

RESEARCH ARTICLE

Artificial intelligence and infrared thermography as auxiliary tools in the diagnosis of temporomandibular disorder

¹Elisa Diniz de Lima, ²José Alberto Souza Paulino, ¹Ana Priscila Lira de Farias Freitas,

¹José Eraldo Viana Ferreira, ¹Jussara da Silva Barbosa, ¹Diego Filipe Bezerra Silva,

¹Patrícia Meira Bento, ¹Ana Marly Araújo Maia Amorim and ¹Daniela Pita Melo

¹Department of Dentistry, State University of Paraíba, Campina Grande, Paraíba, Brazil; ²Computer Science, Federal University, Campina Grande, Brazil

Objective: To assess three machine learning (ML) attribute extraction methods: radiomic, semantic and radiomic-semantic association on temporomandibular disorder (TMD) detection using infrared thermography (IT); and to determine which ML classifier, KNN, SVM and MLP, is the most efficient for this purpose.

Methods and materials: 78 patients were selected by applying the Fonseca questionnaire and RDC/TMD to categorize control patients (37) and TMD patients (41). IT lateral projections of each patient were acquired. The masseter and temporal muscles were selected as regions of interest (ROI) for attribute extraction. Three methods of extracting attributes were assessed: radiomic, semantic and radiomic-semantic association. For radiomic attribute extraction, 20 texture attributes were assessed using co-occurrence matrix in a standardized angulation of 0°. The semantic features were the ROI mean temperature and pain intensity data. For radiomic-semantic association, a single dataset composed of 28 features was assessed. The classification algorithms assessed were KNN, SVM and MLP. Hopkins's statistic, Shapiro–Wilk, ANOVA and Tukey tests were used to assess data. The significance level was set at 5% ($p < 0.05$).

Results: Training and testing accuracy values differed statistically for the radiomic-semantic association ($p = 0.003$). MLP differed from the other classifiers for the radiomic-semantic association ($p = 0.004$). Accuracy, precision and sensitivity values of semantic and radiomic-semantic association differed statistically from radiomic features ($p = 0.008$, $p = 0.016$ and $p = 0.013$).

Conclusion: Semantic and radiomic-semantic-associated ML feature extraction methods and MLP classifier should be chosen for TMD detection using IT images and pain scale data. IT associated with ML presents promising results for TMD detection.

Dentomaxillofacial Radiology (2022) **51**, 20210318. doi: [10.1259/dmfr.20210318](https://doi.org/10.1259/dmfr.20210318)

Cite this article as: Diniz de Lima E, Souza Paulino JA, Lira de Farias Freitas AP, Viana Ferreira JE, Barbosa JS, Bezerra Silva DF, et al. Artificial intelligence and infrared thermography as auxiliary tools in the diagnosis of temporomandibular disorder. *Dentomaxillofac Radiol* 2022; **51**: 20210318.

Keywords: Temporomandibular Joint Disorders; Thermography; Artificial Intelligence

Introduction

Temporomandibular disorders (TMD) are characterized by a set of musculoskeletal and neuromuscular disorders that affect the temporomandibular joint

(TMJ).^{1,2} Masticatory muscle disorders are the most prevalent subgroup of TMD and are characterized by trigger points affecting more than one muscle group and myofascial pain associated with muscle spasm and tenderness during palpation.^{3,4}

Correspondence to: Dr Daniela Pita Melo, E-mail: danipita@gmail.com

Received 07 July 2021; revised 14 September 2021; accepted 23 September 2021

TMD diagnosis is based on the clinical history and clinical examination²; and its detection and classification can be challenging. The Research Diagnostic Criteria for Temporomandibular Disorders (RDC/TMD) is a widely used diagnostic protocol,^{5,6} which requires a certain amount of training and is time- and resource-consuming. A precise TMD diagnosis and effective therapy improves the quality of patients' lives significantly.⁴

Computed tomography (CT), cone-beam computed tomography (CBCT) and magnetic resonance images (MRI) are often indicated as complementary image exams to assist TMD diagnosis and management.⁷ However, these imaging modalities are more useful for evaluating cases of degenerative alterations and disc displacements, not providing information on local microcirculation, which is important in cases of masticatory muscle disorders.⁸

Infrared Thermography (IT) is a promising non-ionizing and non-invasive screening method for measuring skin temperature and temperature variation remotely.^{7,9} Several studies have investigated the use of IT on TMD diagnosis^{7,8,10-16} presenting controversial results. The interpretation of IT exceedingly small temperature variations (0.1–0.5°C) may be a reason for its divergent results. Therefore, this image modality may benefit from computational analytical tools.^{4,17,18} Computer-aided interpretation of thermograms is of great relevance and may improve the information assessed by IT, since the link between the disease and heat pattern is subtle and usually non-linear.

Artificial intelligence (AI) is defined as the ability of a machine to emulate intelligent human behavior to execute complex tasks.¹⁹ In general, computerized methods often employ data mining and machine learning (ML) algorithms, which are considered computer-aided diagnostic tools that assist doctors in making diagnostic decisions.²⁰ Therefore, a computer-aided approach has three benefits: reduced interobserver variability, AI may outperform humans on specific tasks, and the performance of these objective systems can be measured.²¹

The use of AI is progressing on dental practice based on semantic features (data from clinical exams and auxiliary tools) and radiomic features (image data).²² Radiomic features are image-hidden quantitative information, which can be extracted with the aid of algorithms called attribute extractors or feature extractors. Radiomic features represent patterns quantitatively in an image and can be classified into three types of descriptors: shape, texture and frequency.²³ Radiomic features can be associated with semantic features (qualitative data) for a more robust characterization of a given pattern.

ML is one of the subfields of AI that allows a computer model to learn and predict by recognizing patterns.¹⁹ ML classifiers are algorithm classification techniques widely used in machine learning. K-Nearest Neighbors (KNN) classification algorithm is one of

the most used non-parametric classification methods to determine a sample's classification label based on neighboring samples; however, it is limited due to high memory consumption when using large datasets.²⁴ Support Vector Machine (SVM) is a computer algorithm that learns by pattern to assign labels to objects²⁵ and needs long training time for large datasets.²⁶ Artificial neural networks (ANNs) are non-linear models inspired by the neural architecture of the brain and a typical ANN architecture is a multilayer perceptron (MLP) that contains a series of layers, composed of neurons and their connections.²⁷ When assessing ML classifiers, it is important to establish which one shows higher accuracy for the assessed diagnostic task, image modality and type of extracted data (radiomic or semantic).

Therefore, this study aimed to assess three ML attribute extractor methods: radiomic, semantic and the association of semantic and radiomic methods on TMD detection using IT; and to determine which ML classifier, KNN, SVM and MLP, is the most efficient for this purpose.

Material and methods

This cross-sectional observational study was approved by the University's Ethics and Research Committee (protocol n° 73417017.2.0000.5187) and follows the Helsinki Declaration.

Sample selection

IT exams were performed in 41 patients with temporomandibular disorders and 37 patients without temporomandibular disorders (control group), referred to the University's Orofacial Pain clinic. TMD patients and controls were matched for age, sex and body mass index (BMI). The control group was composed of volunteers who did not present TMD signs and symptoms, according to the Fonseca questionnaire during the first patient screening to separate controls from TMD patients. A second screening was done using RDC/TMD (Axes I and II) to ensure the groups categorization.²⁸ The selected patients should be 18 to 60 years old.

Patients with toothache, fever, systemic changes (hypoglycemia, hypothyroidism or hyperthyroidism, hypertension, respiratory diseases, rheumatoid arthritis, fibromyalgia, pregnancy, rheumatological disorders, menopause, neurological disorders); cancer patients; patients under therapy with myorelaxant medication, analgesic and/or anti-inflammatory, or who underwent hormone replacement; and patients who had facial scars or papules were excluded from this study.

Palpation and pain scale assessment

A specialist in TMD and Oral facial pain with 5 years of experience on RDC/TMD examination executed the Orofacial pain examinations. The diagnosis of TMD was performed using RDC/TMD (Axes I and II).²⁸

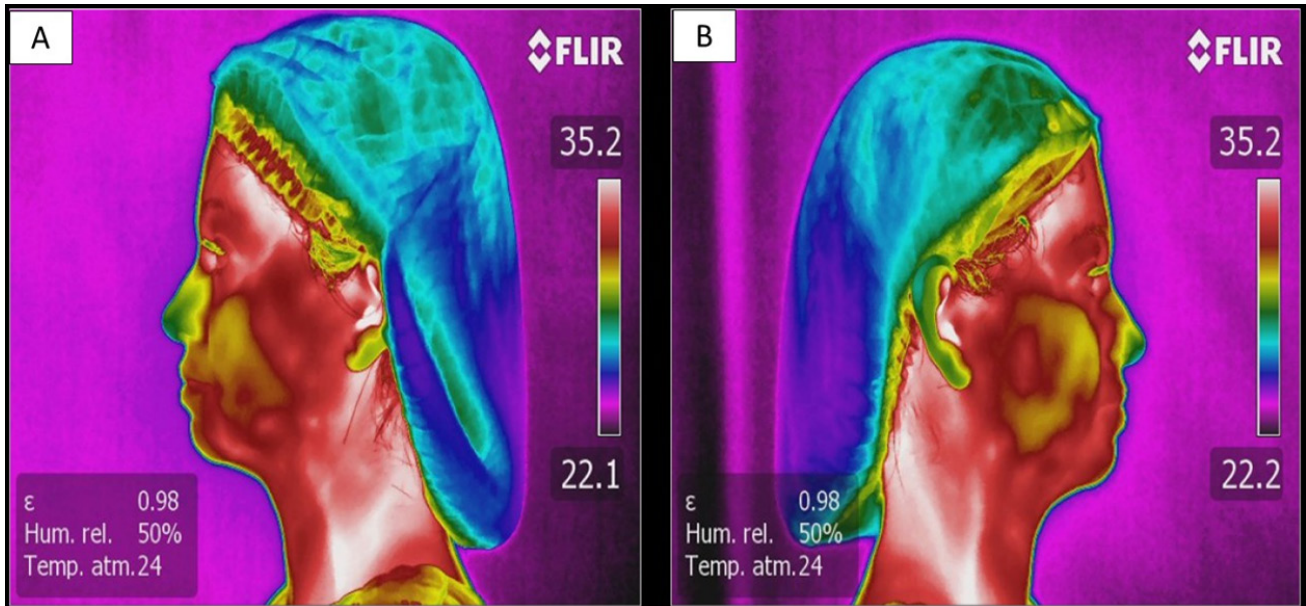


Figure 1 (A) left side-lateral norm thermogram; (B) right side-lateral norm thermogram of the same patient.

Palpation was performed, and pain intensity data were collected according to the following characterization: 0 – without pain; 1 – slight pain; 2 – moderate pain; 3 – severe pain, in accordance with the RDC / TMD.

IT acquisition

For the acquisition of IT exams, a handheld FLIR model T650 infrared sensor camera (FLIR Systems, Wilsonville, USA), with 25 mm lens and 640×480 pixels spatial resolution, capable of capturing images with temperature variation between -40 and 150°C and thermal sensitivity of 0.05 to 30°C was used. The thermographic camera was attached to a tripod and switched on 20 min prior image acquisition.

All images were acquired in the University's Infrared Thermography Laboratory in a room with standardized temperature (22 to 24°C) and no windows. Cold-cathode fluorescent lamps were used for ambient lighting and the walls where the patient's chair was positioned were covered with 25-mm-thick Expanded Polystyrene (EPS-Styrofoam) plates, aluminum foil and Ethylene-vinyl acetate (EVA) in black color.

The exams were performed by an Oral and Maxillofacial Radiologist with 10 years' experience blinded to the clinical information. The image exams were acquired in the lateral (right and left sides) norm (Figure 1) following the recommendations of the American Academy of Thermology.²⁹ All patients received written and oral instructions one day prior the IT exam (do not apply make-up or lotion to the face; do not use sources of heat such as hairdryers or hair straighteners; do not take analgesics, corticoids, anti-inflammatory drugs; do not carry out any kind of physical exercise; and do not touch, rub, or scratch the area of skin that is to be examined). The patients should wait 15 min in

the examination room with an average temperature of $23 \pm 1^\circ\text{C}$ and relative humidity between 40 and 60% to achieve the recommended thermal balance.³⁰

The patient was seated in a swivel-chair, maintaining an erect posture, with the sagittal plane perpendicular to the ground and camper plane parallel to the ground. A standardized distance of 0.80 m between the camera and the patient was set for all image acquisitions. The emissivity of the skin was set at 98%.^{13,14,16,31,32} To minimize the reflection of thermal radiation, only the patient and the operator should be present in the room during image acquisition.

156 thermographic images were acquired and stored in individualized folders. Subsequently, all images were evaluated by the Oral and Maxillofacial Radiologist with the aid of the *FLIR Tools v.6.4* software (FLIR Systems, Wilsonville, USA). The thermographic images were used for both semantics and radiomics attribute extractions.

The regions of interest (ROI) were selected in the image region of the temporal and masseter muscles on the right and left lateral norms, with two points in the temporal region and six points in the masseter region. The "circle" tool (diameter 22 ± 3 mm) was used to check the thermal gradient. Absolute averages of temperatures were obtained for each ROI. The average temperature values for each patient, muscle and ROI were tabulated in a spreadsheet for posterior analysis.

Thermographic images were also analyzed with the aid of AI using three ML methods of feature extraction. The first ML method was the extraction of radiomic features, which are the characteristics extracted from an image through mathematical formulas that describe quantitatively various characteristics related to the image texture within the ROI. The second ML method

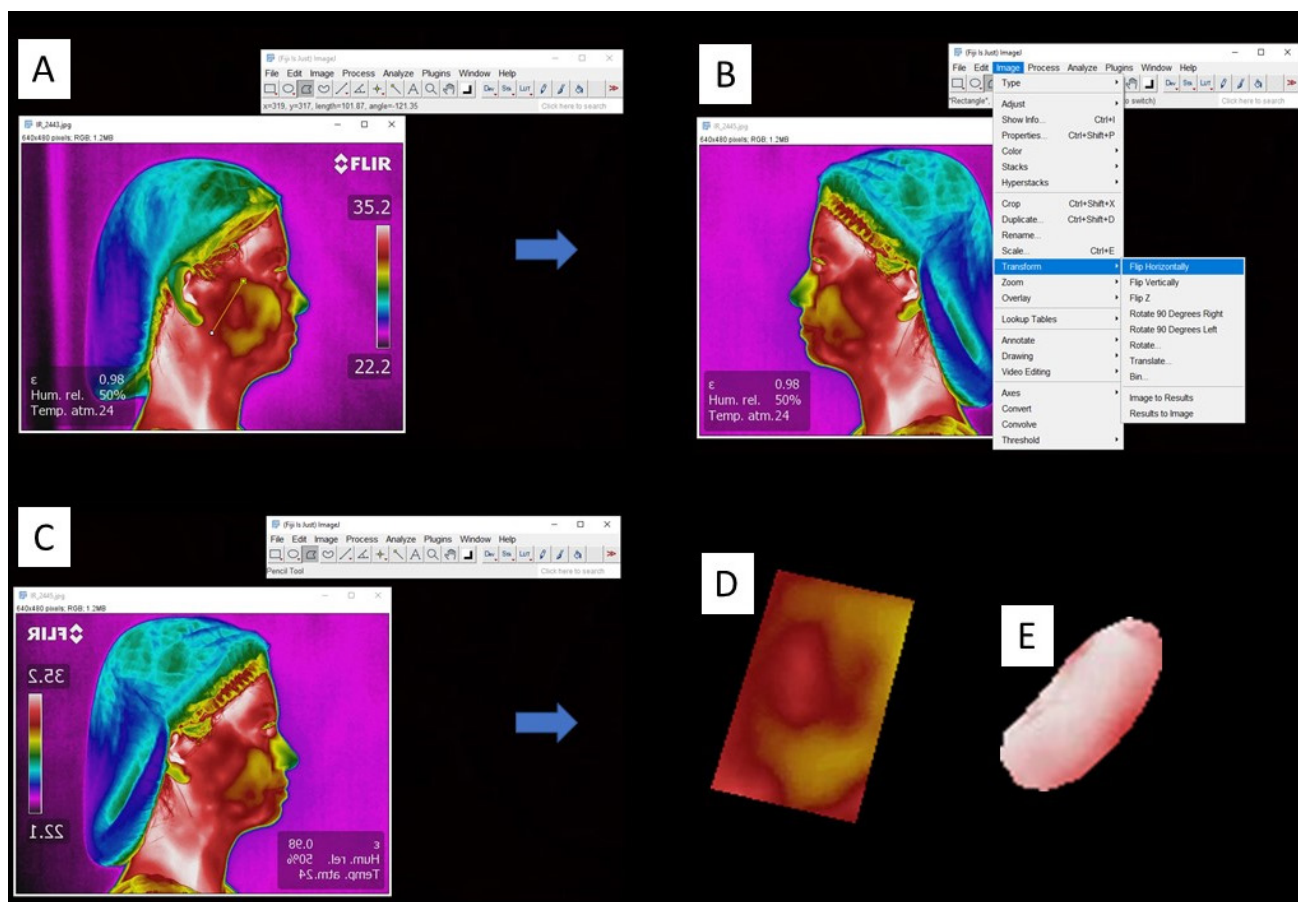


Figure 2 (A) Thermographic image assessed with Fiji ImageJ software using the “polygon” selection tool to select the ROI; (B) Standardization of the lateral right and left norms to the right-side-norm by using the tools “transform” e “flip horizontally”; (C) left-side-lateral norm transformed into right-side-lateral form; (D) ROI of the masseter muscle; (E) ROI of the temporal muscle.

used was the extraction of semantic features, which are qualitative characteristics that are collected according to a subjective assessment done by a radiologist or a health professional.^{33–35} The third method used was an association of radiomic and semantic features to create a single dataset.

To perform the technique of extracting image texture radiomic features, right and left lateral norm images were used (74 images for the control group and 82 images for the TMD group). With the aid of Fiji ImageJ software (64-bit Java 1.8.0_172, National Institutes of Health, Bethesda, MD), the regions corresponding to the masseter and temporal muscles were selected using the “polygon” selection tool. By selecting the “crop” tool, the ROI was removed from the image. To remove the image background, “make-inverse” and “clear” tools were used. The lateral right and left norms were standardized to the right-side-norm by using the tools “transform” e “flip horizontally”, to improve image assessment (Figure 2).

The semantic features were extracted from a set of tabulated data. A spreadsheet containing temperature data from the masseter and temporal regions obtained

using FLIR Tools v.6.4 software in addition to the categorized scale of pain on palpation data.

The third method used radiomic and semantic features extracted previously and transformed these features in a single dataset for further assessment.

Radiomic features (texture attributes)

The vector of attributes construction of each image was done initially with the segmentation of the ROI. For the definition of the ROI, manual segmentation was adopted, isolating the masseter and temporal areas separately, as previously described. This segmentation method is needed to precisely select only the relevant region of the thermogram and discard areas with potential for interference in the attribute classification, such as regions with facial hair.

To characterize the pathology (TMD) and healthy (control) patterns, 20 texture attributes were extracted from the co-occurrence matrices on a 0° standardized angulation. The attributes extracted from the temporal and masseter muscles were contrast, correlation, energy, homogeneity, entropy, trimmed means, kurtosis,

asymmetry, standard deviation, and variance. Therefore, a dataset of 156 samples was created and each sample had 20 attributes (ten attributes for each muscle assessed).

Semantic features (temperature and pain scale data attributes)

The construction of the temperature and pain scale data attribute vectors started during image acquisition. The temperature values were mapped using the infrared camera software (FLIR Tools v.6.4 software). The ROIs used to assess temperature values were the same used for radiomic features.

For each region, the corresponding pain intensity values were assessed in accordance with the RDC/TMD.²⁸ Finally, this extraction process generated a dataset with 156 samples in a tabular format, each sample with eight attributes (masseter temperature at its superior region, pain data at the masseter's superior region, masseter temperature at its middle region, pain data at the masseter's middle region, masseter temperature at its inferior region, pain data at the masseter's inferior region, anterior temporal temperature, and pain at the anterior temporal).

Association of radiomic and semantic attributes

To verify if the association of radiomic and semantic attributes could improve TMD detection, a dataset was created associating both ML methods. For each one of the 156 samples, 28 attributes were extracted. The attributes extracted from the temporal and masseter muscles were contrast, correlation, energy, homogeneity, entropy, trimmed mean, kurtosis, asymmetry, standard deviation, variance, temperature (superior, middle, and inferior masseter and temporal muscles), and pain (superior, middle, and inferior masseter and temporal muscles).

Data processing

The Principal Component Analysis (PCA) method was used to reduce the dimensionality and a 95% variance was considered to generate the new features; therefore, generic features were generated statistically based on the existing features. After applying PCA, the number of radiomic features was reduced from 20 to 8, the number of semantic features was reduced from 8 to 4 and the number of radiomic-semantic association features was reduced from 28 to 11.

Considering these datasets in a tabular format, it is understood that when applying the PCA a new set with the same number of samples (rows) as the original set will be obtained, however with fewer attributes (columns). The datasets are then partitioned into two subsets, following the holdout method, in 70/30 proportions. The subset with 70% was destined to the training stage and the 30% subset was used to carry out the tests. It is important to highlight that the stratification

adopted for this partitioning was proportionality guaranteed for both classes (healthy and pathology).

For subsets with 70% of the samples (training), cross-validation was used, which is a method that consists of using various combinations of the same dataset during training. This method is extremely useful to increase the generalizability of a model. In this research, 10 iterations were defined for resampling the training base via cross-validation.

Classification algorithms (classifiers)

Kernel nearest neighbor (KNN), support vector machine (SVM) and multilayer perceptron (MLP) classifiers were implemented and compared.

The metric adopted to evaluate the classification algorithms was accuracy, which corresponds to the level of correctness of a model. In addition to accuracy, precision, sensitivity, and specificity were assessed.

KNN classifier

The KNN aims to determine a sample's classification label based on neighboring samples. The Kerner value corresponds to the number of neighbors that the algorithm must consider performing the classification. For the construction of this model, kernel value five was used. This value was defined after an iterative execution of the classification varying the k (kernel) from 1 to 50 and comparing the accuracy of the model training in each k value.

SVM classifier

In the SVM technique, classes are separated by a line named the decision boundary dividing the space between classes which must be as large as possible. The greater the space of the decision boundary, the better the classification of the model. Due to the non-linearity of the data used in this research, a degree three polynomial function was used to define the decision boundary. Polynomials of different degrees were also tested, but there was no gain in accuracy.

Multilayer Perceptron (MLP)

In addition to KNN and SVM, a classifier was also created using MLP. The first layer corresponded to the input layer, the second layer has eight neurons and the third and fourth layers have six neurons, the fifth has four neurons and the last layer corresponded to the output. Therefore, for the radiomic, semantic and the radiomic-semantic association features, the MLP had 8, 4, and 11 neurons in the input layer, respectively (Figures 3–5). The input layer corresponds to the number of features after principal components analyses (PCA).

Data analysis

To prove the clustering tendency of the datasets, Hopkins's statistic was used. This statistical test has the threshold of 0.5 ($H = 0.5$) as reference, and the resultant statistical data will be a value between 0 and 1. The closer the H value is to 0, the greater the clustering

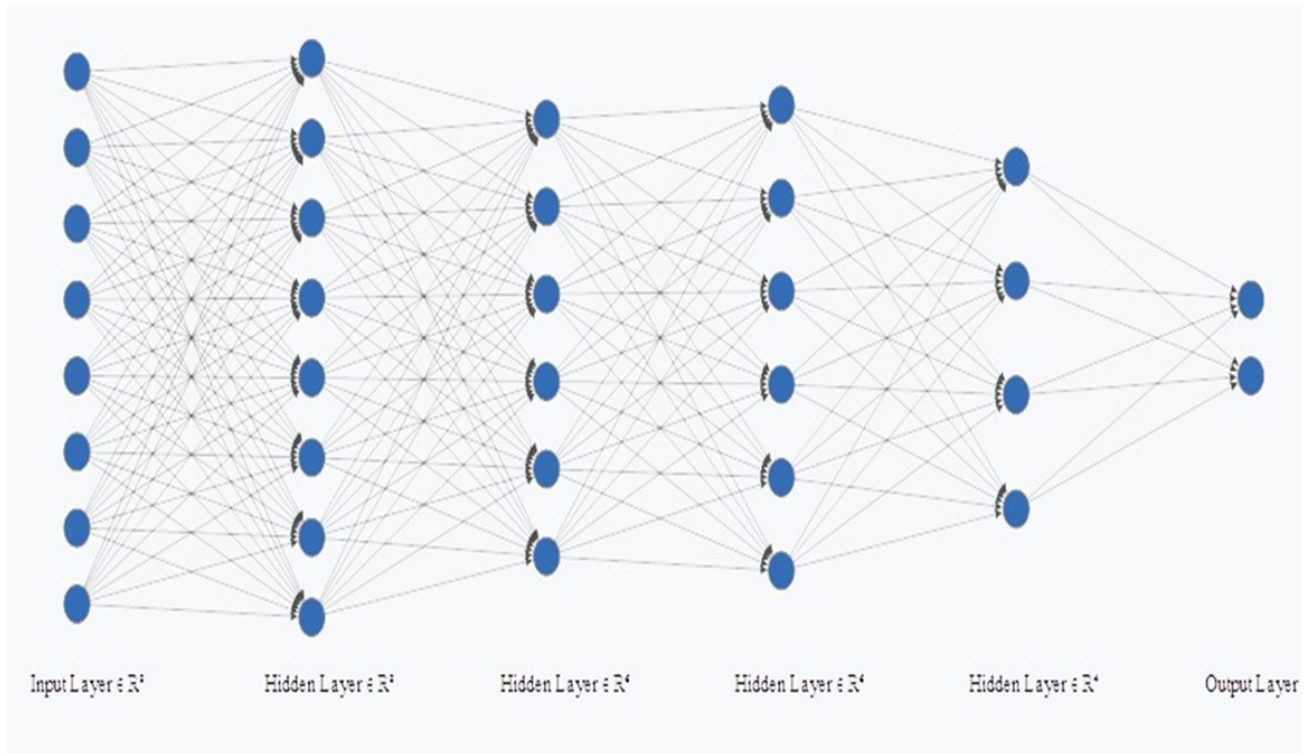


Figure 3 Multilayer Perceptron (MLP) for radiomic features.

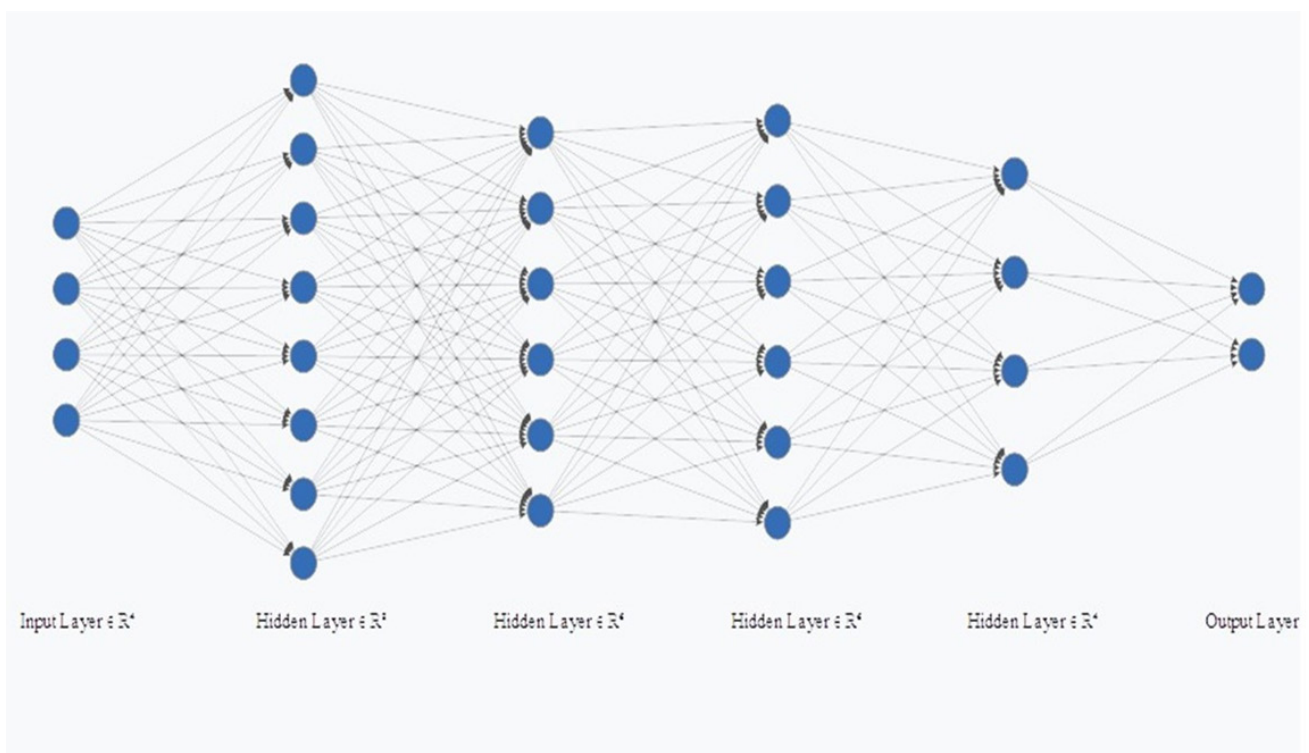


Figure 4 Multilayer Perceptron (MLP) for semantic features.

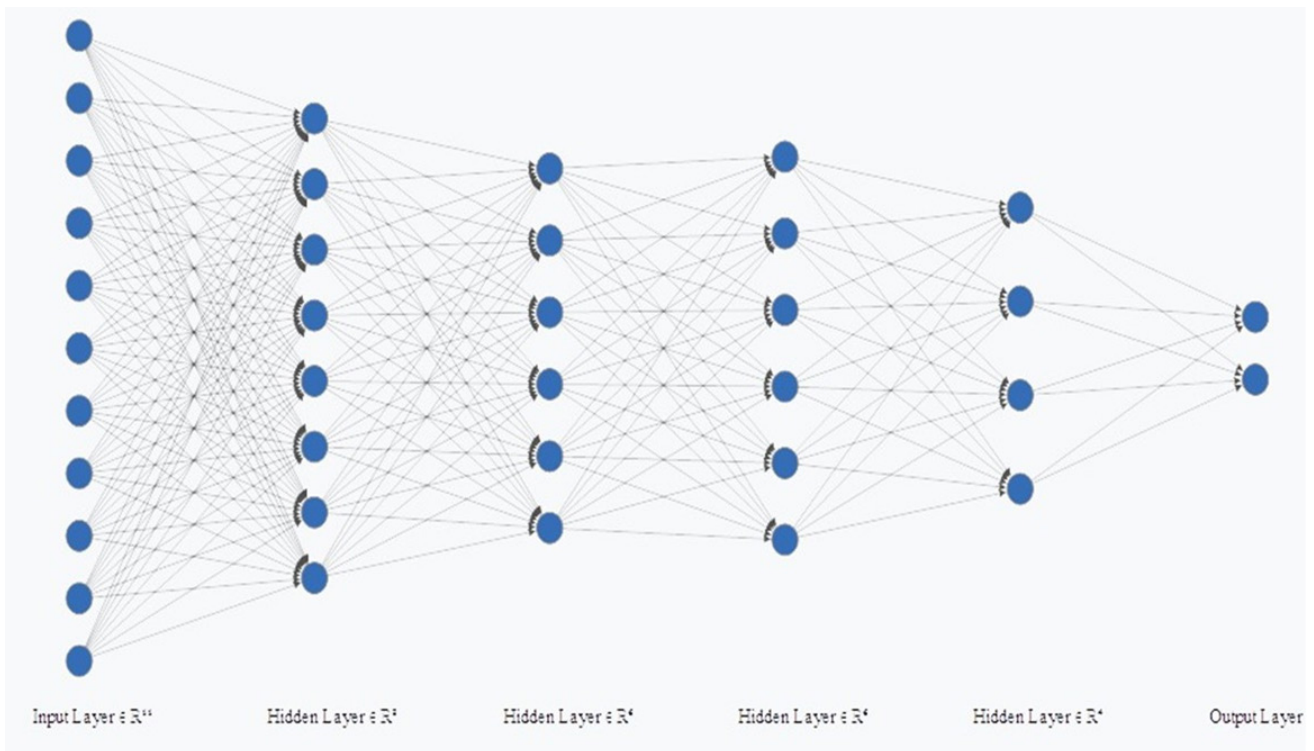


Figure 5 Multilayer Perceptron (MLP) for radiomics and semantics associated features.

tendency of the dataset, and the closer the H value is to the threshold, the greater the randomness of the data distribution (lower tendency of clustering).

Data analysis was done using the Jamovi software (v.1.6, 2021, Sydney, Australia). The Shapiro–Wilk test was used to assess data distribution, which was found to be normal for all groups ($p > 0.05$). For each tested ML method and classifier, accuracy, precision (positive predictive value), sensitivity (true positive rate), and specificity (true negative rate) were calculated. ANOVA for repeated measures and Tukey test were used to compare the studied variables.

Results

The H value obtained in this study for the radiomic, semantic, and radiomic-semantic association features

were 0.19, 0.21, and 0.16, respectively. Therefore, there is a cluster tendency for the studied datasets.

Table 1 shows the accuracy values of the classifiers for the training and testing assessments for each studied feature extraction method. The training and testing accuracy only differed statistically for the radiomic-semantic associated features ($p = 0.003$). MLP presented the best accuracy values and differed from SVM and KNN for the radiomic-semantic associated features ($p = 0.004$). Although SVM presented the best accuracy values for radiomic and semantics features assessed separately, it did not differ from KNN and MLP (Table 1).

When assessing and comparing the testing accuracy, precision and sensitivity values of each feature extraction method, semantic features and radiomic-semantic features associated differed statistically from radiomic features ($p = 0.008$, $p = 0.016$ and $p =$

Table 1 Accuracy values of the ML classifiers for the training and testing datasets and the studied feature extraction methods

ACCURACY	RADIOMIC FEATURES			p-value	SEMANTICS FEATURES			p-value	ASSOCIATED FEATURES			p-value
	KNN ^A	SVM ^A	MLP ^A	Inter	KNN ^A	SVM ^A	MLP ^A	Inter	KNN ^A	SVM ^B	MLP ^C	Inter
TRAINING	71.54% ^a	79.52% ^a	90.90% ^a	0.695	89.90% ^a	89.91% ^a	91.95% ^a	0.499	89% ^a	93.63% ^a	98.68% ^a	0.004
TESTING	68.08% ^a	76.59% ^a	63.82% ^a	0.296	87.23% ^a	95.74% ^a	91.48% ^a	0.758	82.97% ^b	87.23% ^b	91.49% ^b	0.003

^AANOVA test for repeated measures and post Tukey test.

^BIntra corresponds to KNN, SVM and MLP.

^CInter corresponds to training and testing.

^DDifferent letters indicate statistically significant differences between groups.

Table 2 Testing dataset accuracy, precision, sensitivity, and specificity values for the studied feature extraction methods for each assessed classifier

		<i>KNN</i>	<i>SVM</i>	<i>MLP</i>	<i>MEAN</i>	<i>SD</i>	<i>p-value</i>
ACCURACY	RADIOMICS^A	68.08%	76.59%	63.82	69.50%	6.5	0.008
	SEMANTICS^B	87.23%	95.74%	91.48%	91.15%	4.26	
	ASSOCIATED^B	82.97%	87.23%	91.49%	87.20%	4.26	
PRECISION	RADIOMIC^A	70.08%	79.16%	66.66%	72%	6.46	0.016
	SEMANTIC^B	82.75%	96%	88.88%	89.20%	6.63	
	ASSOCIATED^B	84%	85.18%	86.20%	85.10%	1.10	
SENSITIVITY	RADIOMIC^A	68%	76%	64%	69.30%	6.11	0.013
	SEMANTIC^B	96%	96%	96%	96%	0	
	ASSOCIATED^B	84%	92%	100%	92%	8	
SPECIFICITY	RADIOMIC^A	68.18%	77.27%	63.63%	69.70%	6.94	0.045
	SEMANTIC^B	77.27%	95.45%	86.36%	86.40%	9.09	
	ASSOCIATED^{AB}	81.81%	81.81%	81.81%	81.81%	0	

^aDifferent letters indicate statistically significant differences between groups.

^bANOVA test for repeated measures and post Tukey test.

0.013). The semantic features presented the best specificity values, followed by radiomic–semantic associated features and the semantic features attribute extractors differed statistically from radiomic features ($p = 0.045$) (Table 2).

Discussion

IT is a noninvasive and rapid screening method that reveals the dynamics of microcirculation on the skin surface in real-time, detecting the extent of functional, nervous, and vascular changes caused by inflammatory processes, endocrine disorders or oncological conditions.⁷ This diagnostic method has been assessed previously in dentistry in studies on oral candidiasis,³⁶ caries lesions,³⁷ lip herpes³⁸ and endodontic tasks.³⁹

Muscles need oxygen provided by the blood flow to perform their contraction and relaxation movements. The blood flow can also affect the skin temperature which is regulated by the autonomic nervous system.³ In cases of masticatory muscles disorders, the blood flow can be altered and directly affect the muscles thermal patterns⁴⁰; therefore, IT would be indicated to assess muscles thermal alterations. IT diagnostic usefulness in identifying patients with TMD presents limited effectiveness³²; however, the literature is still scarce regarding the reliability of IT in the diagnosis of TMD.⁷

Other image exams have been used to assess the TMJ and aid TMD diagnosis and management. CBCT and CT are useful for evaluating TMJ potentially pathologic changes to the hard tissues such as erosion and osteophyte and morphological changes associated with TMD. CBCT has the advantage of lower doses than CT, and no image distortion when compared to panoramic images. Panoramic images present geometric distortions proper to this image modality technique that may impair the detection of bone alterations⁴¹; therefore, only gross osseous changes can be identified.⁴² Panoramic

images underestimate the radiological findings with higher prevalence, therefore, should not be used as an effectively diagnostic tool for bone components within the TMJ region.⁴³ MRI is useful for evaluating disc morphology, internal disc derangement, joint effusion, abnormalities within the bone marrow of the condyle, and within the muscles and surrounding soft tissues without using ionizing radiation. In many institutions MRI is the preferred examination for TMJ soft tissue pathology; however, its cost is a limitation to MRI indication in many countries. IT is a low-cost noninvasive image exam that can aid TMD diagnosis and management, permitting multiple image acquisition.

Regarding the use of AI in the medical field, studies on the detection of breast cancer,⁴⁴ respiratory frequency,⁴⁵ cellulite stages,^{4,46} differentiation of potentially malignant lesions and lower lip cancer,⁴⁷ facial analysis⁴⁸ and thermographic findings related to orofacial pain have shown promising results.⁴⁹ Recently, at the eighth European Medical and Biological Engineering Conference, a conference paper sought to investigate the ability of infrared thermal image on TMD detection, using AI and found promising AUC values (0.71), thus suggesting that IT can be a relatively inexpensive and simple to use tool for assessing TMD.⁵⁰ Therefore, the use of AI may improve IT diagnostic efficiency increasing productivity and diagnostic accuracy, aiding the subjective interpretation of the clinician.

According to Hung (2020),¹⁹ reports on AI techniques in the dentomaxillofacial radiology field focus on four main applications including automatic localization of cephalometric landmarks, diagnosis of osteoporosis, classification/segmentation of the maxillofacial cysts and/or tumors, and identification of periodontitis/periapical disease. Previous studies have used AI on TMD diagnosis using CBCT scans and have shown satisfactory results.^{20,51–54} A recently published study used AI through the extraction of radiomic features (texture

attributes) to assess thermographic findings related to TMD symptoms of 19 TMD patients and 21 control patients, using the texture attribute homogeneity.⁴⁹ The authors stated that there is a potential for texture attribute analysis assessing IT on TMD diagnosis, as TMD pain affects all masticatory system and can affect thermal changes in muscles unevenly.

Automatic segmentation tools are paramount for the creation of diagnostic task specific software and apps so AI can be used as an auxiliary tool in the dental clinic routine. IT automatic segmentation is difficult because heat distribution is not delimited by anatomical structures, segmentation errors due to similar temperatures between the environment and object of study can occur and different shapes and sizes of the object of interest may difficult the machine learning process.⁵⁴ Therefore, a large dataset must be used with the purpose of obtaining high accuracy automatic segmentation of IT images. Several semi-automatic methods have been proposed for segmentation; however, this process often depends on hand-crafted image features and preprocessing operations.⁵⁵ Studies with large sample sizes are needed to establish the best IT automatic segmentation method.

The total number of published predictive modeling studies using radiomic features has been rapidly rising; however, a consensus on which features are repeatable and reproducible has not yet emerged.⁵⁶ Radiomic features characteristics have values in a continuous scale, which can provide details distinguishing for example odontogenic tumors.^{33,57} Technically, the images are transformed into dimensional data to search for correlations that can define a radiographic phenotype useful before, during or after therapy.^{57,58} Entropy has been consistently reported as one of the most stable radiomic features and contrast appears among the least reproducible radiomic features.⁵⁶ In this study, 10 texture attributes (contrast, correlation, energy, homogeneity, entropy, trimmed means, kurtosis, asymmetry, standard deviation, variance) were assessed, in order to extract as much quantitative texture attributes as possible to improve TMD diagnostic efficiency. A previous study⁴⁹ assessing TMD pain using IT and AI in a smaller study sample used only one radiomic feature (homogeneity) obtaining lower accuracy values than this study.

Additionally, three classification algorithms were used in this study (KNN, SVM and MLP) to assess ML feature extraction accuracy. KNN and SVM classifiers are easy to implement and widespread in the field of ML.⁴⁷ SVM is less computationally demanding than KNN and is easier to interpret but can identify only a limited set of patterns.⁵⁹ SVM presented satisfying accuracy values for all assessed ML feature extraction

methods and performed better than the other studied classifiers for semantic features.

ANN are non-linear models inspired by the brain's neural architecture and were developed aiming to model the learning capacity of biological neural systems. Multi-layer perceptron (MLP) is a typical artificial networks architecture containing a series of layers, composed of artificial neurons and their connections, which can calculate the weighted sum of its inputs and then apply an activation function to obtain a signal that will be transmitted to the next neuron. A MLP must have a good architecture with the necessary connections to solve the problem in question and training with the right amount of data, which can be time-consuming for large datasets.²⁷ Convolutional Neuro Networks (CNN) are a class of artificial neural networks that has increased in the radiology field with promising results; however, this class of ANN needs larger image samples than the one used in this study. Although this study sample was larger than previous studies, the limited sample size is a limitation and studies with larger samples sizes and using CNN are needed.

In this study, the Fonseca questionnaire was used as a first screening tool to assess the volunteers to establish a control group. According to Pires (2018),⁶⁰ the Fonseca questionnaire is a short form that can be used as a screening tool with adequate level of diagnostic accuracy and reliability for myogenous TMD; however, it is important to have a proper diagnosis to guide TMD therapy using the RDC/TMD.⁶¹ Therefore, a second screening was done, assessing all patients with RDC/TMD to correctly categorize the studied groups.

Our study has a certain degree of novelty and a methodology with potential to be used in future studies. Furthermore, more studies assessing the use of AI as an auxiliary tool of image exams (IT, CBCT, MRI) on TMD detection are needed to create a software to be used in dental clinics and radiology centers.

Conclusion

Semantic and radiomic-semantic-associated ML feature extraction methods perform better than radiomic features for TMD detection. The radiomic-semantic attribute extraction method associated with the MLP classifier should be chosen for TMD detection using IT images and pain scale data. IT associated with AI presents promising results for TMD detection.

Acknowledgment

The authors deny any conflicts of interest related to this study.

REFERENCES

- Conville RM, Moriarty F, Atkins S. The management of temporomandibular disorders: a headache in general practice. *Br J Gen Pract* 2019; **69**: 523–4. doi: <https://doi.org/10.3399/bjgp19X705977>
- Gauer RL, Semidey MJ. Diagnosis and treatment of temporomandibular disorders. *Am Fam Physician* 2015; **91**: 378–86.
- Altindış T, Güngörmüş M. Thermographic evaluation of occlusal splint and low level laser therapy in myofascial pain syndrome. *Complement Ther Med* 2019; **44**(January): 277–81 Available from. doi: <https://doi.org/10.1016/j.ctim.2019.05.006>
- Machoy M, Szyszka-Sommerfeld L, Rahnama M, Koprowski R, Wilczyński S, Woźniak K. Diagnosis of temporomandibular disorders using Thermovision imaging. *Pain Res Manag* 2020; **2020**: 5481365. doi: <https://doi.org/10.1155/2020/5481365>
- Brochado FT, Jesus LHde, Carrard VC, Freddo AL, Chaves KD, Martins MD. Comparative effectiveness of photobiomodulation and manual therapy alone or combined in TMD patients: a randomized clinical trial. *Braz Oral Res* 2018; **32**: e50PMID. doi: <https://doi.org/10.1590/1807-3107bor-2018.vol32.0050>
- de Paiva Bertoli FM, Bruzamolín CD, de Almeida Kranz GO, Lasso EM, Brancher JA, de Souza JF. Anxiety and malocclusion are associated with temporomandibular disorders in adolescents diagnosed by RDC/TMD. A cross-sectional study. *J Oral Rehabil* 2018; **45**: 747–55. Epub 2018 Jul 20 PMID. doi: <https://doi.org/10.1111/joor.12684>
- de Melo DP, Bento PM, Peixoto LR, Martins S, Martins CC. Is infrared thermography effective in the diagnosis of temporomandibular disorders? A systematic review. *Oral Surg Oral Med Oral Pathol Oral Radiol [Internet]* 2019; **127**: 185–92 Available from.
- Haddad DS, Brioschi ML, Vardasca R, Weber M, Crosato EM, Arita ES. Thermographic characterization of masticatory muscle regions in volunteers with and without myogenous temporomandibular disorder: preliminary results. *Dentomaxillofac Radiol* 2014; **43**: 20130440. doi: <https://doi.org/10.1259/dmfr.20130440>
- Lahiri BB, Bagavathiappan S, Jayakumar T, Philip J. Infrared Physics & Technology Medical applications of infrared thermography : A review. *Infrared Phys Technol* 2012; **55**: 221–35.
- Biagioni PA, Longmore RB, McGimpsey JG, Lamey PJ. Infrared thermography. its role in dental research with particular reference to craniomandibular disorders. *Dentomaxillofac Radiol* 1996; **25**: 119–24 PMID. doi: <https://doi.org/10.1259/dmfr.25.3.9084259>
- Haddad DS, Brioschi ML, Arita ES. Thermographic and clinical correlation of myofascial trigger points in the masticatory muscles. *Dentomaxillofac Radiol* 2012; **41**: 621–9. doi: <https://doi.org/10.1259/dmfr/98504520>
- Dibai-Filho AV, Costa ACdeS, Packer AC, Rodrigues-Bigaton D. Correlation between skin surface temperature over masticatory muscles and pain intensity in women with myogenous temporomandibular disorder. *J Back Musculoskelet Rehabil* 2013; **26**: 323–8. doi: <https://doi.org/10.3233/BMR-130387>
- Rodrigues-Bigaton D, Dibai-Filho AV, Packer AC, Costa ACdeS, de Castro EM. Accuracy of two forms of infrared image analysis of the masticatory muscles in the diagnosis of myogenous temporomandibular disorder. *J Bodyw Mov Ther* 2014; **18**: 49–55. doi: <https://doi.org/10.1016/j.jbmt.2013.05.005>
- Dibai-Filho AV, Packer AC, Costa ACdeS, Rodrigues-Bigaton D. The chronicity of myogenous temporomandibular disorder changes the skin temperature over the anterior temporalis muscle. *J Bodyw Mov Ther* 2014; **18**: 430–4. doi: <https://doi.org/10.1016/j.jbmt.2013.11.001>
- Dibai-Filho AV, Costa AC de S, packer AC, de Castro em, Rodrigues-Bigaton D. women with more severe degrees of temporomandibular disorder exhibit an increase in temperature over the temporomandibular joint. *Saudi Dent J* 2015; **27**: 44–9.
- Barbosa JS, Amorim A, Arruda M, Medeiros G, Freitas A, Vieira L, et al. Infrared thermography assessment of patients with temporomandibular disorders. *Dentomaxillofac Radiol* 2020; **49**: 20190392. doi: <https://doi.org/10.1259/dmfr.20190392>
- Kawano W, Kawazoe T, Tanaka M, Hikida Y. Deep thermometry of temporomandibular joint and masticatory muscle regions. *J Prosthet Dent* 1993; **69**: 216–21. doi: [https://doi.org/10.1016/0022-3913\(93\)90143-c](https://doi.org/10.1016/0022-3913(93)90143-c)
- Rahmayani L, Yahya M, Soraya C, Syahreza S. Thermal condition of muscle area around the temporomandibular joint in patient with systemic lupus erythematosus using infrared thermography application: a case report. *J Int Soc Prev Community Dent* 2020; **10**: 674–9 PMID PMIDCID. doi: https://doi.org/10.4103/jisped.JISPCD_126_20
- Hung K, Montalvao C, Tanaka R, Kawai T, Bornstein MM. The use and performance of artificial intelligence applications in dental and maxillofacial radiology: a systematic review. *Dentomaxillofac Radiol* 2020; **49**: 20190107. doi: <https://doi.org/10.1259/dmfr.20190107>
- de Dumast P, Mirabel C, Cevidanes L, Ruellas A, Yatabe M, Ioshida M, et al. A web-based system for neural network based classification in temporomandibular joint osteoarthritis. *Comput Med Imaging Graph* 2018; **67**: 45–54. Epub 2018 May 1 PMID PMIDCID. doi: <https://doi.org/10.1016/j.compmedimag.2018.04.009>
- Faust O, Rajendra Acharya U, Ng EYK, Hong TJ, Yu W. Application of infrared thermography in computer aided diagnosis. *Infrared Phys Technol* 2014; **66**: 160–75. doi: <https://doi.org/10.1016/j.infrared.2014.06.001>
- Bindushree V, Sameen RJ, Vasudevan V, Shrihari TG, Devaraju D, Mathew NS. Artificial intelligence: in modern dentistry. *J Dent Res Rev* 2020; **7**: 27–31.
- Gonzalez RC, Woods RE, Prentice Hall P. *Digital Image Processing*. Third Edition. **954**: Pearson International Edition prepared by Pearson Education; 2008.
- Salvador-Meneses J, Ruiz-Chavez Z, Garcia-Rodriguez J. Compressed kNN: K-Nearest Neighbors with Data Compression. *Entropy* 2019; **21**: 1–20. doi: <https://doi.org/10.3390/e21030234>
- Noble WS. What is a support vector machine? *Nat Biotechnol* 2006; **24**: 1565–7. doi: <https://doi.org/10.1038/nbt1206-1565>
- Nalepa J, Kawulok M. Selecting training sets for support vector machines: a review. *Artif Intell Rev [Internet]* 2019; **52**: 857–900 Available from.
- Castro W, Oblitas J, Santa-Cruz R, Avila-George H. Multilayer perceptron architecture optimization using parallel computing techniques. *PLoS One* 2017; **12**: 1–17. doi: <https://doi.org/10.1371/journal.pone.0189369>
- Dworkin SF, LeResche L. Research diagnostic criteria for temporomandibular disorders: review, criteria, examinations and specifications, critique. *J Craniomandib Disord* 1992; **6**: 301–55 PMID-Fall.
- Schwartz RG. Temporomandibular disorders. *N Eng J Med* 2008; **359**: 447–66.
- Brioschi ML, Macedo JF, Macedo RAC. Skin thermometry: new concepts. *J Vasc Bras* 2003; **2**.
- Brioschi ML. Principios E indicações da termografia médica first edition prepared by Andreoli. 2010;.
- Woźniak K, Szyszka-Sommerfeld L, Trybek G, Piątkowska D. Assessment of the sensitivity, specificity, and accuracy of thermography in identifying patients with TMD. *Med Sci Monit* 2015; **21**: 1485–93. doi: <https://doi.org/10.12659/MSM.893863>
- Yip SSF, Liu Y, Parmar C, Li Q, Liu S, Qu F, et al. Associations between radiologist-defined semantic and automatically computed radiomic features in non-small cell lung cancer. *Sci Rep* 2017; **7**: 3519 PMID PMIDCID. doi: <https://doi.org/10.1038/s41598-017-02425-5>
- Coroller TP, Bi WL, Huynh E, Abedalthagafi M, Aizer AA, Greenwald NF, et al. Radiographic prediction of meningioma grade by semantic and radiomic features. *PLoS One* 2017; **12**: e0187908. doi: <https://doi.org/10.1371/journal.pone.0187908>
- Pan S, Ding Z, Zhang L, Ruan M, Shan Y, Deng M, et al. A nomogram combined radiomic and semantic features as imaging biomarker for classification of ovarian cystadenomas. *Front Oncol* 2020; **10**: 895 PMID PMIDCID. doi: <https://doi.org/10.3389/fonc.2020.00895>

36. Iosif L, Preoteasa CT, Murariu-Măgureanu C, Preoteasa E. Clinical study on thermography, as modern investigation method for Candida-associated denture stomatitis. *Rom J Morphol Embryol* 2016; **57**: 191–5PMID.
37. Zakian CM, Taylor AM, Ellwood RP, Pretty IA. Occlusal caries detection by using thermal imaging. *J Dent* 2010; **38**: 788–95. Epub 2010 Jul 3PMID. doi: <https://doi.org/10.1016/j.jdent.2010.06.010>
38. Biagioni PA, Lamey PJ. Electronic infrared thermography as a method of assessing herpes labialis infection. *Acta Derm Venereol* 1995; **75**: 264–8PMID. doi: <https://doi.org/10.2340/000155575264268>
39. Kilic K, Er O, Kilinc HI, Aslan T, Bendes E, Sekerci AE, et al. Infrared thermographic comparison of temperature increases on the root surface during dowel space preparations using circular versus oval fiber dowel systems. *J Prosthodont* 2013; **22**: 203–7. Epub 2012 Sep 17PMID. doi: <https://doi.org/10.1111/j.1532-849X.2012.00919.x>
40. Costa ACS, Dibai Filho AV, Packer AC, Rodrigues-Bigaton D. Intra and inter-rater reliability of infrared image analysis of masticatory and upper trapezius muscles in women with and without temporomandibular disorder. *Braz J Phys Ther* 2013; **17**: 24–31. doi: <https://doi.org/10.1590/s1413-35552012005000058>
41. Honey OB, Scarfe WC, Hilgers MJ, Klueber K, Silveira AM, Haskell BS, et al. Accuracy of cone-beam computed tomography imaging of the temporomandibular joint: comparisons with panoramic radiology and linear tomography. *Am J Orthod Dentofacial Orthop* 2007; **132**: 429–38. doi: <https://doi.org/10.1016/j.ajodo.2005.10.032>
42. Brooks SL, Brand JW, Gibbs SJ, Hollender L, Lurie AG, Omnell KA, et al. Imaging of the temporomandibular joint: a position paper of the American Academy of oral and maxillofacial radiology. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 1997; **83**: 609–18. doi: [https://doi.org/10.1016/s1079-2104\(97\)90128-1](https://doi.org/10.1016/s1079-2104(97)90128-1)
43. Ladeira DBS, da Cruz AD, de Almeida SM. Digital panoramic radiography for diagnosis of the temporomandibular joint: CBCT as the gold standard. *Braz Oral Res* 2015; **29**: S1806–303. doi: <https://doi.org/10.1590/1807-3107BOR-2015.vol29.0120>
44. Mambou SJ, Maresova P, Krejcar O, Selamat A, Kuca K. Breast cancer detection using infrared thermal imaging and a deep learning model. *Sensors* 2018; **18**: 2799.
45. Jagadev P, Giri LI. Human respiration monitoring using infrared thermography and artificial intelligence. *Biomed Phys Eng Express* 2020; **6**: 035007. doi: <https://doi.org/10.1088/2057-1976/ab7a54>
46. Bauer J, Hoq MN, Mulcahy J, Tofail SAM, Gulshan F, Silien C, et al. Implementation of artificial intelligence and non-contact infrared thermography for prediction and personalized automatic identification of different stages of cellulite. *Epma J* 2020; **11**: 17–29. doi: <https://doi.org/10.1007/s13167-020-00199-x>
47. Paulino JAS, Da Silva Alves AV, De Sousa RP, Galdino KE, Alves PM, Oliveira EJV. The use of hybrid classifier to assist in the diagnosis of lip lesions in thermal images: a preliminary study. *Proc - 2019 IEEE Int Conf Bioinform Biomed BIBM 2019* 2019; 1585–8.
48. Müller D, Ehlen A, Valeske B. Convolutional neural networks for semantic segmentation as a tool for multiclass face analysis in thermal infrared. *J Nondestr Eval* 2021; **40**: 9. doi: <https://doi.org/10.1007/s10921-020-00740-y>
49. Rytivaara R, Nöpänkangas R, Kainulainen T, Sipola A, Kallio-Pulkkinen S, Raustia A. Thermographic findings related to facial pain—a survey of 40 subjects. *Cranio - J Craniomandib Pract [Internet]* 2021; **00**: 1–8Available from.
50. Perpetuini D, Trippetti N, Cardone D, Breda L, D’Attilio M, Merla A. Detection of Temporomandibular Joint Dysfunction in Juvenile Idiopathic Arthritis Through Infrared Thermal Imaging and a Machine Learning Procedure. In: Jarm T, Cvetkoska A, Mahnič-Kalamiza S, Miklavcic D, eds. *8th European Medical and Biological Engineering Conference: EMBEC 2020*. IFMBE Proceedings, vol 80. Springer, Cham; 2021.
51. Yi PH, Lin A, Wei J, Yu AC, Sair HI, Hui FK, et al. Deep-Learning-Based semantic labeling for 2D mammography and comparison of complexity for machine learning tasks. *J Digit Imaging* 2019; **32**: 565–70. doi: <https://doi.org/10.1007/s10278-019-00244-w>
52. Bianchi J, de Oliveira Ruellas AC, Gonçalves JR, Paniagua B, Prieto JC, Styner M, et al. Osteoarthritis of the temporomandibular joint can be diagnosed earlier using biomarkers and machine learning. *Sci Rep* 2020; **10**: 8012PMIDPMCID. doi: <https://doi.org/10.1038/s41598-020-64942-0>
53. Lee KS, Kwak HJ, Oh JM, Jha N, Kim YJ, Kim W, et al. Automated detection of TMJ osteoarthritis based on artificial intelligence. *J Dent Res* 2020; **99**: 1363–7. Epub 2020 Jul 1PMID. doi: <https://doi.org/10.1177/0022034520936950>
54. Liu Z, Wang S, Dong D, Wei J, Fang C, Zhou X, et al. The applications of radiomics in precision diagnosis and treatment of oncology: opportunities and challenges. *Theranostics* 2019; **9**: 1303–22. doi: <https://doi.org/10.7150/thno.30309>
55. Tayel MB, Elbagoury AM. Breast infrared thermography segmentation based on adaptive tuning of a fully Convolutional network. *Curr Med Imaging* 2020; **16**: 611–21Available from. doi: <https://doi.org/10.2174/1573405615666190503142031>
56. Traverso A, Wee L, Dekker A, Gillies G. Repeatability and reproducibility of radiomic features: a systematic review. *Int J Radiat Oncol Biol Phys* 2018; **102**: 1143–58Available from. doi: <https://doi.org/10.1016/j.ijrobp.2018.05.053>
57. Wu W, Pierce LA, Zhang Y, Pipavath SNJ, Randolph TW, Lastwika KJ, et al. Comparison of prediction models with radiological semantic features and radiomics in lung cancer diagnosis of the pulmonary nodules: a case-control study. *Eur Radiol* 2019; **29**: 6100–8. Epub 2019 May 21PMIDPMCID. doi: <https://doi.org/10.1007/s00330-019-06213-9>
58. Hawkins S, Wang H, Liu Y, Garcia A, Stringfield O, Krewer H, et al. Predicting malignant nodules from screening CT scans. *J Thorac Oncol* 2016; **11**: 280–1; . . . Epub 2016 Jul 13PMIDPMCIDErratum in: *J Thorac Oncol*. 2018 Feb11(12):2120-2128.. doi: <https://doi.org/10.1016/j.jtho.2016.07.002>
59. Bzdok D, Krzywinski M, Altman N. Machine learning: supervised methods, SVM and kNN. *Nature Methods, Nature Publishing Group* 2018; pp: 1–6ffhal-01657491f.
60. Pires PF, de Castro EM, Pelai EB, de Arruda ABC, Rodrigues-Bigaton D. Analysis of the accuracy and reliability of the short-form Fonseca anamnestic index in the diagnosis of myogenous temporomandibular disorder in women. *Brazilian J Phys Ther [Internet]* 2018; **22**: 276–82.
61. Campos J, Gonçalves DAG, Camparis CM, Speciali JG. Reliability of a questionnaire for diagnosing the severity of temporomandibular disorder. *Rev Bras Fisioter* 2009; **13**: 38–43.