.15	0.75 ± 0.21	0.70 ± 0.19	0.76 ± 0.18
RF	0.80 ± 0.19	0.75 ± 0.24	0.74 ± 0.23	0.75 ± 0.28

F1 score	MRMR	Chi-Square	RFE	LASSO
LUCCK	0.70 ± 0.16	0.51 ± 0.12	0.68 ± 0.15	0.68 ± 0.18
SVM	0.77 ± 0.12	0.69 ± 0.16	0.70 ± 0.18	0.77 ± 0.13
RF	0.67 ± 0.17	0.66 ± 0.17	0.63 ± 0.18	0.58 ± 0.21

Specificity	MRMR	Chi-Square	RFE	LASSO
LUCCK	0.54 ± 0.23	0.52 ± 0.25	0.49 ± 0.27	0.59 ± 0.23
SVM	0.66 ± 0.23	0.68 ± 0.30	0.69 ± 0.20	0.69 ± 0.29
RF	0.83 ± 0.16	0.75 ± 0.23	0.76 ± 0.20	0.83 ± 0.17

Accuracy	MRMR	Chi-Square	RFE	LASSO
LUCCK	0.68 ± 0.15	0.66 ± 0.16	0.65 ± 0.16	0.67 ± 0.17
SVM	0.76 ± 0.13	0.72 ± 0.13	0.72 ± 0.16	0.77 ± 0.14
RF	0.73 ± 0.11	0.70 ± 0.14	0.68 ± 0.14	0.68 ± 0.14

AUPRC	MRMR	Chi-Square	RFE	LASSO
LUCCK	0.52 ± 0.11	0.51 ± 0.12	0.49 ± 0.12	0.51 ± 0.13
SVM	0.57 ± 0.09	0.53 ± 0.10	0.48 ± 0.11	0.59 ± 0.06
RF	0.51 ± 0.09	0.52 ± 0.13	0.48 ± 0.10	0.49 ± 0.12

Figure 1. Heatmaps of the performance of the feature selection methods and machine learning approaches tested. The color code dark green to red indicated respectively lower to improved performance.

We evaluated the predictive risk of TMJ OA cure/improvement or persistence/worsening of clinical comorbidities, imaging features extracted from the articular fossa, condyle, and joint space as well as biological features. The best performance was obtained with the SVM model using LASSO that achieved AUC, sensitivity and precision of 0.92 ± 0.08 , 0.85 ± 0.19 and 0.76 ± 0.28 , respectively.

Figure 2 allows the visualization of the predictive performance separately for each feature selection method and machine learning approach. LASSO and MRMR are the top performing feature selection methods when considering the AUC, F1 score and sensitivity. SVM presents stronger performance in terms of AUC, F1 score, sensitivity, and accuracy with slightly lower precision than RF. LUCCK outperformed RF in items of sensitivity using any feature selection method.

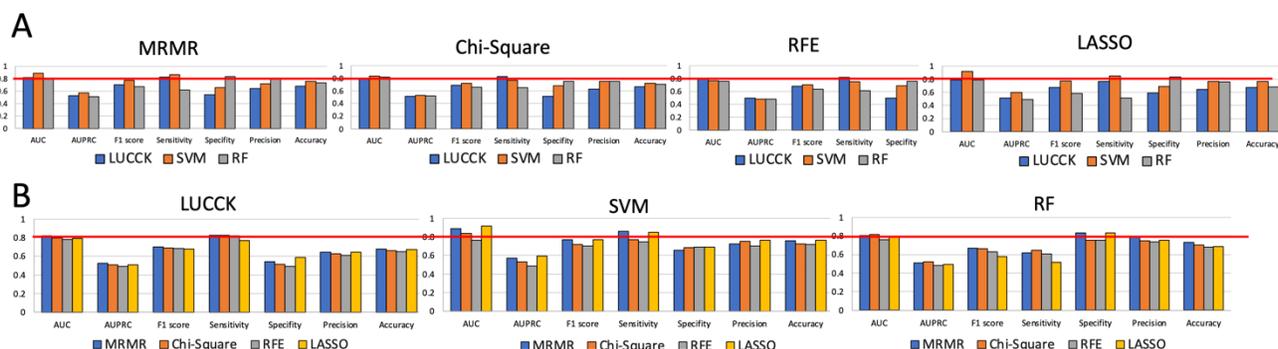


Figure 2. A. Graphic visualization of performance of each feature selection method. B. Graphic visualization of the performance of each machine learning approach.

To interpret the prediction of our proposed model, we utilized feature occurrence which calculates the number of times a feature is selected by the SVM model using LASSO among the total 50 models. The more often a feature occurs, the more reliable its importance is. Contributing features are shown in Table 1 according to the feature occurrence, which indicates the impact of each feature on the model performance. According to the order of features in Table 1, higher values of headache, LongRunHighGreyLevelRunEmphasis and muscle soreness increased the probability of assigning TMJ OA diagnosis. Baseline levels of headaches, lower back pain, restless sleep, muscle soreness, articular fossa bone surface/bone volume and trabecular separation, condylar HighGreyLevelRunEmphasis and ShortRunHighGreyLevelEmphasis, saliva levels of 6Ckine, OPG and Angiogenin, and serum levels of 6ckine and BDNF were the most frequently occurring features in the SVM model using LASSO to predict more severe TMJ OA prognosis.

	Clinical comorbidities				Imaging				Biological					
	Headaches	Lower Back Pain	Restless Sleep	Muscle Soreness	Articular Fossa		Condyle		Saliva				Serum	
					Bone Surface/ Bone Volume	Trabecular Separation	HighGreyLevel RunEmphasis	ShortRunHigh GreyLevelEmphasis	OPG	6Ckine	Angiogenin	ENA78	6Ckine	BDNF
Times of occurrence	50	46	11	8	34	8	32	9	39	38	28	15	41	8

Table 1. Times of occurrence showing how many times a feature was selected by the SVM model using LASSO.

4. CONCLUSION

In this paper we proposed three machine learning models and four feature selection methods to predict the progression of TMJ OA disease. We designed 5-fold cross validation and calculated the feature occurrence to determine which features have stronger predictive power. The SVM predictive model of TMJ OA using LASSO for feature selection achieved AUC, sensitivity and precision of 0.92 ± 0.08 , 0.85 ± 0.19 and 0.76 ± 0.18 , respectively. Baseline levels of headaches, lower back pain, restless sleep, muscle soreness, articular fossa bone surface/bone volume and trabecular separation, condylar HighGreyLevelRunEmphasis and ShortRunHighGreyLevelEmphasis, saliva levels of 6Ckine, OPG and Angiogenin, and serum levels of 6ckine and BDNF were the most frequently occurring features in the SVM model using LASSO to predict more severe TMJ OA prognosis.

REFERENCES

- [1] Abrahamsson, A. K., Kristensen, M., Arvidsson, L. Z., Kvien, T. K., Larheim, T. A., & Haugen, I. K. "Frequency of temporomandibular joint osteoarthritis and related symptoms in a hand osteoarthritis cohort". *Osteoarthritis and cartilage*, 25(5), 654-657 (2017).
- [2] O'Connor, R. C., Fawthrop, F., Salha, R., & Sidebottom, A. J. "Management of the temporomandibular joint in inflammatory arthritis: Involvement of surgical procedures". *European journal of rheumatology*, 4(2), 151 (2017).
- [3] Goldring, S. R., & Goldring, M. B. "Changes in the osteochondral unit during osteoarthritis: structure, function and cartilage–bone crosstalk". *Nature Reviews Rheumatology*, 12(11), 632-644 (2016).
- [4] Cevidanes, L. H. S., Hajati, A. K., Paniagua, B., Lim, P. F., Walker, D. G., Palconet, G., ... & Phillips, C. "Quantification of condylar resorption in temporomandibular joint osteoarthritis". *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology*, 110(1), 110-117 (2010).
- [5] Cevidanes, L. H., Walker, D., Schilling, J., Sugai, J., Giannobile, W., Paniagua, B., ... & Styner, M. "3D osteoarthritic changes in TMJ condylar morphology correlates with specific systemic and local biomarkers of disease". *Osteoarthritis and Cartilage*, 22(10), 1657-1667 (2014).
- [6] Ding, C., & Peng, H. "Minimum redundancy feature selection from microarray gene expression data". *Journal of bioinformatics and computational biology*, 3(02), 185-205 (2005).
- [7] Yang, Y., & Pedersen, J. O. "A comparative study on feature selection in text categorization". In *ICML*, Vol. 97, No. 412-420, p. 35. (1997).
- [8] Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. "Gene selection for cancer classification using support vector machines". *Machine learning*, 46(1), 389-422 (2002).
- [9] Muthukrishnan, R., & Rohini, R. "LASSO: A feature selection technique in predictive modeling for machine learning". In *2016 IEEE international conference on advances in computer applications (ICACA)* (pp. 18-20). IEEE. (2016).
- [10] Trafalis, T. B., & Ince, H. "Support vector machine for regression and applications to financial forecasting". In *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium* (Vol. 6, pp. 348-353). IEEE. (2000).
- [11] Maimon, O. Z., & Rokach, L. "Data mining with decision trees: theory and applications" (Vol. 81). World scientific. (2014).
- [12] Sabeti, E., Gryak, J., Derksen, H., Biwer, C., Ansari, S., Isenstein, H., ... & Najarian, K. "Learning using concave and convex kernels: applications in predicting quality of sleep and level of fatigue in fibromyalgia". *Entropy*, 21(5), 442. (2019).