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

Najla Al Turkestani (alnajla@umich.edu)
University of Michigan
Lingrui Cai
University of Michigan
Lucia Cevidanes
University of Michigan
Jonas Bianchi
University of the Pacific
Winston Zhang
University of Michigan
Marcela Gurgel
University of Michigan
Maxime Gillot
University of Michigan
Baptiste Baquero
University of Michigan
Kayvan Najarian
University of Michigan
Reza Soroushmehr
University of Michigan

Research Article

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

Posted Date: July 14th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1840348/v1

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.

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 (Fig. 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.

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 $(x_1,y_1),...,(x_l,y_l)$, where $x_i \in X$, $y_i \in \{-1,1\}$, $i = 1,...,l$, 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 $f(x,\alpha)$, $\alpha \in \wedge$, 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

$(x_1,x*_1,y_1),...,(x_l,x*_l,y_l)$, $x_i \in X$, $x*_i \in X*$, $y_i \in \{-1,1\}$, $i = 1,...,l$, 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 (Fig. 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.

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.

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 A). Shapley values were used to interpret the contribution of individual features into the prediction of the trained model. Contributing features are shown in Fig. 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.

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 (Fig. 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$\sqrt{}$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. Importanly, 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; 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. 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/ 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., Giacovelli, 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*Biology*Physics **94**(5), 1121–1128 (2016). https://doi.org/10.1016/ 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., Gon¸calves, 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/ 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., Gon¸calves, 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$_7$*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. 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)

Figures

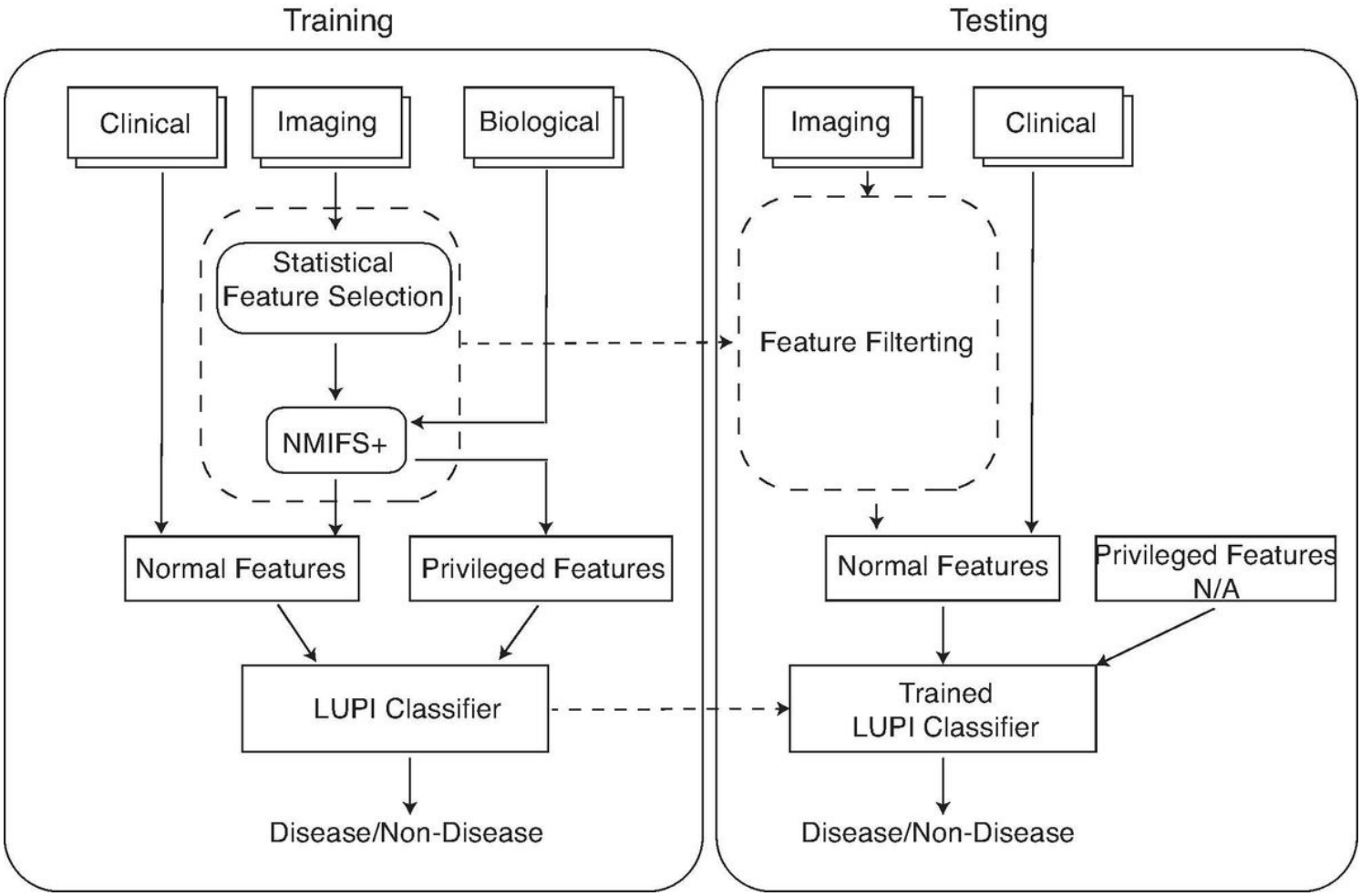

Figure 1

Diagram of training and testing process

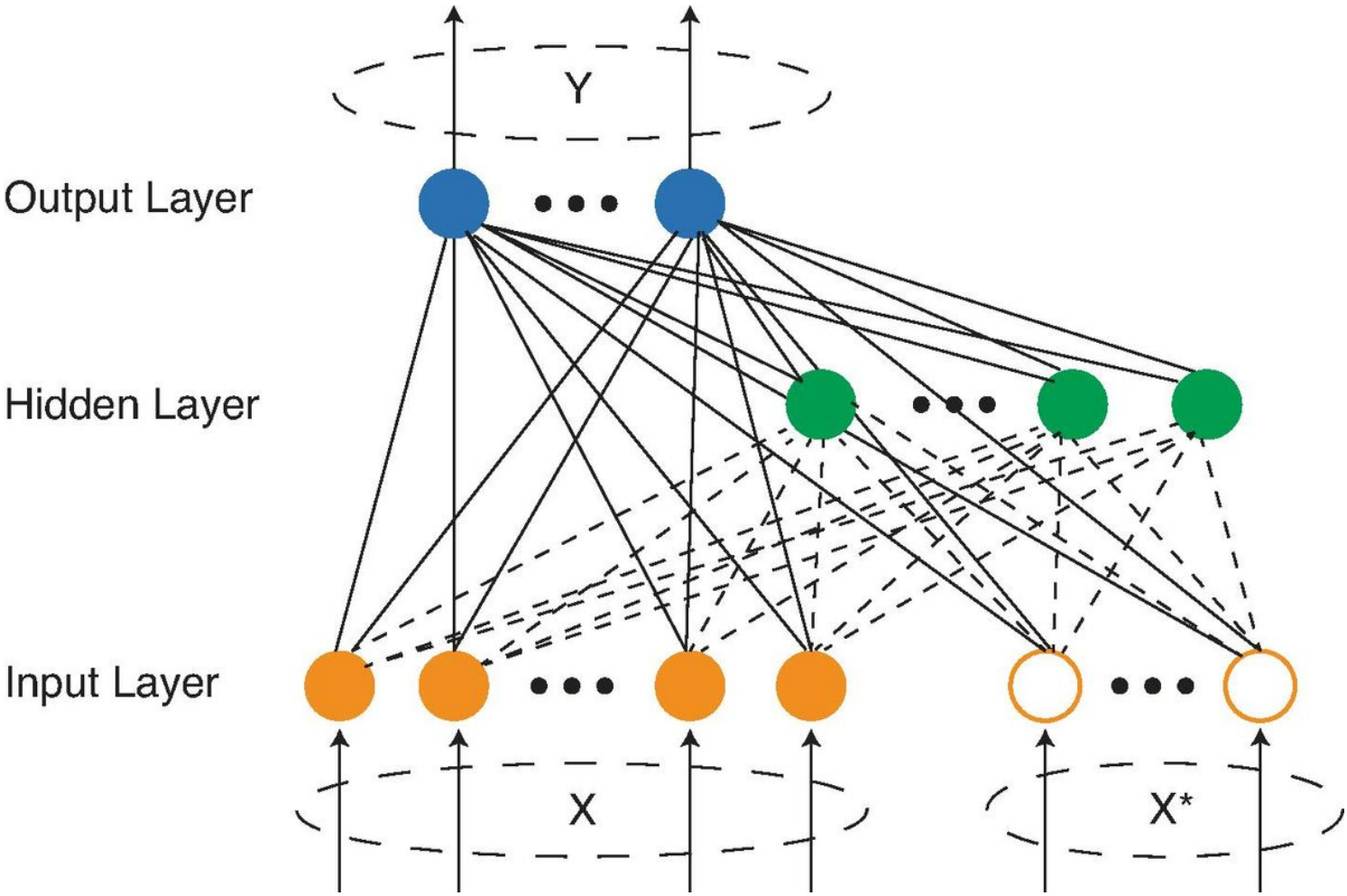

Figure 2

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

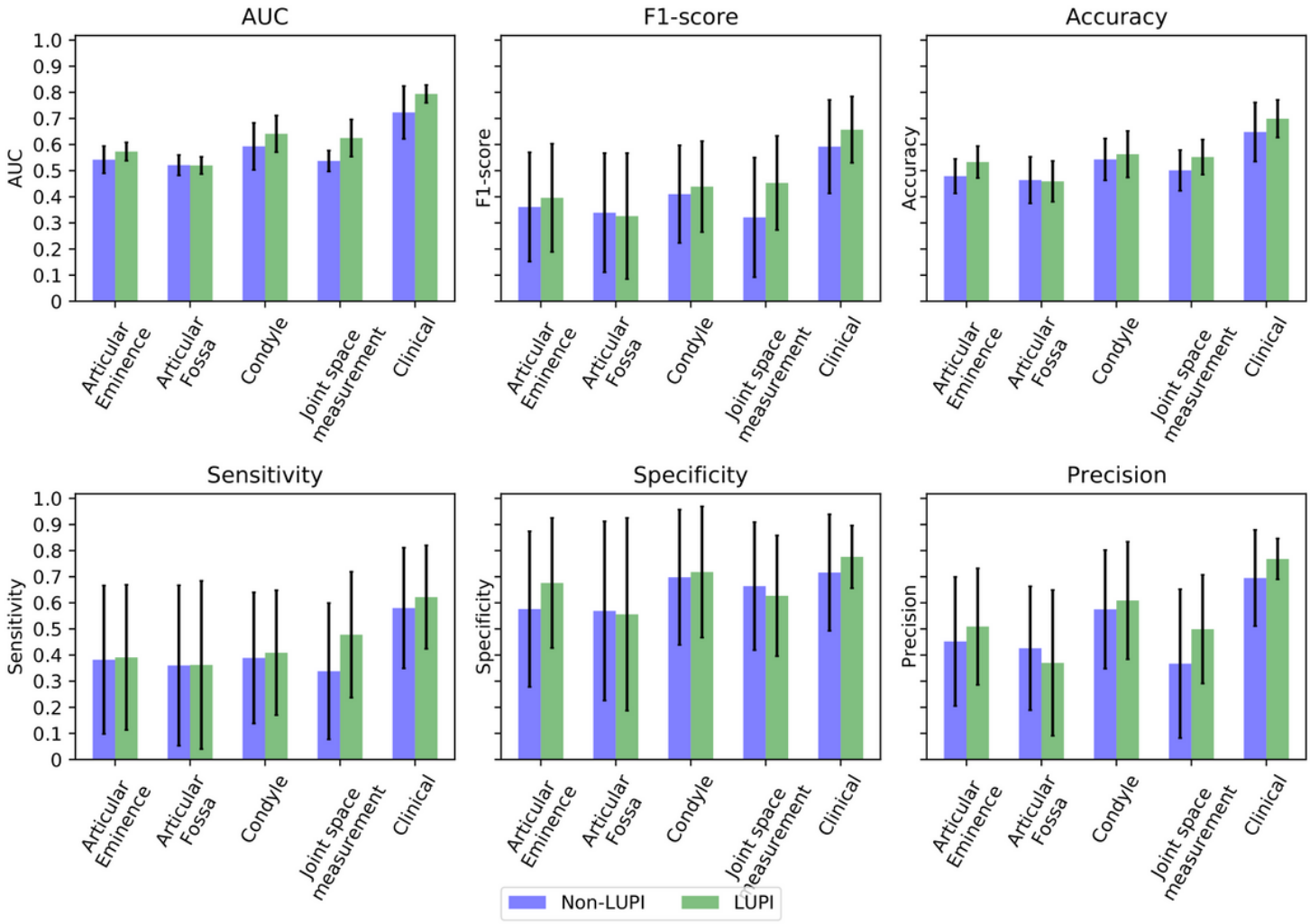

Figure 3

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.

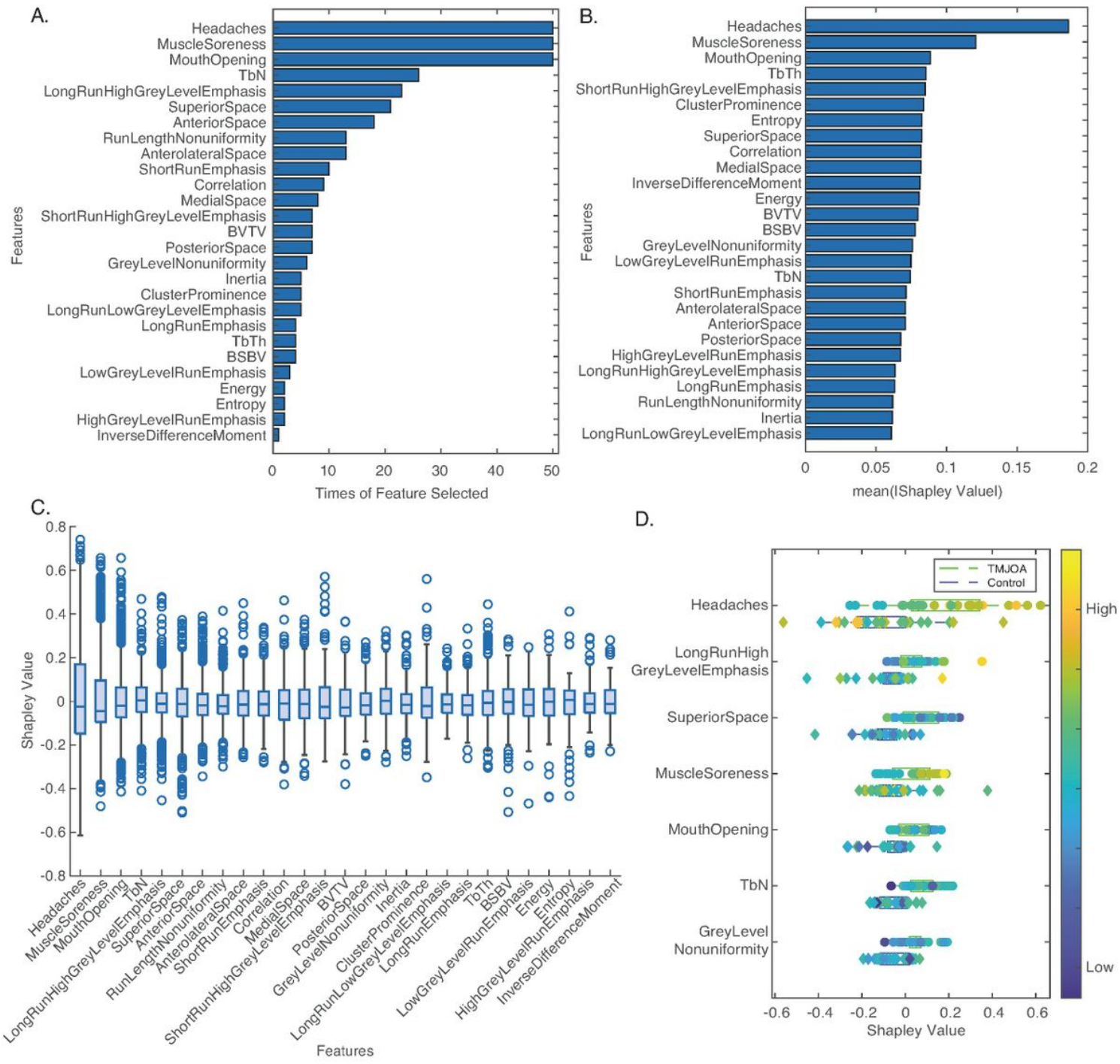

Figure 4

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.