IOTeeth: Intra-Oral Teeth Sensing System for Dental Occlusal Diseases Recognition

ZHIZHANG HU*, University of California, Merced, USA

AMIRMOHAMMAD RADMEHR*, University of Massachusetts Amherst, USA

YUE ZHANG, University of California, Merced, USA

SHIJIA PAN†, University of California, Merced, USA

PHUC NGUYEN†, University of Massachusetts Amherst, USA

While occlusal diseases – the main cause of tooth loss – significantly impact patients' teeth and well-being, they are the most underdiagnosed dental diseases nowadays. Experiencing occlusal diseases could result in difficulties in eating, speaking, and chronicle headaches, ultimately impacting patients' quality of life. Although attempts have been made to develop sensing systems for teeth activity monitoring, solutions that support sufficient sensing resolution for occlusal monitoring are missing. To fill that gap, this paper presents *IOTeeth*, a cost-effective and automated intra-oral sensing system for continuous and fine-grained monitoring of occlusal diseases. The *IOTeeth* system includes an intra-oral piezoelectric-based sensing array integrated into a dental retainer platform to support reliable occlusal disease recognition. *IOTeeth* focuses on biting and grinding activities from the canines and front teeth, which contain essential information of occlusion. *IOTeeth*'s intra-oral wearable collects signals from the sensors and fetches them into a lightweight and robust deep learning model called Physio-aware Attention Network (PAN Net) for occlusal disease recognition. We evaluate *IOTeeth* with 12 articulator teeth models from dental clinic patients. Evaluation results show an F1 score of 0.97 for activity recognition with leave-one-out validation and an average F1 score of 0.92 for dental disease recognition for different activities with leave-one-out validation.

CCS Concepts: • **Human-centered computing** → **Ubiquitous and mobile computing systems and tools**; • **Applied computing** → *Consumer health*.

Additional Key Words and Phrases: Intra-oral Sensing, Occlusal Disease Recognition

1 INTRODUCTION

Dental diseases, encompassing conditions such as bruxism [70], temporomandibular joint disorders (TMD) [84], and periodontal diseases [7], detrimentally affect oral health and the overall quality of life of millions of people

---

*Equal contribution.

†Co-corresponding authors.

Authors' addresses: Zhizhang Hu, zhu42@ucmerced.edu, University of California, Merced, USA; Amirmohammad Radmehr, aradmehr@ umass.edu, University of Massachusetts Amherst, USA; Yue Zhang, yzhang58@ucmerced.edu, University of California, Merced, USA; Shijia Pan, span24@ucmerced.edu, University of California, Merced, USA; Phuc Nguyen, phuc@umass.edu, University of Massachusetts Amherst, USA.

Fig. 1. *IOTeeth*'s concept: embedding customized small-sized piezoelectric sensors into dental retainers for recognizing and monitoring user's occlusal disease.

worldwide [38, 52, 64]. In addition to physical discomfort, dental diseases can lead to psychological distress, lowered self-esteem, and reduced social interactions because of aesthetic concerns and functional limitations [47]. Among dental diseases, occlusal disease is often overlooked and is one of the most underdiagnosed dental diseases [7, 10, 25, 66]. However, it can significantly impact patients' teeth and well-being, given it is the main cause of tooth loss [75] and can result in high-impact chronic pain [35]. The term "occlusion" refers to the way the upper and lower teeth fit together (bite) [6], and occlusal disease refers to pathological effects when upper and lower jaws, facial muscles and joints work in disorder and teeth do not fit [43]. This musculoskeletal disorder results in imbalanced force distribution and often leads to worn/damaged teeth, sensitive teeth, and tooth loss [24, 25]. It can also cause excessive muscle activities in the jaw and neck to compensate for the imbalance, resulting in muscle tension and a chronicle headache [77, 79].

The current dental practice relies on the biannual preventive examination for occlusal disease screening, which does not capture the fine-grained changes of teeth activities (e.g., stress-induced excessive grinding) and their impact on the occlusion. These teeth activities and early-stage symptoms (e.g., worn teeth, slight crack) may be unnoticed by the person or misunderstood as aging issues, and patients' self-reported outcomes during dentist visits are not reliable [25]. This kind of neglect could lead to delaying early treatment opportunities and causing unnecessary aggravation of the occlusal disease condition [76]. Additionally, alterations in people's occlusion may occur after dental surgeries, like teeth replacement [4]. Profiling the pre- and post-surgery occlusal condition is crucial for tracking and rectifying any unwanted alterations in occlusion, ensuring the success of the surgical intervention and preventing potential occlusal diseases. Therefore, it is essential to conduct continuous monitoring of the oral occlusal conditions for in-time disease recognition [14, 25], occlusion profiling [8], and proactive prevention and treatment [9, 25].

Although attempts have been made to achieve intra-oral sensing, existing systems mainly focus on activity (e.g., clenching) monitoring [26, 30, 34, 56] rather than occlusal disease inference. These systems leverage sensors (e.g., piezoresistive) that focus on sensing static or quasi-static mechanical stress to measure bite force and detect activity events. However, they cannot be used to infer the occlusion conditions due to the limited resolution (slow response to dynamic stress changes) [42], which can not capture the transient information of teeth contact during activities for deriving occlusion conditions. TeethVib integrates piezoelectric sensors, which focus on capturing high-dynamic mechanical stress changes (e.g. vibrations) [42], into layers of dental retainers for functional occlusion monitoring [58]. Nonetheless, it only demonstrates the feasibility of using piezoelectric sensors over different teeth to acquire signals of different patterns, without inferring any dental conditions or activities.

We propose *IOTeeth*, a dental retainer-form intra-oral sensing system for continuously recognizing occlusal diseases to remind the user for an in-time dental diagnose, which is illustrated in Figure 1. The *IOTeeth* system monitors and analyzes teeth contact-induced vibration signals to recognize the existence of occlusal disease, and it can report an aggregated result to the user with the user's preferred frequency (e.g., daily, weekly, etc.) *IOTeeth* comprises an intra-oral sensing platform that captures the vibration caused by teeth contact during oral

activities via customized piezoelectric sensors embedded between dental retainers. Note that in this work, the oral activities-induced vibration signals are all collected from articulator-mounted dental models replicated from dental clinic customers, instead of real humans. The acquired vibration signals are then processed and analyzed to recognize the presence of occlusal diseases.

Realizing such a system to sense and recognize occlusal diseases is not trivial and faces the following challenges. (C1) The human mouth has limited capacity and oral occlusion is sensitive to alignment and position, thus it is challenging to design and manufacture the sensing wearable with minimum interference to the occlusion. (C2) People's facial musculoskeletal profile varies significantly, therefore, the human variance (both physical and behavioral) would make it difficult to generalize the learning model with limited data. To minimize the interference to the occlusion (C1), we design the intra-oral sensing wearable in the form of thin dental retainer, and fabricate small-sized sensors with a radius of four millimeters to minimize teeth gaps after the placement. To ensure the quality of sensors, we develop an automated benchmarking platform for the sensor characterization. We also optimize the placement of small-sized sensors on the internal side of canines and front teeth, to minimize the introduced interference. To tackle the human variance-induced significant data distribution shift while achieving accurate disease recognition (C2), we propose a Physio-aware Network (PAN Net) to process and extract the most informative patterns from the signal for generalizable modeling. The reason that PAN Net is able to recognize occlusal disease from teeth contact-induced vibration is that such vibration encodes the relationship of the dental arches, the teeth morphology, and the musculoskeletal condition implicitly, which are the indicators that dentists use to determine the occlusal diseases [22, 43]. By extracting and decoding these features, the PAN Net is able to conduct accurate occlusal disease recognition. The contributions of this work are as follows.
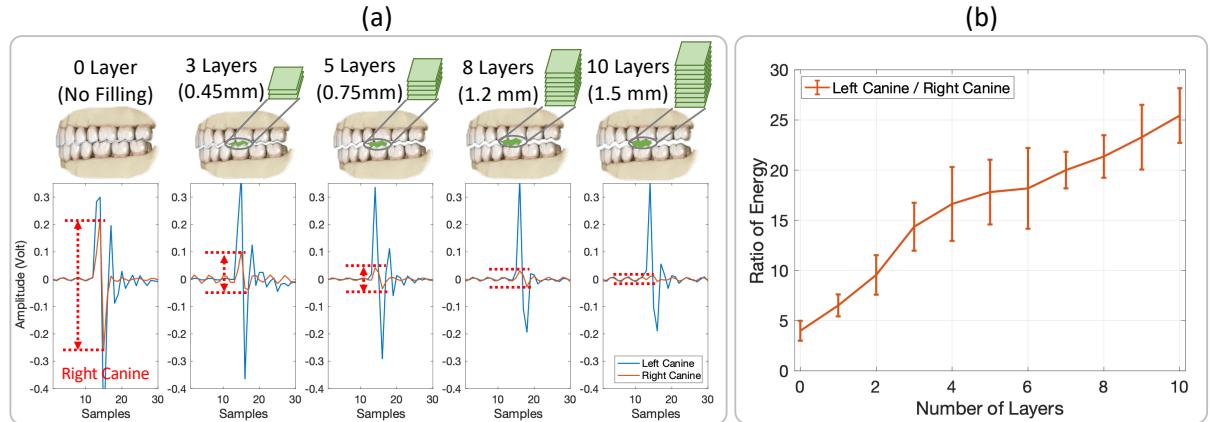
- We present *IOTeeth*, an intra-oral teeth sensing system that consists of a dental retainer-form sensing platform and data-driven algorithms to capture teeth contact-induced vibration signals during oral activities and recognizes occlusal diseases from them.
- We design and custom small-sized piezoelectric-based sensors and strategically integrate them into the dental retainers to form the sensing platform's wearable of the *IOTeeth*, with a minimized interference to the occlusion.
- We propose a Physio-aware Network as the *IOTeeth*'s occlusal disease recognition algorithm that allows generalized modeling while focusing on the teeth that are representative of the individual user.
- We evaluate *IOTeeth* with 12 articulator-based teeth models built with patients' bite registrations by the dentist and achieve over 0.92 average micro-F1-score for leave-one-out occlusal disease recognition.

The rest of the paper is organized as follows. Section 2 introduces the background for occlusion-related dental diseases and intra-oral vibration sensing, and showcase a feasibility study of measuring fine-grained occlusion changing with piezoelectric sensing. Section 3 presents the *IOTeeth* overview, followed by Section 4 and 5 for *IOTeeth*'s sensing platform and algorithm details. Section 6 demonstrates the evaluation with real-world data collection using models with patients' bite registrations. Section 7 discusses the future directions that can be further explored and applications beyond occlusal disease monitoring, followed by the related work discussion in Section 8. Finally, we conclude in Section 9.

2 FUNDAMENTAL OF INTRA-ORAL OCCLUSAL SENSING

2.1 Primer on Occlusion, Malocclusion, and Occlusal Diseases

Dental occlusion refers to the way the teeth meet. When people grow, their occlusion changes due to various reasons (e.g., bad habit, dental filling, bruxism, etc.) [12]. These changes may develop into adverse oral conditions, such as malocclusion and occlusal diseases. Malocclusion refers to misalignment or incorrect relation between the teeth of the two dental arches as the jaws close (e.g., crossbite, open bite, overbite, underbite), and is typically identifiable through visual examination [51]. Occlusal diseases (such as temporomandibular joint disorders, also

Fig. 2. Feasibility experiment of detecting fine-grained occlusion change. Figure (a) shows examples when there are different levels of thin film fillings that change the occlusion. The highlighted right canine's signal energy gradually decreases with the number of filling increases. Figure (b) plots the ratio between two signals when there are different numbers of layers placed.
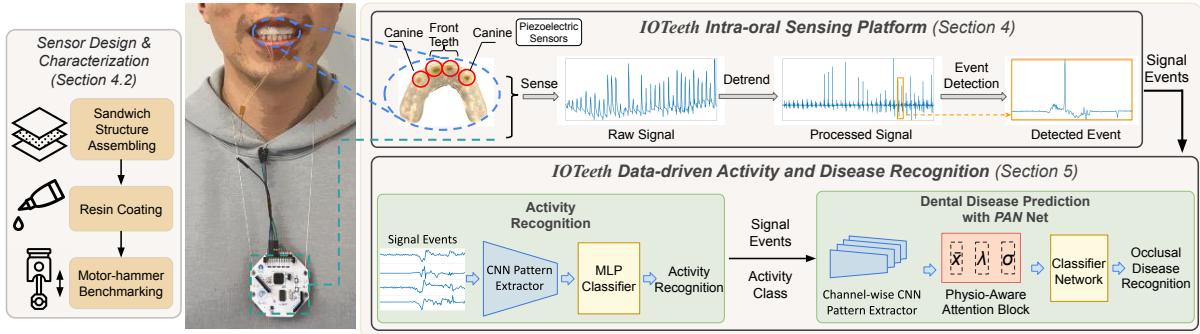
known as TMD) are pathological effects resulting from facial musculoskeletal disorders, leading to excessive tooth wear, tooth fracture, and complications such as chronic headaches and jaw pain [25]. Among these occlusion-related problems, occlusal disease is the most destructive process [17] and the most overlooked one due to misleading symptoms. Patients may consider occlusal disease symptoms to be caused by reasons like lack of rest, aging, or sports injuries [25]. Unlike malocclusion can be visually recognized, dentists often overlook the musculoskeletal system's functionality in their checking routine and miss the signs of occlusal disease [76]. Therefore, we are targeting to recognize it in this work.

2.2 Intra-oral Vibration Sensing

Our upper and lower teeth often contact each other when we close our mouths during daily activities such as eating, grinding, biting, and others. Characterizing such contact is important to understand how well our teeth are aligned. Multiple sensors can be used to monitor those vibrations, including IMU, microphone, or piezoelectric sensor. Yet, the constrained space of the oral cavity has restricted the options, ultimately making piezoelectric sensing the most practical solution. The piezoelectric sensor is commonly used to measure vibrations and mechanical stress with high sensitivity. When the piezoelectric material experiences mechanical stress, its atomic bonding structure deforms and results in a change in the electric dipole distribution. This change generates an electric charge across the material, resulting in measurable voltage. Guided by this physic principle, there is a linear relationship between the mechanical stress applied to the piezoelectric material and the amplitude of output voltage [1, 5]:

$$V = \frac{dLF}{\epsilon_0 \epsilon_r A},$$ (1)

where V is the output voltage, d is the piezoelectric coefficient, L is the material length, F is the applied mechanical stress, ϵ_0 and ϵ_r are permittivity coefficients of the material, and A is the area that the stress is applied on. Hence, by placing a tiny piezo sensor at the contact area, we might be able to accurately and reliably measure and characterize the biting behavior reliably.

Fig. 3. *IOTeeth* system overview.
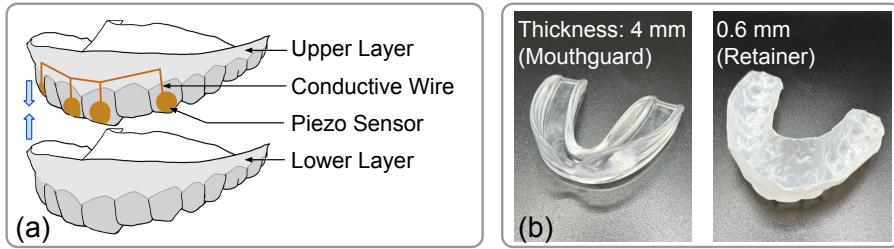
2.3 Feasibility Study

We conduct a preliminary study to demonstrate the feasibility of using the intra-oral piezoelectric sensor to capture changes in occlusion. A dentist records a person's dental anatomical landmarks via bite registration, which describes the relationship between upper and lower dental arches and demonstrates nearly balanced horizontal teeth contact. A dental plaster model is then cast from the bite registration. We place two piezoelectric sensors on the internal surface of canines to capture the vibration induced by teeth contact while preserving the original occlusion. To simulate the changes in occlusion, we use thin films (one-layer thickness $\approx 0.15mm$) to fill in the groove of the left first molar. This offset usually happens after a tooth-filling surgery, when the dimension of the filling material is different from the original tooth structure. With an increasing number of films, the changes in occlusion increase. For each setting, we simulate the biting motion with the articulator and collect the vibration signals captured by two sensors. Figure 2 (a) depicts bite-induced vibration signals, which show observable signal energy differences between the left and right sensors when the occlusion condition changes gradually. Figure 2 (b) quantifies the ratio of signal energy with different numbers of filling layers. This result depicts the feasibility of using intra-oral vibration sensors to capture changes in occlusion, which is the foundation for more complicated occlusal disease recognition.

3 *IOTEETH* SYSTEM OVERVIEW

We design the *IOTeeth* system as an intra-oral sensing system for accurately recognizing the existence of occlusal diseases from teeth contact-induced vibration during oral activities. Figure 3 illustrates the overview of the *IOTeeth* system, comprising two modules, namely intra-oral sensing platform and data-driven activity and disease recognition algorithms.

The proposed intra-oral sensing platform collects and pre-processes the data, with three parts: a retainer-form wearable, a data acquisition device, and a signal pre-processing unit. The retainer-form wearable device is designed for comfortable long-term use. And as demonstrated in the feasibility study, small changes in tooth morphology would result in significant changes in occlusion. Therefore, in order to **minimize the interference** to natural occlusion, we design, manufacture, and benchmark piezoelectric sensors that are small enough ($2mm$-radius) and place them on the internal surface of front teeth and canines. When the user wears *IOTeeth*, their oral activities such as bite and grind will induce vibration that can be captured via the data acquisition device. These signals are then sent to signal pre-processing unit, which can be deployed on the cloud or the local server, to remove the noise and extract segments of interest. These segments are the input to *IOTeeth*'s data-driven activity and disease recognition algorithms. We introduce details of this module in Section 4.

Fig. 4. (a): Retainer-form design of the *IOTeeth*'s intra-oral sensing platform. (b): Comparison of a mouthguard and two-layer retainers. The annotation indicates the thickness of one layer.

The data-driven activity and disease recognition algorithms need to efficiently extract anatomical characteristics of the masticatory system. It is challenging because 1) what the sensor directly measures is the skeletal characteristics and the algorithm needs to infer the musculoskeletal condition indirectly, and 2) the labeled data is costly and limited. Therefore, in order to extract effective features that are **robust across people with different dental anatomic landmarks**, we propose a Physio-aware Attention Network (PAN Net) which is able to automatically select the most informative sensor for each person for occlusal disease recognition. Signal segments of different activities represent different musculoskeletal motions and show distinctive signal characteristics, thus it requires activity-specific occlusal disease recognition models. *IOTeeth* predicts occlusal disease for each segment of interest, and these predictions can be used to further generate aggregated results. The detail of the module is elaborated in Section 5.
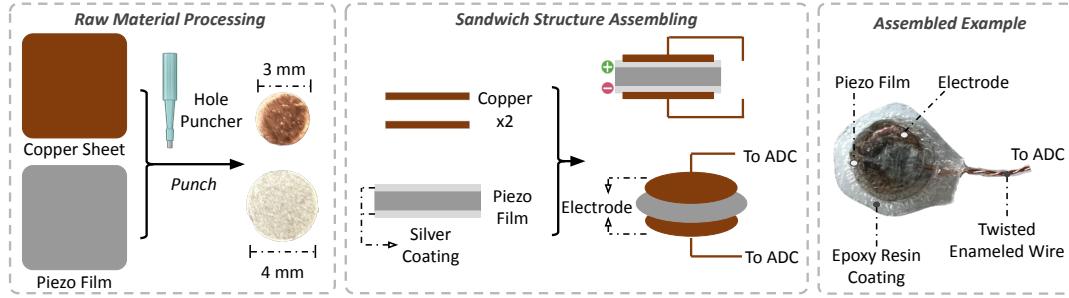
4 INTRA-ORAL SENSING PLATFORM

In this section, we first introduce the intra-oral wearable device with the retainer-form design in Section 4.1. Then, we present the design, fabrication, and characterization process of the customized piezoelectric teeth sensor in Section 4.2. In Section 4.3, we illustrate how different components are integrated as a sensing platform. Lastly, in Section 4.4 we present how the vibration signals captured by the intra-oral wearable are processed for the next module.

4.1 Retainer-Form Intra-oral Wearable Device

We design *IOTeeth*'s intra-oral wearable in the form of a dental retainer, which is shown in Figure 4 (a). A dental retainer is made from a thin (0.5-1.2 mm[1]) thermoplastic sheet customized to fit the people's teeth shapes. Compared to other dental appliances, building the intra-oral wearable upon dental retainer helps to reduce the overall thickness of the platform, hence minimizing the impact on the user's natural oral activities [78]. Figure 4 (b) shows the comparison of one single-layer mouthguard and two-layer retainers and the two-layer retainers' thickness (1.2 mm) is less than half of the mouthguard's thickness (4 mm). This less intrusive factor enables the system to minimize the interference to the natural occlusal condition. Also, it enables us to design a vertical placement of piezoelectric sensors. As shown in Figure 3 and Figure 4 (a), we mount piezoelectric sensors onto the inner surface of the anterior teeth and canines, placing them securely between two layers of retainers. The retainer's snug, customized fit ensures the piezoelectric sensors remain securely in place and facilitates close attachment to the teeth for optimal vibration sensing. Additionally, the vertical placement minimizes the added discrepancy between the lower and upper jaw. Therefore, it avoids the discomfort from the altered teeth fitting from the amplified upper and lower teeth gaps caused by the hinge joint structure of the human mouth.

---

[1]This is the common thickness range in North America.

Fig. 5. Overview of the sensor fabrication process, including individual components, assembling procedure and an exemplar fabricated sensor.
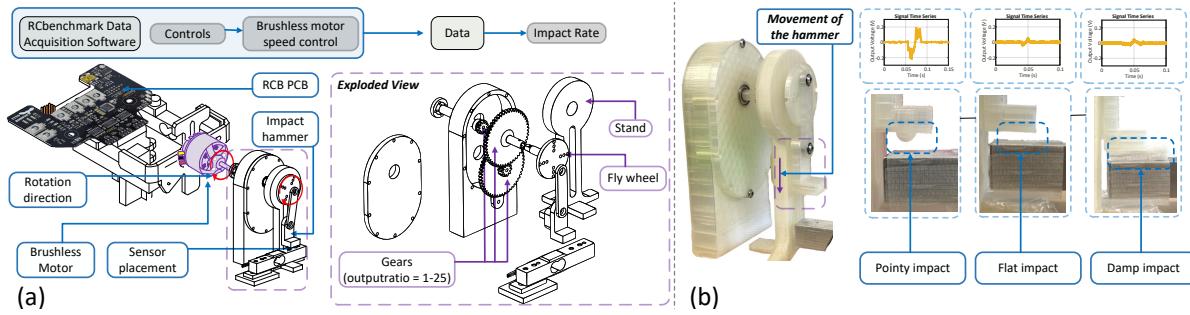
The placement location also coheres with the function of front teeth and canines during oral activities (e.g., canine-guided grinding) enabling the system to effectively capture comprehensive information on the user's occlusal condition.

4.2 Sensor Design and Characterization
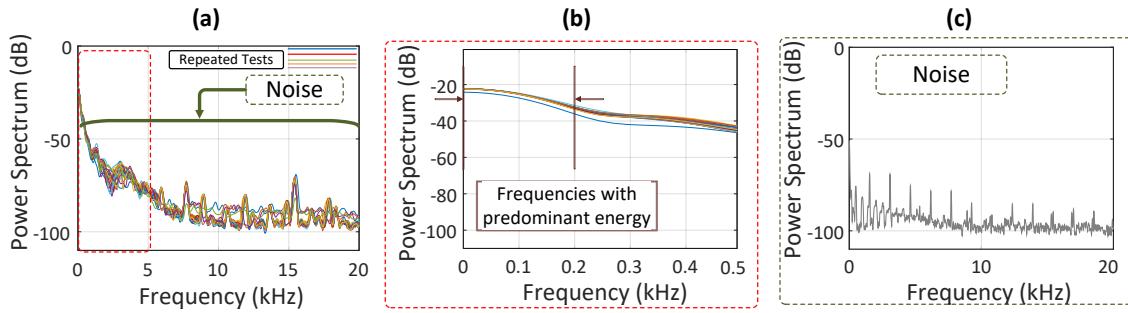
Sensor Design. We design the sensor for the wearable to be sensitive for capturing vibration signals while being small and thin enough for the designed canines and front teeth placement. To the best of our knowledge, there are no commodity sensors that suit our requirements. Therefore, we design a customized piezoelectric sensor. Figure 5 illustrates the design and fabrication process of the sensor. We customize the sensor upon PVDF piezoelectric films for its high sensitivity of vibration sensing and flexibility in fitting into curved surfaces [42] as can be seen in Figure 4 (b). The sensor is designed in a circular shape with a four-millimeter diameter. Compared with other shapes, circular piezoelectric sensors are known to have lower noise and higher sensitivity [2, 50]. The 4mm diameter dimension also enables the sensor to fit the canine and front tooth's size (around seven millimeters with a variance of three millimeters in width [93]). Given such small size, a silver coating is applied to the sensor to improve its sensitivity. This silver coating provides a conductive layer that allows the sensor to capture fine-grained variations in the generated voltage. To read the output voltage, we place three two circular copper films (3mm-diameter) on each side of the coated piezo film as electrodes. Under the in-mouth working conditions, the sensor will be exposed to saliva, which may cause short circuits. Therefore, another epoxy resin coating is designed to isolate the sensor from the humid working environment and protect the sensor. With the proposed multi-layer design, the sensor meets our requirements for sensitivity, comfort, and durability.

Sensor Characterization. To ensure optimal performance of the fabricated sensor, a dedicated sensor characterization is required to benchmark sensors' sensitivity and consistency. Also, the characterization provides guidance for selecting a data acquisition device compatible with the specific output of the customized sensor.

We first design a procedure to characterize the sensor with conditions that simulate real-world scenarios: 1) different rates of impact, 2) impacts of different types of impact heads, and 3) consistent long-term impacts. The first test mimics different frequencies and extent of intra-oral activities. The second test helps us understand the sensor's behavior when embedded in the dental retainer. A mild impact head can function as a damper to mimic the mechanical stress from vibration when the sensor is embedded in the retainer, while the hard surface impact heads can be used as a reference. The third test mimics long-term usage, which can help us understand how durable our customized sensor is. These different characterizing tests can provide insights into the sensor's behavior and enable us to optimize the sensor's design and performance for specific applications. To ensure sensor quality, we examine the sensor in three aspects: The first aspect is examining if there is a broken sensor.

Fig. 6. a) This shows an overview of the developed test bench and the principle of impact testing, which is used to evaluate the performance of materials under mechanical stress. b) This shows the multiple hammer endpoint used to test the sensors, which are designed to measure the impact frequency for characterization and testing.
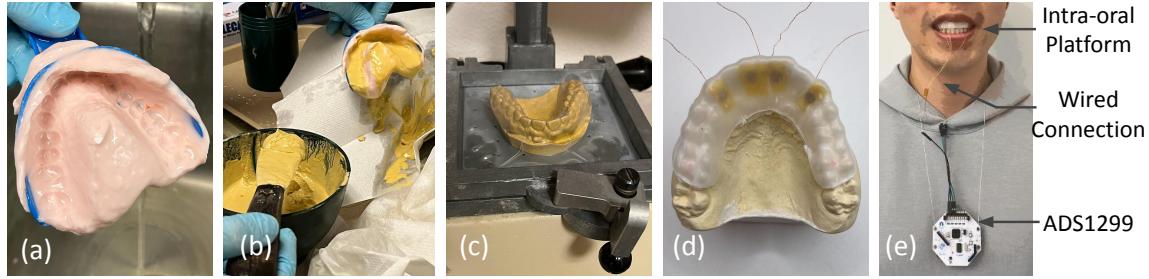
Fig. 7. a) Power spectrum of the output signal of the custom-built sensor in one of the tests b) Showcasing its energy concentration c) Power spectrum of the output voltage in the absence of any input stimuli (noise)

We consider sensors that cannot output voltage when mechanical stress is applied as broken sensors. Secondly, we verify the fabricated sensors are generating a sufficient signal amplitude relative to the noise level. In this paper, we empirically set the signal-to-noise ratio of 40 dB when the impact rate is 240 RPM as the standard. We also check the consistency of each sensor compared to other fabricated sensors. Based on the pattern of the generated signals under these tests, we can identify whether the sensor is functioning as expected.

We design the test bench as a platform that can repetitively and constantly hit the sensors to simulate the impact and vibration caused by human teeth. The test bench comprises of three main components, illustrated in Figure 6 (a). The first and main component is an impact head consisting of a flywheel, a hammer, and a rod. The rotation of the flywheel moves the hammer up and down to create impacts on the testing sensor. The driven power of the flywheel comes from the second component: a brushless motor with a 1-25 rate gearbox. The gearbox reduces the high-speed motor rotation 2500 revolutions per minute (RPM) to 12000 RPM into a slow range of 100 to 480 RPM, which is suitable for testing operations. The third component is an RCbenchmark data collection circuit [72] for controlling and recording the brushless motor speed. We design three different types of impact heads for the second test in the characterization procedure: an acoustic panel for the damping characterization, a flat hard surface and a pointy hard surface for the reference.

To measure the sensitivity of the fabricated sensor and guide the selection of the data acquisition device for the customized sensor, a PicoScope, 16-bit high-resolution oscilloscope [65], is used to read the sensor's output during

Fig. 8. Sub-figure a,b, and c show the making procedure of a teeth model plus its retainer. Sub-figure d shows the embedding of customized sensors between retainers. Sub-figure e depicts the wearing of *IOTeeth* system.

characterization. Figure 7 (a) and (b) shows the power spectrum of impact-induced signal from the PicoScope and zoomed-in visualizations, which illustrates the frequency domain's power distribution. Figure 7 (c) depicts noise's power spectrum. We find the energy of noise is lower than $-40dB$ and most of the energy of the output signals falls below 200 Hz. These findings indicate we can select a low-cost, low-power data acquisition device.

4.3 *IOTeeth* Sensing Platform Integration

We adopt a customized procedure to integrate the *IOTeeth* intra-oral sensing platform so that the wearable device can fit each user's unique facial musculoskeletal profile and maintain a low-interference wearing experience. The integration consists of two steps: 1) building dental retainers and embedding the piezoelectric sensors to build the wearable device, and 2) connecting the wearable device with the data acquisition device. Figure 8 (a-c) illustrates the process of making one layer of retainer by a professional dentist. First, the dentist duplicates the person's teeth shape and size with an alginate impression, as shown in Figure 8 (a). Then an inverted model is built by plaster from the alginate impression (Figure 8 (b)). After the plaster model is solidified, it is then put on a thermovacuum machine. One sheet of retainer material is heated and softened by a heat source above the machine. When the sheet has reached an appropriate temperature and softness, as confirmed by the dentist, it is formed into the shape of the plaster model with the thermovacuuming technique, as shown in Figure 8 (c). Finally, the dentist cuts out the outskirt and trims the retainer to a suitable size based on expert knowledge. By repeating the step in Figure 8 (c) with one retainer placed on the teeth model, the second retainer is built with a tight fit to the first retainer. We embed four customized sensors between retainers on locations of front teeth and canines' center area and fix them with one thin layer of Kapton tape, as shown in Figure 8 (d). The center alignment ensures a standardized approach while allowing for individualized customization of each person's unique oral dimension.

Next, we connect the intra-oral wearable to the data acquisition device. Guided by the sensitivity characterization (Figure 7), we select the commodity OpenBCI Cyton board as the data acquisition device for its high resolution (24-bit Σ-Δ), multi-channel input, convenience to use, and lightweight for comfort wearing. Its high resolution enables the *IOTeeth* to capture a fine-grained pattern of transit teeth contact-induced vibration signals. Its ADS1299 ADC has eight channels, each samples at 250 Hz, which is sufficient for capturing the information from customized sensors. The low power consumption of ADS1299 also allows the system to be portable without the need to change the battery during daily long-term wearing. Figure 8 (e) depicts the wearing of *IOTeeth*. The person wears the intra-oral wearable in the mouth, and the platform connects to the data acquisition device hung on the chest.

4.4 Signal Processing

Signals captured by the intra-oral wearable are further processed to 1) remove non-stationary trends, and 2) extract segments of interest for following inference steps. The piezoelectric sensor has been known for showing a non-stationary trend, which is a systematic increase or decrease in the obtained signal due to movements of the system or change of the temperature [27, 88]. This trend adds the noise for the following prediction tasks, therefore, the first step of signal processing is detrending the signal. We adopt the temporal differencing method for the detrending process. Given a time-series signal T consists of $[t_0, t_1, ..., t_n]$, the temporal difference can be written as $[t_2 - t_1, t_3 - t2, ..., t_n - t_{n-1}]$. This process removes the series dependence on time, hence it eliminates the trend while keeps other information in the data [19]. The comparison of a signal before and after the detrending is illustrated in Figure 3.

The teeth contact-induced vibrations result impulsive piezoelectric signals, which we define as *event*s. To extract those events, we conduct an energy-based event detection. We first apply a sliding window to the acquired sensor signal. When there is no event happening (e.g., before the sensor is worn), we establish a Gaussian noise model with the energy of the windowed signal. Then we conduct an outlier detection to new incoming windowed signals based on this Gaussian noise model [49]. The windows whose signals are detected as outliers are considered events. The event detection is conducted individually for each sensor then aggregate to generate the final event detection result. We adopt a greedy algorithm over all sensors to avoid missing detection and aggregate the detected events. For each step in the sliding, if there is one window that detects an event, *IOTeeth* considers it an event for all sensors. When there are multiple consecutive windows with detected events, if the last window's start index is before the first window's end index, we consider them as one event, and select the window with the highest signal energy to represent this event. On the other hand, if the last window's start index is after the first window's end index, we conduct a backward search in these consecutive windows. Once the window whose start index first comes after the first window's end index is found, we separate the consecutive windows into two parts and consider them as (at least) two events. We repeat the separation process to the second part until there is no more separation needed. The practice of selecting the window with the highest energy as the representative window ensures the extracted events are of the same length.
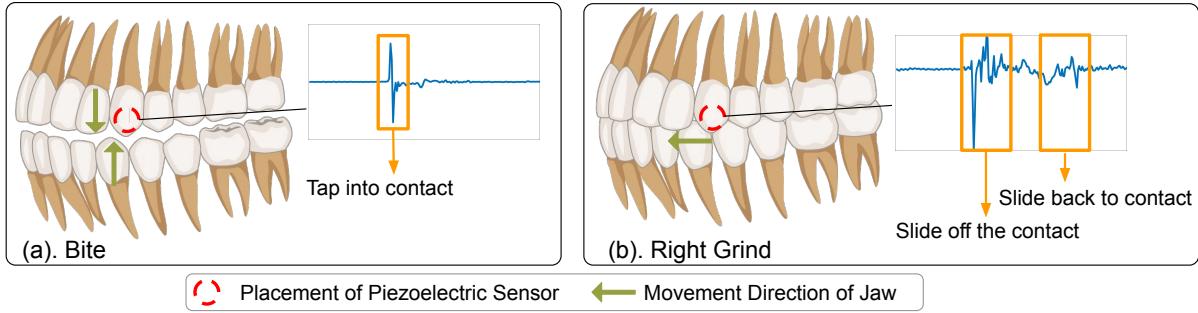
5 ACTIVITY AND OCCLUSAL DISEASE RECOGNITION ALGORITHMS

In this section, we introduce the second module of *IOTeeth*. This section contains three parts, namely, the activity recognition model (Section 5.1), the motivation experiment for the design of the occlusal disease recognition model (Section 5.2.1), and the proposed Physio-aware Attentional Network for predicting occlusal disease (Section 5.2.2).

5.1 Activity Recognition

IOTeeth focuses on recognizing biting and grinding, which are the anatomically atomic oral activities and they respectively represent functional and parafunctional activities in the human mouth [15, 70]. Biting refers to the activity that the upper and lower teeth fit together ***from untouched to contact*** [74], which can be mainly considered as a movement in the vertical direction. Grinding refers to the jaw moving side to side when upper and lower teeth are ***in contact***, hence it induces the sliding of upper and lower teeth against each other [89]. Figure 9 illustrates two activities' movements and exemplary signals captured by a piezoelectric sensor placed on the left canine's internal surface. The movement difference makes the signal have distinctive patterns. The biting signal is more impulsive and has fewer and shorter fluctuations than the grinding signal.

Given the distinct pattern, we propose to leverage the Convolutional Neural Network (CNN)-based model to recognize activities from signals for its strong performance in local salience pattern extraction [54, 95]. The input to the model is signal events normalized by its energy, hence the signal pattern won't be dominated by the absolute

Fig. 9. Two types of oral activities and respective signal examples.

amplitudes. The pattern extraction network consists of three stacked convolution layers. Each convolution layer extracts features from the input signal or the output from the previous layer through a convolution operation with a kernel. The 2-D kernel has a width of four, which equals the number of signal channels, and it moves along the time direction only, to simultaneously extract four channels' temporal pattern. Between each convolution layer, there is one batch normalization layer [21] to stable the training and one rectified linear unit (ReLU) as the activation function. After the second convolution layer, we apply a max pooling layer to extract the most representative feature and downsample the feature size. The output of the CNN pattern extractor is sent to the classifier network, which consists of three feedforward layers and ReLU activation functions in between. The output from the classifier network (denoted as o) is sent to a softmax function σ to calculate the prediction probability of each class:

$$\sigma(\mathbf{o})_i = \frac{e^{\mathbf{o}_i}}{\sum_{j=1}^{J} e^{\mathbf{o}_j}}, \tag{2}$$

where i is the ith class in the total number of J classes. The class with the highest prediction probability is the predicted label $\hat{y}$

$$\hat{y} = \underset{i}{\mathrm{argmax}}\ \sigma(\mathbf{z})_i \tag{3}$$
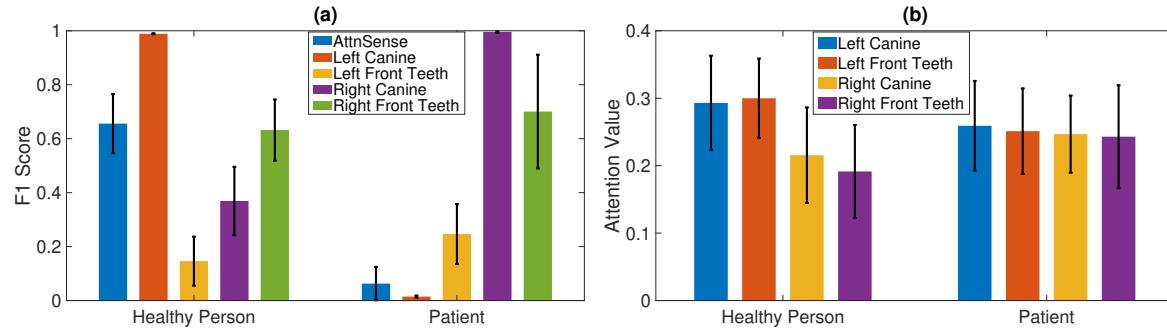
We employ the cross-entropy loss to train the activity recognition model:

$$\mathcal{L}_{Act} = -\sum_{j=1}^{J} y_j log(\sigma(\mathbf{o})_j) \tag{4}$$

, where y_j is the label of current input signal event. After being predicted the activity, each signal event will be sent to the occlusal disease recognition module to let *IOTeeth* recognize if the user has occlusal diseases. In the deployment, users can also set a prediction confidence threshold (such as 80%) for selecting signal events. In cases where a user performs a complex motion, like closing the mouth followed by sliding the teeth, *IOTeeth*'s event detection might return multiple signal events. However, only those events with high confidence, where signal events predominantly align with one of the two activities, will be considered for occlusal disease recognition.

5.2 Occlusal Disease Recognition

The occlusal disease recognition algorithm analyzes the signal events and predicts the possibility of having occlusal diseases. Even with just two classes, the task is challenging because the vibration sensor captures the skeletal information, and the algorithm needs to infer the underlying musculoskeletal health. Also, different people have their own unique dental profiles, making it critical to learn a robust feature for inference. On the
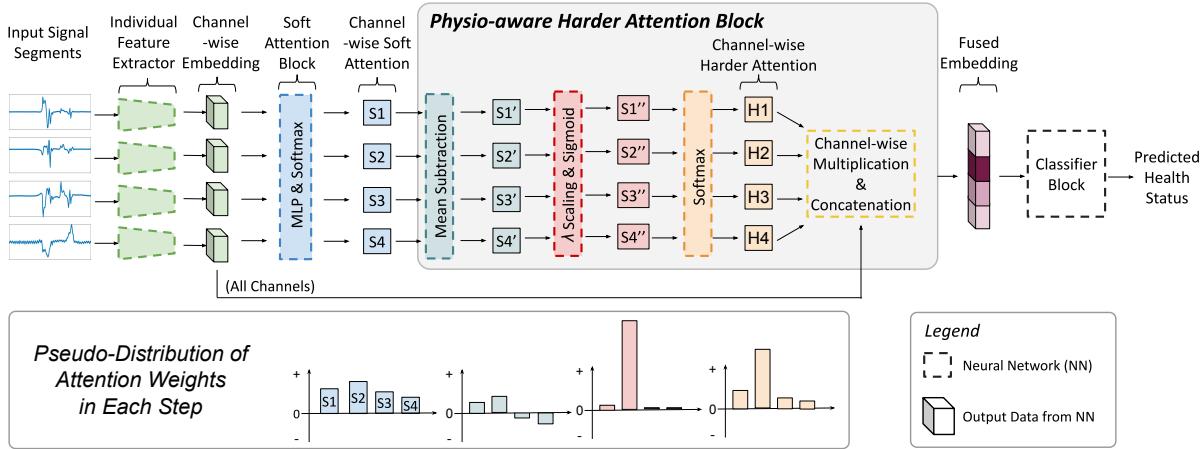
Fig. 10. Motivation study: Ten-time leave-one-out test result of randomly selected one healthy person's and one patient's biting signal events using AttnSense model (with four channels' signal) and individual channel signal. Sub-figure (a) shows the test F1 score of each model. Sub-figure (b) shows the attention value of each channel calculated by the AttnSense model.

other hand, the vibrations resulting from teeth contact inherently reflect the dynamics of dental arches, teeth morphology, and musculoskeletal condition, serving as implicit occlusal disease indicators for the algorithm. Considering the critical difference in the signal pattern between biting and grinding activities, we design occlusal disease recognition to have two independent but identical neural network models for signals of different activities, respectively. This subsection is organized into two parts. In the first part, we conduct a motivation study to investigate the efficacy and challenge of fusing the signal from multiple sensors in the *IOTeeth*. The insights from the result motivate the design of our proposed Physio-aware Attentional network (PAN Network).

5.2.1 A Motivation Study. Sensor fusion has been successfully applied to multi-sensor systems for different inference tasks, e.g., human activity recognition [96]. The state-of-the-art fusion algorithms like AttnSense [44] and Cosmo [55] propose to leverage attention-based fusion techniques to combine the information in all sensor's signal. These works assign a weight of 0 to 1 for each sensor channel, known as soft attentions [11]. However, predicting occlusal disease is a much more complex task than human activity recognition due to the high variance across each individual [81]. A patient diagnosed with occlusal disease may not have every tooth or dental area impacted by the condition, and the severity of the disease can vary greatly among affected teeth area [53]. In such circumstances, if the model pays attention to the tooth that is not impacted, the model is prone to fail. Similarly, a dentist-identified healthy individual might exhibit early signs of occlusal diseases in one or two teeth area[13]. The signals from those teeth are *noisy and misleading* to the occlusal disease recognition task. Soft attention is prone to have difficulty in distinguishing the important channels from the irrelevant ones when the inputs are noisy [33, 46]. Therefore, a model that can emphasize the most informative one while suppressing the contribution of potentially noisy sensors is needed for occlusal disease recognition. To justify our argument, we conduct a feasibility experiment.

We randomly select two test persons, one is dentist-labeled healthy and the other one is a patient with occlusal diseases. The experiment is conducted in a leave-one-out setting: the test person's biting signal events will be used to evaluate the model trained by the other 11 person's biting signal events. The procedure and detail of data collection will be elaborated in Section 6.1. We select the AttnSense model as the representative soft attention model, which has an attention-based module that dynamically learns the weights for combining features of different sensors [44]. The input to the AttnSense model is the four channels' signal. Also, we have a CNN-based model that takes only one channel as the input to train and test the model. The CNN-based model has an identical architecture to the model proposed in Section 5.1. We repeat the experiment 10 times and report the averaged F1 score as the metric.

Fig. 11. PAN Net overview. The PAN Net takes extracted signal events as the input, then sends them to the channel-wise feature extractor. The extracted channel-wise embedding from feature extractors is then sent to the soft-attention block and physio-aware harder attention block to calculate the attention weights. The embeddings are finally fused with attention weights and are sent to the classifier block for prediction. We depict the pseudo-distribution of attention weights in each step to show the designed change of those weights.

The recognition F1 score of five models for two test persons' data is illustrated in Figure 10 (a). We observe the AttnSense model has limited performance. It achieves an average of 0.656 and 0.062 F1 scores for healthy person and the patient's data. On the other hand, the CNN model with individual sensor input has drastically varied results. For the healthy user, if the model takes the left canine's signal into account, it can achieve an F1 score of 0.989. However, if the model takes the left front tooth's signal as the input, the test F1 drops to 0.146. Also, the left canine is not the gold sensor placement location for other cases. For the patient, the model with the left canine's signal achieves a 0.0146 F1 score, while the model with the right canine's signal input achieves a 0.995 F1 score. These results show that sensors placed on different teeth contribute significantly different to occlusal disease recognition. When the signal from negatively contributing sensors is combined, it adds noise to the data and reduces the information of the data from the most contributing sensor. Figure 10 (b) shows the distribution of the mean of the AttnSense model's attention weights put to each sensor signal in test data. We observe that the distribution of both test cases is centering around 0.25. This distribution indicates that the model takes four sensor inputs almost equally, instead of paying more attention to the informative sensor channel and neglecting the sensor which has misleading information. This experiment leads us to the requirement of a generalizable occlusal disease recognition model: a model can automatically highlight the most informative sensor channel for inferring the unseen new data.

5.2.2 Physio-aware Attentional Network. In this section, we introduce our proposed Physio-aware Attentional Network (PAN Net). The design goal is: the model can automatically amplify the attention weight of the most informative sensor/channel for occlusal disease recognition given the input, and surpass the attention weight of other channels. We propose to sharpen the distribution of attention weights with a harder attention block. In this study, we build our harder-attentional network upon the AttnSense model. The overview of the network is illustrated in Figure 11. The network consists of individual channel-wise feature extractors, a soft attention block, a harder attention block, and a classifier block.

Feature Extractor. The feature extractor takes extracted signal events as the input and output embedding (feature) vector of each channel. We adopt the same structure of three CNN layers used for activity recognition. After each layer, we append a dropout layer to regularize the training [82]. The different part from that in the activity recognition model is we utilize individual CNN feature extractors for each sensor's input. As shown in Figure 11, the extracted signal event of k-th sensor, x_k is sent to its own CNN feature extractor, which outputs a channel-wise feature (embedding) vector v_k. All K sensor's embedding vectors are sent to the soft attention layer.

Soft Attention Block. The soft attention block, which is identical to the one used in the AttnSense model, takes sensors' embedding vectors $[v_1, v_2, ..., v_K]$ as the input and returns an attention weight $[S_1, S_2, ..., S_K]$ for each sensor. It first projects each sensor's embedding vector to a latent space by one feedforward layer, and we denote this latent representation as μ_k. Then the block calculates the weight of k-th sensor by measuring the similarity between μ_k and a learnable sensor context vector C [44]. Finally, the calculated weights of each sensor are normalized via a softmax function to get the final soft attention weights. The calculation steps can be formed as follows:

$$\mu_k = tanh(W \cdot v_k + b), \tag{5}$$

$$S_k = \frac{exp((\mu_k)^T C)}{\sum_1^K exp((\mu_k)^T C)}, \tag{6}$$

where W and b are the weight and bias parameters in the feedforward layer, and $tanh$ is the hyperbolic tangent activation function. The parameter W, b, and sensor context vector C are randomly initialized and jointly learned during the training process [44].

Physio-aware Harder Attention Block. The physio-aware harder attention block automatically amplifies the most informative channel(s)'s attention weight while decreasing other uninformative ones'. The harder attention block consists of four arithmetic steps: mean subtraction, scaling, sigmoid function, and softmax function. The mean subtraction step lets each sensor's soft attention weight minus the mean value of all sensor's weights. This step pushes attention weights center around zero. After this step, we denote the mean-subtracted S_k as S'_k. The S'_k will be further scaled by a scalar parameter, λ, then pass a sigmoid function. This step pushes the positive values close to 1 and negative values close to 0, i.e., semi-binarized distribution. This is realized by leveraging the sigmoid function's property: the λ is also known as the gain value that controls how steep the sigmoid curve is. The gain value is a hyperparameter to be set for the PAN Net. When it equals one, for input that is bigger than 10 or smaller than -10, the output is almost one (> 0.999) or zero (< 0.001). The higher the gain value, the more sensitive the output is to the input. We denote the processed S'_k as S''_k. Each person's musculoskeletal profile is unique , thus the input data also has a high variance across different persons. The semi-binarized attention distribution is prone to cause training oscillations when the input data has a high variance [36, 69], hence we append an additional softmax function to slightly smooth the distribution. The calculation can be formed into the following:

$$S'_k = S_k - \frac{\sum_1^K S_k}{K}, \tag{7}$$

$$S''_k = \frac{1}{1 + exp(-\lambda S'_k)}, \tag{8}$$

$$H_k = \frac{exp(S''_k)}{\sum_1^K exp(S''_k)}, \tag{9}$$

where the H_k is the harder attention weight. The last step is the multiplication between each channel's harder attention weight and its embedding vector, then concatenate into the fused embedding Z for the classification:

$$Z = \langle H_1 \cdot v_1, H_2 \cdot v_1, ..., H_K \cdot v_K \rangle, \tag{10}$$

where $\langle \rangle$ represents the concatenation operation.

Classifier Block. The classifier block takes the fused embedding as the input and outputs the occlusal disease recognition, i.e., if these events show a sign of having a occlusal disease. Inspired by AttnSense and other previous signal classification studies, we adopt a two-layer Gated Recurrent Unit (GRU) with a three-layer MLP as the classifier block. The logits output from the MLP pass a softmax function to calculate the prediction probability of each class and the highest one is the predicted label. We employ a cross-entropy loss for training the occlusal disease recognition model. The forward pass's calculation can be expressed as:

$$Q = \text{MLP}(\text{GRU}(Z)), \tag{11}$$

$$\sigma(\mathbf{Q})_i = \frac{e^{\mathbf{Q}_i}}{\sum_{j=1}^{J} e^{\mathbf{Q}_j}}, \tag{12}$$

$$\mathcal{L}_{Health} = -\sum_{j=1}^{J} y_j log(\sigma(\mathbf{Q})_j) \tag{13}$$
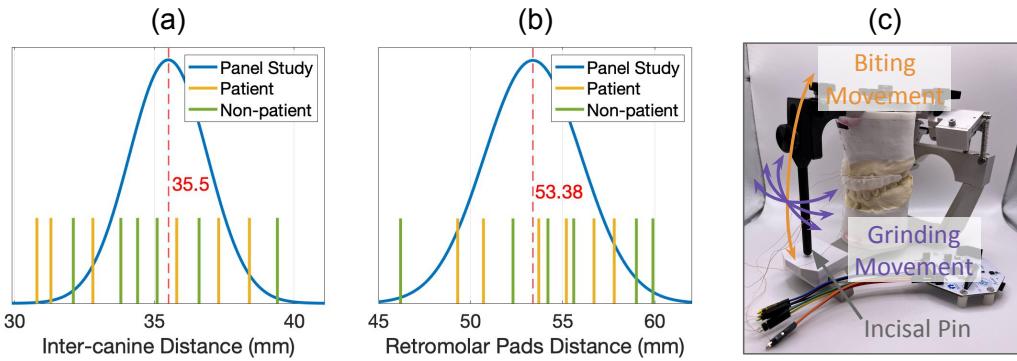
where i is the ith class in the total number of J classes.

6 EVALUATION

In this section, we describe the implementation details and setups of data collection (Section 6.1), introduce data-driven learning baselines and the metrics for evaluation (Section 6.2), and present the system's performance in activity and occlusal disease recognition in Section 6.3. We also examine our customized piezoelectric sensors (Section 6.4), assess sensor placement's interference to teeth fitting (Section 6.5), and conduct the ablation study and sensitivity analysis for the proposed PAN Net in Section 6.6 and Section 6.7.

6.1 Data Collection

6.1.1 Teeth Activity and Occlusal Disease Data. Despite our best efforts, we were unable to ensure that all required source materials for building the *IOTeeth* intra-oral sensing wearable are FDA-approved. Therefore, in regard to bio-safety, we use teeth models installed on articulators to replicate teeth contact and intra-oral activities for data collection. An articulator is a mechanical tool used by dentists to fix plaster teeth models and reproduce jaw movements. In total there are 12 plaster teeth models provided by a professional dentist clinic are used for data collection. Those teeth models are randomly selected and they reproduce the teeth contact conditions of the people when they are in their natural states. Six out of 12 teeth models are built from patients diagnosed with occlusal diseases, while the rest six models are built from people who a dentist identifies to be healthy. To understand how representative these 12 models are, we compare two anatomic landmarks with the large-population general distribution from a panel study [3]. The inter-canine distance and retromolar pad distance are anatomic landmarks commonly used for assessing the overall dental arch form, potential spacing or crowding issues, and occlusions [71, 92]. Results shown in Figure 12 (a) and (b) illustrate these randomly selected teeth models encompass a comprehensive range of both anatomical landmarks, ensuring a broad representation of dental variations and a fair evaluation of the *IOTeeth*'s generalizability. To collect the data, we mount each teeth model to the articulator as shown in Figure 12 (c). Guided by the dentist, we operate the incisal pin of the articulator to replicate the jaw movement of human's biting and grinding activities. For each teeth model, we collect 150 times of biting movements and 150 grinding movements for left- and right-directional grinding. The

Fig. 12. Testing models' dental anatomic landmark illustration (sub-figure a and b), and occlusal disease data collection illustration (c). Sub-figures (a) and (b) show the comparison of the dental anatomic landmark between testing models' measurements and general human adults' distribution from a panel study [3]. Blue curves show the distribution of anatomic landmarks. Each vertical bar indicates one tooth model's measurement. The random-selected testing models cover most of the range. Sub-figure (c) shows the mounting of the articulator teeth model and the illustration of data collection movements.

force, speed, angle, and extent of those movements are randomly adjusted each time to mimic human nature moves. During the data collection, we log the time of each movement. In total, we collect over 5000 movements from 12 different teeth models.

6.1.2 Sensor Sensitivity Data. To validate the design of the sensor and a comprehensive characterization, we conduct two sensor sensitivity experiments and collect the sensor output. First, we collect the sensor output of different diameter sizes (2/3/4/5/6/7/8/9/15 *mm*) with impacts generated by the test bench. Then, we collect the output of the 4*mm* sensor with different testing settings following the designed sensor characterization test introduced in Section 4.2. For each characterization experiment, the speed of the test bench starts at approximately 90 *RPM*, i.e., every two seconds there are three impacts. We then gradually increased the speed by an incremental step of 18 *RPM* and stop at 460 *RPM*, recording the signals generated by each impact at the same time with the OpenBCI Cyton board. In total, for each experiment, there are around 10 impacts. Throughout the experiment, we monitored the RPM of the brushless motor with RCBenchmark at 40*Hz* to ensure a precise speed adjustment.

6.2 Baseline Models and Implementation

We compare our system's performance with the following baseline models.
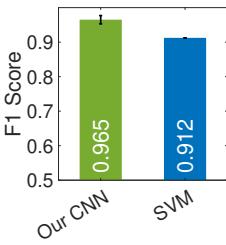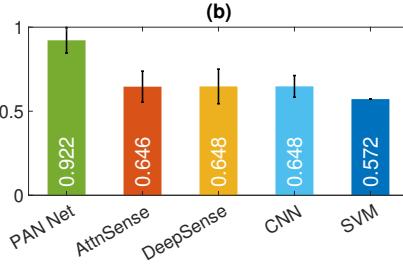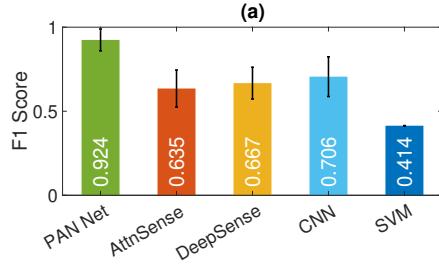
Support Vector Machine (SVM) SVM is a shallow machine learning model that has been successfully applied to vibration signal classification tasks. Inspired by prior works, we extract each sensor signal's frequency (1-125 *Hz*) component and concatenate them together as the feature [60].

Convolutional Neural Network (CNN) A CNN-based model has a same structure to the one for activity recognition is used as the baseline for the disease prediction task.

DeepSense DeepSense is one of the state-of-the-art sensor fusion approaches for multi-sensor signal classification task. Features from different sensors are concatenated together for fusion [96].

AttnSense AttenSense is an attention-based supervised multi-sensor fusion algorithm which dynamically assigns weights to features from different sensors, then fuses the weighted feature [44].

Given the complexity of activity and disease recognition models in *IOTeeth*, the activity recognition task will be evaluated with SVM only. The SVM model is tuned with a hyperparameter grid search with 3-fold cross-validation in the training data, and the best hyperparameter setting is used for predicting the testing data. All deep learning

Fig. 13. Average leave-one-out occlusal disease detection results across all participants of two activities. Sub-figure (a) is with biting activities and sub-figure (b) is with grinding activities.

Fig. 14. Average leave-one-out activity recognition results across all participants.
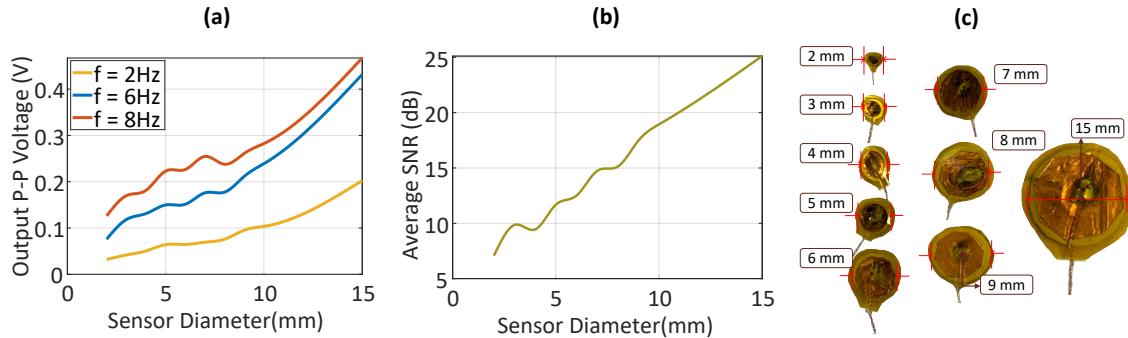
models are implemented with PyTorch [62] framework. To ensure a fair evaluation, all deep learning models are trained with the same hyper-parameters: the batch size is set as 512, the learning rate is 0.002 and the training epoch number is 50 with the early stopping. The λ in the physio-aware harder attention block of the PAN Net is empirically set as 100. The *IOTeeth* system is designed to report an aggregated result to the user, but for a quantitative test and a fair comparison with baseline models, we evaluate all models' performance at the event level. We use the F1 score as the evaluation metric, calculated as:

$$ \text{F1} = \frac{\text{TP}}{\text{TP} + \frac{1}{2}(\text{FP} + \text{FN})}, \tag{14} $$

where TP is the number of signal events that are true positives, FP is the number of false positives, and FN is the number of false negatives. Each experiment is in the leave-one-out setting: one person's data is used as the testing data and the rest 11 person's data are used for training. We repeat the experiment of each person's data as the testing data 20 times and report the mean F1 score as the final result.

6.3 Overall Performance of Benchmarking Test

We conduct the benchmarking test to evaluate the effectiveness of *IOTeeth* in detecting occlusal disease and recognizing different activities, with both training and testing data collected from articulator-mounted dental models replicated from real-world dental clinic customers. The occlusal disease detection results are evaluated and discussed independently since there is no dental reference for a fair fusion of the two activities' detection results. Figure 13 shows the average leave-one-out test F1 score of occlusal detection across all participants' models (hereafter abbreviated as "participants"). *IOTeeth*'s PAN Net achieves an average of 0.924 and 0.922 F1 score in the leave-one-out test setting, which indicates that our system is accurate and robust to various participants. All deep learning-based models significantly outperform the SVM baseline by a clear margin, which illustrates the efficacy of extracting the signal feature with the CNN-based feature extractor layers. Compared with deep learning baseline models, PAN Net achieves at least 0.2 higher F1 score. This confirmed the necessity of adding the physio-aware harder attention block to the network for selecting the most contributing sensor/channel. The AttnSense model and DeepSense model have comparable results in both activities. Considering the only difference between these two models is that DeepSense combines the channel-wise embedding without a soft attention weight, the comparable results of the two models indicate that soft attention fails to pay more attention to the informative sensor but pays roughly equaled attention to all sensors' data. The CNN model achieves the second-to-best result, outperforming both AttnSense and DeepSense models. This phenomenon suggests a few

Fig. 15. (a) Relation between the output voltage of the sensors compared to its size (b) Average SNR in recorded tests for different sizes of the sensor (c) Some of the fabricated sensors in different sizes.
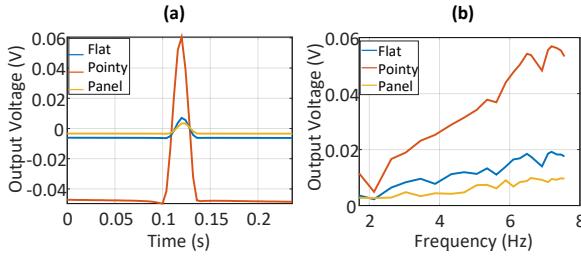
key insights. First, the CNN model processes all four channels' features simultaneously, which is able to model inter-channel relationships more effectively in the early stage. Also, the max pooling operation in the CNN model acts like the Physio-aware Harder Attention Block that keeps the most informative pattern in the signal and discards the rest. Unlike AttnSence and DeepSense models that extract individual channel's features and then attempt ineffective fusion, the combination of early-stage inter-channel modeling and pooling operation makes the CNN model perform better in the occlusal disease recognition task. Additionally, the absence of recurrent neural network (RNN) layers after the CNN model's feature extractor indicates for intra-oral activities-induced vibration signals, RNN layers for modeling the temporal dependency are not a core contributor to the model's performance. Also, we observe that each model achieves a comparable F1 score in both activities. We believe this phenomenon is potentially related to the nature of these oral activities and their relationship to occlusal disease. Biting and grinding are two distinct activities that can provide different perspectives on a person's occlusal health, as they can contribute to the development and progression of various occlusal disease conditions [41, 67]. These comparable results from both activities also align with the real-world dental practice, where dentists evaluate and consider multiple factors, including both functional and parafunctional activities, when diagnosing and managing occlusal diseases [16, 80].

Figure 14 depicts the average leave-one-out test F1 score of activity recognition of the CNN-based model and SVM baseline model. We observe that both models achieve an over 0.9 F1 score, and our CNN-based model outperforms the baseline model by a 0.05 F1 score. This result indicates that signals of different activities show consistently distinctive patterns, and the design of a CNN-based activity recognition model that is driven by extracting patterns is validated. Also, these results show the system can accurately send each signal event to the according disease detection model, which ensures the *IOTeeth*'s overall disease detection accuracy.

6.4 Sensor Sensitivity Analysis

This subsection characterizes the sensors' performance, sensitivity, and reliability with multiple sizes in multiple settings. We particularly focus on the *4mm* sensor as it is the smallest sensor that can be manually crafted.

Signal outputs of different sizes of sensors. Figure 15 shows the comparison of sensors in different diameter sizes and measurements of output. From subfigure (a), we observed a positive correlation between the size of the sensor and the output voltage it produces. This phenomenon is aligned with Eq. 1, the output voltage (V) increases with the size (L). Also, from subfigure (b), we observe that a bigger size of the sensor can output signals with a higher signal-to-noise ratio (SNR). This result indicates that during the design of customized PVDF piezoelectric sensors,

Fig. 16. (a) Signal pattern of the 4mm sensor with different impact heads. (b) Relation between the frequency of the impact and the peak-to-peak voltage of the output signal.

Table 1. Results obtained from the test using OpenBCI.

	SNR dB (min, max)	Average Correlation
Pointy	38, 44	99.8%
Panel	20, 24	98%
Flat	9, 12	99%
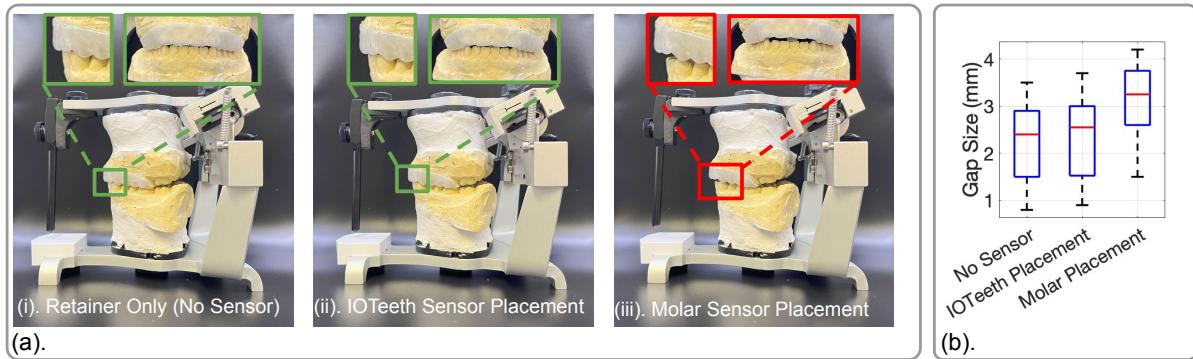
we should select the biggest size that is acceptable. Given the statistic of human canines' and front teeth width of seven millimeters with a variance of three millimeters [93], the 4-*mm* diameter design is at the balanced point of adapting to most person's teeth and keeping a high sensitivity and low noise.

Signal outputs of different impact heads. Figure 16 (a) shows the signal pattern of different impact heads. From this figure, we observe that compared to the output from pointy impact, both acoustic panel vibration and flat impact induce a weaker signal. This phenomenon coheres with Eq. 1, the output voltage (V) decreases when the area of receiving mechanical stress (A) increases. Table 1 shows the quantitative results of the measured signal-to-noise ratio. The output signal from the pointy impact head shows the highest SNR range, indicating that the sensor is most sensitive to direct impact with a small impact area. Comparing the SNR of signals from the acoustic panel and signals from the flat impact head, the sensor is more sensitive to vibration. The SNR of signals from the acoustic panel is significantly higher than the other, even they have similar amplitudes. This result shows that the customized sensor is suitable for being embedded in between retainers for vibration sensing.
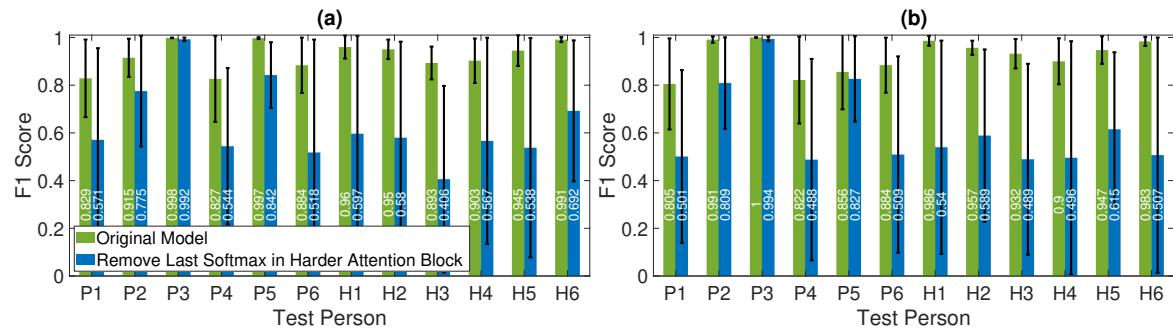
Signal outputs of different impact frequency. Figure 16 (b) shows the relationship between the frequency of mechanical stress applied and the output's amplitude. We observe that the output with the panel test head has the least change over different frequencies, indicating high stability and consistency. This result indicates the sensor is able to stably generate signals with identical patterns when the oral activity frequency changes. Table 1 shows cross-correlation between each sensor's output given the same mechanical stress. All three tests' results show an over 98 % average correlation coefficient, which indicates the sensor fabrication error and the variance between each sensor are low. The results from two testing illustrate the customized sensor has a high consistency.

6.5 Sensor Placement

This experiment aims to examine the efficacy of *IOTeeth*'s sensor placement design in reducing the interference. We compare the occlusal fitting conditions with three sensor placements: 1). there is only a retainer with no sensor, 2). the sensor is placed with *IOTeeth*'s front teeth and canines placement design, and 3). placing two sensors in the front teeth side and two sensors in the location of the left and right first molars. The molar placement is commonly adopted in previous works [26, 58]. The "no sensor" setting represents the natural occlusion condition when a person is wearing a retainer, which is used as the reference. Figure 17 (a) shows an example of occlusal fitting with different settings. We observe that the *IOTeeth* placement does not introduce significant interference to the occlusal fitting compared to the "no sensor" setting. On the other hand, the molar placement causes a gap between the upper and lower front teeth due to the amplification effect of the human jaw's hinge joint structure. This gap indicates the deviation from the natural fitting. In order to quantify the interference, we measure the

Fig. 17. Sub-figure (a) shows an example of sensor placement and its interference with the teeth fit. (a-i) shows the teeth fit with the retainer but with no sensor placement. (a-ii) shows the teeth fit with the *IOTeeth*'s placement. (a-iii) shows the teeth fit when moving two canine sensors to the place of the two first molars. The molar placement causes an open bite which significantly interferes the normal teeth fitting. Sub-figure (b) shows the upper and lower front teeth gap size distribution of different sensor placements.

Fig. 18. Ablation study of removing the last softmax layer in the leave-one-out test setting. "P" stands for the patient, while "H" stands for the healthy person. Sub-figure (a) is evaluated with biting activities, and sub-figure (b) is evaluated with grinding activities. For simplicity, both sub-figures share the same legend.
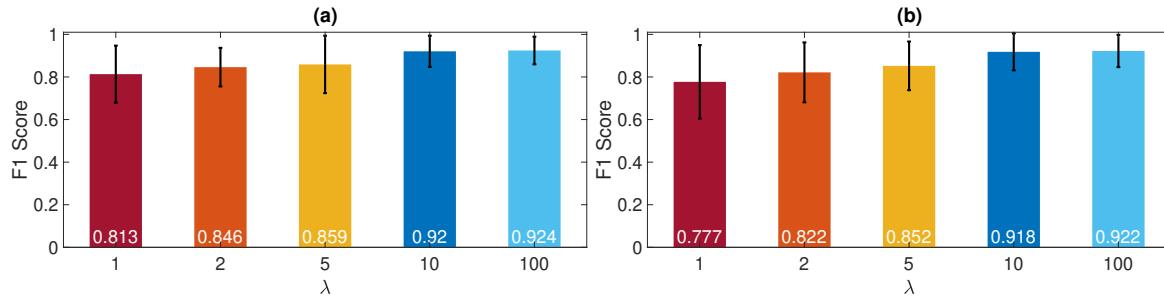
gap distance between the upper and lower front teeth with a feeler gauge as an indicator of the occlusal fitting, which is commonly used in dental practice [48, 87]. The feeler gauge we use has a precision of 0.05 *mm*. The distribution of 12 models' front teeth gap distance in three settings is shown in Figure 17 (b). From the result, we observe the molar placement significantly enlarges the front teeth gap, and the median value of the gap increases from 2.4 *mm* ("no sensor" setting) to 3.25 *mm*. On the other hand, the *IOTeeth* placement has a median of 2.55 *mm*, which reduces **82%** less interference compared to the molar placement. This result validates the design of *IOTeeth*'s sensor placement can minimize the interference to natural occlusal fitting.

6.6 PAN Net Ablation Study

In the design of the PAN Net's physio-aware harder attention block, we add one more softmax layer in the end to slightly smooth the amplified attention weights. In this ablation study, we compare the difference between models with and without the last softmax layer (denoted as the baseline model in this section). Figure 18 depicts

the F1 score of two models. We observe that for most participants, the PAN Net's mean F1 score of 20 rounds of experiments is clearly higher than that of the baseline model. One of the main reasons for this reduced F1 score is the high instability of the baseline model's performance. We notice that the baseline model's results have a significantly bigger range than that of the PAN Net, and the baseline model has a bigger standard deviation in the result for each of the test users. This phenomenon indicates that the training is unstable if the attention weights after the λ-scaling and sigmoid layer. The reason behind this is the semi-binarized attention weight introduces high randomness and uncertainty in the sensor selection process and makes the model easily fit into the data noise or outliers [32, 36, 69]. The added last softmax layer serves as the normalization to regularize the uncertainty for stable training. From the standard deviation of individual testing participants' results, we observe that the added last softmax layer achieves different levels of stabilization. For instance, for patients 1, 4, and healthy participant 4, PAN Net's test results show a high standard deviation. We explore different network designs and methods of regularization (e.g., batch normalization, layer normalization, dropout, weight decay, and schedule of training), but there are no fundamental improvements. In summary, a method for better regularization is an open question left for us.

6.7 PAN Net λ Value Sensitivity Analysis

Fig. 19. Sensitivity analysis of different λ value and average leave-one-out testing result. Sub-figure (a) is evaluated with biting activities and sub-figure (b) is evaluated with grinding activities.

The λ in the physio-aware harder attention block is the sigmoid function's gain value that controls how steep the sigmoid curve is. In this experiment, we examine the model's sensitivity to different λ values. For a vanilla sigmoid function ($\lambda = 1$), if the input is bigger than 10, the output is close to 1, and for the input value is smaller than 10, the output is close to 0. Therefore, we choose five different λ values for investigation: $\lambda = 1, 2, 5, 10, 100$. The value five is in the middle between 1 and 10, while the value two is in the middle between 1 and 5. The value 100 is the big value we select to test the model's performance when a very small (absolute value bigger than 0.1) input to the sigmoid can have an output close to 0 or 1. Figure 19 illustrates the testing result. From the result, we observe when $\lambda < 10$, the model is sensitive to the value. A bigger value can help the model to achieve a higher average testing F1 score. The model with $\lambda = 10$ and $\lambda = 100$ achieves a comparable result. For the two activities, we can observe identical sensitivity trends. This sensitivity trend is related to the function of the sigmoid layer in the PAN Net: a bigger λ value can amplify the attention weight of the most informative sensor closer to one. Therefore, compared to models with small λ values, models with bigger λ values are more effective in emphasizing the contribution of the most informative sensor's signal and surpassing others. Note that this does not equal that the PAN Net only focuses on one single channel. Even with a large λ, the softmax layer is able to mild the semi-binarized attention weight, letting the model take input from channels with lower

attention. Also, the model saturates with λ between 10 and 100, this indicates that there's an optimal range for λ. Beyond this, increasing its value doesn't provide significant performance gains.

7 DISCUSSION

7.1 User Study

We recognize one of the limitations is the lack of a real-person user study. The primary reason for this limitation stems from the challenges associated with acquiring FDA-approved materials for fabricating the sensors. As the sensors are to be embedded within retainers and placed in mouth, ensuring the safety and biocompatibility of the materials is of paramount importance. In the future, we will continue our efforts to identify and obtain medically safe materials for fabricating the sensors. Once the appropriate materials are found, we plan to conduct a user study with real human participants to validate the effectiveness of our system in recognizing oral activity and predicting dental disease. To enhance the precision of our data collection, we will integrate Electromyography (EMG) technology and record oral motion videos, thereby capturing detailed muscle activity and joint coordination during oral activities. With the data collected from human participants and labels from dental professionals, we envision the *IOTeeth* can not only recognize more complicated oral activities and fine-grained occlusal diseases, but also contribute to the broader understanding of dental health research.

7.2 Beyond Occlusal Disease Recognition

We envision *IOTeeth* as a versatile system extendable for a range of applications, including: 1) wellness monitoring, 2) assistive dental technology, and 3) interactive systems. One practical application is in the monitoring of involuntary oral activities during sleep. By tracking these activities, *IOTeeth* can offer valuable insights into an individual's daily stress levels, serving as an indicator of overall well-being and sleep quality [45]. Another application falls into orthodontics treatment [4]. By continuously monitoring the vibration patterns generated during teeth movements, the system could serve as a teeth fitting profiler and for pre- and post- teeth realignment treatments' reference. This would not only offer orthodontists a handy tool to gauge the progress of treatments but also provide them with timely data to make any necessary adjustments. The *IOTeeth* can also be integrated with other facial wearable systems, broadening its scope of applications. For instance, when combined with the TeethTap system [85], they could create a mouth-based human-computer interaction platform, offering users a convenient way to interact with their devices via both jaw movements and teeth contacts.

7.3 Musculoskeletal Sensing

This study focuses on skeletal sensing using teeth models on an articulator. In the future, we plan to extend our approach with a combination of related facial sensing systems to encompass musculoskeletal sensing. The primary motivation behind this direction is the recognition that real human data encompasses not only the anatomical bone structure and bone movement relationship but also the contribution of muscles to the motion of the masticatory system. For instance, the MuteIt system [83] provides the motion monitoring of the lower jaw and temporomandibular joint, while the FaceSense system [23] measures the facial physiological signals. Integrating those systems' functions with *IOTeeth* enables us to develop a more complete and accurate model of the masticatory system, potentially leading to better diagnostic and therapeutic strategies beyond occlusal diseases recognition, like fine-grained temporomandibular disorders diagnostic.

7.4 Smaller Sensor

In our current design and implementation, we utilize piezoelectric sensors with a dimension of $4mm$. While this configuration allows us to sense signals from oral activities like biting and grinding, there is room for improvement in terms of sensor size. As shown in previous sections we can still capture signals with a sensor size

of only 2*mm*. Smaller sensors would enable more comprehensive sensing, as we could potentially place multiple sensors on a single tooth. This increased sensing resolution may lead to finer-grained activity monitoring, as well as facilitate the identification of specific occlusal issues that may be challenging to detect with larger sensors. However, the current sensor size presents significant challenges for mass production, particularly in terms of soldering and assembly. Therefore, we plan to investigate advanced fabrication methods that can enable the production of smaller sensors in the future.

7.5 Data Acquisition Platform

We adopt the OpenBCI Cyton board for data collection, while it shows promise, there is an opportunity to enhance its usability and comfort by exploring new data acquisition platforms. For example, the current setting requires a wired connection, which limits the comfort for a long-time wearing. A wireless data collection platform can enhance convenience and comfort for users, as it eliminates the need for cumbersome cables and allows for greater freedom of movement during data collection. In the future, we will extend the current *IOTeeth* setting with a low-energy Bluetooth function. However, a wireless solution could raise safety concerns due to the presence of a battery inside the user's mouth and may significantly interfere with the user's occlusion. Therefore, we want to explore the capabilities of piezoelectric sensors [20]. By utilizing their voltage output, we aim to develop a wireless system that does not require an in-mouth battery. In addition, the extent of the wired solution's impact on users' life and experience is under-explored. The future real-life user study is important to gather feedback for guiding the wired or wireless design.

8 RELATED WORKS

8.1 Oral Sensing Systems

Attention has been paid to oral sensing systems with applications ranging from health characterization to hands-free device control. Table 2 shows a comparison of related oral sensing systems. These systems can be categorized into intra-oral and extra-oral systems. Intra-oral systems are generally built upon dental appliances that are to be worn inside the user's mouth, while extra-oral systems are wearable devices mounted around the user's ear and jaw. Notable extra-oral systems include (but are not limited to) MuteIt by Srivastava et al. [83] and TeethTap by Sun et al. [85], both utilizing jaw motion for device control. Pataranutaporn et al.'s Wearable Lab sticks to the cheek for facial and saliva sensing [63]. These systems, however, face challenges in capturing specific dental-level information due to the complex anatomy of the human mouth. In contrast, intra-oral systems focus on teeth motion and kinematics, using various sensors like piezoresistive force [26, 28, 31], IMU [28, 34, 94], pressure [57, 86], and piezoelectric [59]. Despite their ability to monitor teeth motion or kinematics, these systems fall short in recognizing occlusal diseases, primarily due to design limitations like mouthguard-based builds or large sensors that interfere with occlusion and limit sensing resolution. Additionally, the placement of batteries inside the mouth raises comfort and safety concerns. *IOTeeth*, however, stands out by recognizing occlusal diseases with minimal interference to occlusion. Its design allows for fine-grained teeth-level sensing without the need for an internal battery, addressing the limitations of existing systems. The retainer-form design of *IOTeeth* is also cost-effective compared to mouthguard-based systems like s-Guard [34]. Customizing two layers of retainer costs under 100 US dollars[2] and we envision the total cost of the *IOTeeth* system to be approximately 180 US dollars[3],

---

[2]The calculation is based on the dental lab technician's average wage in the USA (22 USD/hour in 2023, source: www.indeed.com/career/dental-technician/salaries) and material cost.
[3]This calculation is based on one STM32H743 Microcontroller Unit (MCU), one ADS1299 Analog-to-Digital Converter (ADC) and accessory materials.

Table 2. Comparison of Oral Sensing Systems

Systems	System Type	Sensing Modality	Motion/Kinematics Monitoring	Occlusal Disease Recognition	Low Interference	No Battery In Mouth	Fine-grained (Teeth-level) Sensing	Thickness (Lower is better)
Wu et al. [94]	Intra-oral	IMU	**Yes**	No	No	No	No	-
Kinjo et al. [26]	Intra-oral	Piezoresistive Force Sensor	**Yes**	No	No	-	No	4mm
Lantada et al. [31]	Intra-oral	Piezoresistive Force Sensor	**Yes**	No	No	**Yes**	No	5mm
Toma et al. [86]	Intra-oral	Pressure Sensor	**Yes**	No	No	-	No	<3mm
Kuo et al. [28]	Intra-oral	IMU, Force Sensor	**Yes**	No	No	No	No	>4mm
s-Guard [34]	Intra-oral	IMU	**Yes**	No	No	No	No	-
iBrux [57]	Intra-oral	Pressure Sensor	**Yes**	No	**Yes**	No	No	-
TeethVib [59]	Intra-oral	Piezoelectric Sensor	No	No	**Yes**	**Yes**	**Yes**	<3mm
MuteIt [83]	Extra-oral	IMU	**Yes**	No	**Yes**	**Yes**	No	-
TeethTap [85]	Extra-oral	IMU, Microphone	**Yes**	No	**Yes**	**Yes**	No	-
Wearable Lab [63]	Extra-oral	IMU, Biosensor (e.g., saliva)	**Yes**	No	**Yes**	**Yes**	No	-
IOTeeth (Ours)	Intra-oral	Piezoelectric Sensor	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	<3mm

more affordable than systems built on a customized mouthguard, which typically cost over 300 US dollars[4]. This affordability enhances *IOTeeth*'s accessibility.

8.2 Piezoelectric Sensing

There has been active research on wearable sensors and devices leveraging piezoelectric sensing. Wang et al. developed a piezoelectric MEMS sensor for unvoiced speech recognition based on oral airflow, which recognizes words based on their unique voltage-signal characteristics and is less affected by external noise [91]. Park et al. introduce an earable device using piezoelectric sensors to capture the in-ear pulse waves on the surface of the ear canal for heart rate monitoring [61]. Yi et al. present a continuous blood pressure monitoring wearable that leverages a piezoelectric sensor to measure arterial pulse waves on the elbow [98]. Roopa Manjunatha et al. integrate piezoelectric sensors into the eyeglass frame [73]. Sensors are hit by airflow from each nostril, then the system derives the user's respiration pattern from signals. Lanata et al. et al. propose a chest-mounted wearable system for cardiopulmonary activity monitoring [29]. By placing the piezoelectric sensor near the heart and lung, the system can measure both heart apex movements and the respiration pattern. These studies illustrate the versatility and potential of piezoelectric sensing in health monitoring applications. Different from these systems, our work is the first effort to leverage piezoelectric sensors for continuous occlusal disease recognition.

8.3 Multi-sensor Learning for Human Sensing

Multiple studies have been published in the area of multi-sensor learning for human sensing, along with the progress in deep learning algorithms. Learning from multiple sensors provides a holistic view of the sensing target hence increasing the model's inference accuracy. One of the most explored areas is human activity recognition. Yao et al. propose the DeepSense model that concatenates latent features from different sensors together to fuse their information [96]. Beyond concatenation, there are studies that model sensor interactions and fuse the information via shared neural network layers [18, 37, 40, 68]. While these multi-sensor learning approaches show promising results in combining individual sensor data to form global information, the weight of individual sensors is predefined or static [39]. The static weight limits the learning model's generalizability and adaptability to variance in the real-world sensing data. Hence there is a challenging requirement for multi-sensor learning: dynamically assess and adapt to which inputs or features are most relevant or important in any given context,

---

[4]The estimated cost of over 300 US dollars is derived from market research on mouthguard customization and additional expenses for accessory materials such as piezoelectric sensors, MCU and ADC.

and the attention mechanism is actively studied. These works propose to model attention weights as learnable parameters for each sensor input, thus the model can dynamically adjust its emphasis by weighting the features of each sensor accordingly [39, 44, 90, 97, 99–101]. Even though they show promising results in human activity recognition, it is not a plug-and-play transfer to our occlusal disease recognition task. Unlike human activity recognition, signals captured by the in-mouth vibration sensor array contain misleading information (e.g., an occlusal disease patient may not have all dental areas showing symptoms.) Therefore, *IOTeeth*'s PAN Net is built upon previous attention studies and extends its capability to signify the most indicating sensor.

9 CONCLUSION

In this paper, we presented *IOTeeth*, an intra-oral sensing system with a dental retainer-form design to continuously monitor teeth activities and recognize occlusal diseases. *IOTeeth* consists of an intra-oral sensing wearable that leverages customized mini-sized PVDF piezoelectric sensors embedded between dental retainers to sense teeth activities-induced vibration. To ensure the sensitivity and consistency of customized sensors, we design and implement an automated testing bench for validating and characterizing the sensor. The vibration signals are then further processed for teeth activity recognition. The signals of different activities are sent to the corresponding model for occlusal disease recognition. We propose a Physio-aware Network (PAN Net) to automatically extract the most informative patterns from the signal for generalizable occlusal disease recognition. We evaluate *IOTeeth* with 12 real-world patients' articulator-based teeth models from a dental clinic. The *IOTeeth* achieves a leave-one-out testing F1 score of 0.97 for activity recognition and an average leave-one-out testing F1 score of 0.92 for dental disease recognition for different activities.

ACKNOWLEDGMENTS

REFERENCES

[1] ANSI/IEEE Std 176-1987. 1987. IEEE standard on piezoelectricity. The Institute of Electrical and Electronics Engineers New York.

[2] Dr Olurotimi A. Adeleye, Muritala Yusuf, and Oluwaseyi . Balogun. 2020. Dynamic Analysis of Viscoelastic Circular Diaphragm of a MEMS Capacitive Pressure Sensor using Modified Differential Transformation Method. *Karbala International Journal of Modern Science* 6, 3 (Oct. 2020). https://doi.org/10.33640/2405-609X.1706

[3] Sinan Akdemir, Selen Ergin Tokgöz, Esra Kurt, and Hakan Bilhan. 2022. Determination of intercanine distance from a proportion calculation: A pilot clinical study. *The Journal of Prosthetic Dentistry* (2022).

[4] Feras AlMofreh AlQahtani, Sudhir Rama Varma, Sam Thomas Kuriadom, Basma AlMaghlouth, and Nasser AlAsseri. 2023. Changes in occlusion after orthognathic surgery: a systematic review and meta-analysis. *Oral and Maxillofacial Surgery* (2023), 1–12.

[5] Antonio Arnau et al. 2004. *Piezoelectric transducers and applications*. Vol. 2004. Springer.

[6] Merete Bakke. 2006. Bite force and occlusion. In *Seminars in orthodontics*, Vol. 12. Elsevier, 120–126.

[7] Monish Bhola, Leyvee Cabanilla, and Shilpa Kolhatkar. 2008. Dental occlusion and periodontal disease: what is the real relationship? *Journal of the California Dental Association* 36, 12 (2008), 925–930.

[8] Francesco Carini, Margherita Mazzola, Chiara Fici, Salvatore Palmeri, Massimo Messina, Provvidenza Damiani, and Giovanni Tomasello. 2017. Posture and posturology, anatomical and physiological profiles: overview and current state of art. *Acta Bio Medica: Atenei Parmensis* 88, 1 (2017), 11.

[9] Peter E Dawson. 2006. *Functional occlusion: from TMJ to smile design*. Elsevier Health Sciences.

[10] JA De Boever, GE Carlsson, and IJ Klineberg. 2000. Need for occlusal therapy and prosthodontic treatment in the management of temporomandibular disorders. Part II: Tooth loss and prosthodontic treatment. *Journal of oral rehabilitation* 27, 8 (2000), 647–659.

[11] Alana de Santana Correia and Esther Luna Colombini. 2022. Attention, please! A survey of neural attention models in deep learning. *Artificial Intelligence Review* 55, 8 (2022), 6037–6124.

[12] Mustafa Demirci, Ferda Karabay, Meriç Berkman, İlknur Özcan, Safa Tuncer, Neslihan Tekçe, and Canan Baydemir. 2022. The prevalence, clinical features, and related factors of dentin hypersensitivity in the Turkish population. *Clinical Oral Investigations* 26, 3 (2022), 2719–2732.

[13] Rena N D'Souza, Francis S Collins, and Vivek H Murthy. 2022. Oral health for all—realizing the promise of science. *New England Journal of Medicine* 386, 9 (2022), 809–811.

[14] Jingyuan Fan and Jack G Caton. 2018. Occlusal trauma and excessive occlusal forces: Narrative review, case definitions, and diagnostic considerations. *Journal of periodontology* 89 (2018), S214–S222.

[15] M Farella, S Palla, S Erni, A Michelotti, and LM Gallo. 2008. Masticatory muscle activity during deliberately performed oral tasks. *Physiological Measurement* 29, 12 (2008), 1397.

[16] Robert Gauer and Michael J Semidey. 2015. Diagnosis and treatment of temporomandibular disorders. *American family physician* 91, 6 (2015), 378–386.

[17] Irving Glickman. 1963. Inflammation and trauma from occlusion, co-destructive factors in chronic periodontal disease. *The Journal of Periodontology* 34, 1 (1963), 5–10.

[18] Zhizhang Hu, Yue Zhang, Tong Yu, and Shijia Pan. 2022. VMA: Domain Variance-and Modality-Aware Model Transfer for Fine-Grained Occupant Activity Recognition. In *2022 21st ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*. IEEE, 259–270.

[19] Rob J Hyndman and George Athanasopoulos. 2018. *Forecasting: principles and practice*. OTexts.

[20] Kenta Ichikawa and Wataru Hijikata. 2022. Energy harvesting from biting force with thin sheet harvester based on electret and dielectric elastomer. *Nano Energy* 99 (2022), 107357.

[21] Sergey Ioffe and Christian Szegedy. 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning*. pmlr, 448–456.

[22] OP Ivanyuta, G Al-Harasi, and DA Kuleshov. 2020. Modification of the dental arch shape using graphic reproduction method and its clinical effectiveness in patients with occlusion anomalies. *Archiv EuroMedica* 10, 4 (2020), 181.

[23] Vimal Kakaraparthi, Qijia Shao, Charles J Carver, Tien Pham, Nam Bui, Phuc Nguyen, Xia Zhou, and Tam Vu. 2021. FaceSense: sensing face touch with an ear-worn system. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 3 (2021), 1–27.

[24] NJ Kassebaum, E Bernabé, M Dahiya, Bishal Bhandari, CJL Murray, and W Marcenes. 2014. Global burden of severe tooth loss: a systematic review and meta-analysis. *Journal of dental research* 93, 7_suppl (2014), 20S–28S.

[25] Mohd Toseef Khan, Sanjeev Kumar Verma, Sandhya Maheshwari, Syed Naved Zahid, and Prabhat K Chaudhary. 2013. Neuromuscular dentistry: Occlusal diseases and posture. *Journal of oral biology and craniofacial research* 3, 3 (2013), 146–150.

[26] Rio Kinjo, Takahiro Wada, Hiroshi Churei, Takehiro Ohmi, Kairi Hayashi, Kazuyoshi Yagishita, Motohiro Uo, and Toshiaki Ueno. 2021. Development of a Wearable Mouth Guard Device for Monitoring Teeth Clenching during Exercise. *Sensors* 21, 4 (2021), 1503.

[27] Karthik Krishnamurthy, Frederic Lalande, and Craig A Rogers. 1996. Effects of temperature on the electrical impedance of piezoelectric sensors. In *Smart structures and materials 1996: smart structures and integrated systems*, Vol. 2717. SPIE, 302–310.

[28] Calvin Kuo, Lyndia C Wu, Brad T Hammoor, Jason F Luck, Hattie C Cutcliffe, Robert C Lynall, Jason R Kait, Kody R Campbell, Jason P Mihalik, Cameron R Bass, et al. 2016. Effect of the mandible on mouthguard measurements of head kinematics. *Journal of biomechanics* 49, 9 (2016), 1845–1853.

[29] Antonio Lanata, Enzo Pasquale Scilingo, and Danilo De Rossi. 2009. A multimodal transducer for cardiopulmonary activity monitoring in emergency. *IEEE Transactions on Information Technology in Biomedicine* 14, 3 (2009), 817–825.

[30] Andres Diaz Lantada, Carlos González Bris, Pilar Lafont Morgado, and Jesús Sanz Maudes. 2012. Novel system for bite-force sensing and monitoring based on magnetic near field communication. *Sensors* 12, 9 (2012), 11544–11558.

[31] Andres Diaz Lantada, Carlos González Bris, Pilar Lafont Morgado, and Jesús Sanz Maudes. 2012. Novel System for Bite-Force Sensing and Monitoring Based on Magnetic Near Field Communication. *Sensors (Basel, Switzerland)* 12, 9 (Aug. 2012), 11544–11558. https://doi.org/10.3390/s120911544

[32] Dieterich Lawson, Chung-Cheng Chiu, George Tucker, Colin Raffel, Kevin Swersky, and Navdeep Jaitly. 2018. Learning hard alignments with variational inference. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 5799–5803.

[33] Jinyoung Lee and Hong-Goo Kang. 2023. Real-time neural speech enhancement based on temporal refinement network and channel-wise gating methods. *Digital Signal Processing* 133 (2023), 103879.

[34] Seo-Joon Lee, Il-Do Jeong, Eo-Bin Kim, Jin-Young Park, In-Hwan Jo, Jae-Hoon Han, and Tae-Young Jung. 2021. s-Guard: Multisensor Embedded Obstructive Sleep Apnea and Bruxism Real-Time Data Transmission Intraoral Appliance Device. *Applied Sciences* 11, 9 (2021), 4182.

[35] Vladimir Leon-Salazar, Leesa Morrow, and Eric L Schiffman. 2012. Pain and persistent occlusal awareness: what should dentists do? *The Journal of the American Dental Association* 143, 9 (2012), 989–991.

[36] Xilai Li, Wei Sun, and Tianfu Wu. 2020. Attentive normalization. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XVII 16*. Springer, 70–87.

[37] Zexin Li, Xiaoxi He, Yufei Li, Shahab Nikkhoo, Wei Yang, Lothar Thiele, and Cong Liu. 2023. Mimonet: Multi-input multi-output on-device deep learning. *arXiv preprint arXiv:2307.11962* (2023).

[38] Mei Lin, Susan O Griffin, Barbara F Gooch, Lorena Espinoza, Liang Wei, Chien-Hsun Li, Gina Thornton-Evans, Michele L Junger, Valerie A Robison, Eleanor B Fleming, et al. 2019. Oral health surveillance report: trends in dental caries and sealants, tooth retention, and edentulism, United States: 1999–2004 to 2011–2016. (2019).

[39] Shengzhong Liu, Shuochao Yao, Jinyang Li, Dongxin Liu, Tianshi Wang, Huajie Shao, and Tarek Abdelzaher. 2020. Giobalfusion: A global attentional deep learning framework for multisensor information fusion. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 1 (2020), 1–27.

[40] Zuozhu Liu, Wenyu Zhang, Tony QS Quek, and Shaowei Lin. 2017. Deep fusion of heterogeneous sensor data. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 5965–5969.

[41] F Lobbezoo and M Naeije. 2001. Bruxism is mainly regulated centrally, not peripherally. *Journal of oral rehabilitation* 28, 12 (2001), 1085–1091.

[42] Lijun Lu, Ning Zhao, Jingquan Liu, and Bin Yang. 2021. Coupling piezoelectric and piezoresistive effects in flexible pressure sensors for human motion detection from zero to high frequency. *Journal of Materials Chemistry C* 9, 29 (2021), 9309–9318.

[43] J D Lytle. 2001. Occlusal disease revisited: Part I–Function and parafunction. *The International journal of periodontics & restorative dentistry* 21 3 (2001), 264–71.

[44] Haojie Ma, Wenzhong Li, Xiao Zhang, Songcheng Gao, and Sanglu Lu. 2019. AttnSense: Multi-level attention mechanism for multimodal human activity recognition.. In *IJCAI*. 3109–3115.

[45] Cristiane R Macedo, Elizeu C Macedo, Maria R Torloni, Ademir B Silva, and Gilmar F Prado. 2014. Pharmacotherapy for sleep bruxism. *Cochrane database of systematic reviews* 10 (2014).

[46] Jianguo Miao, Congying Deng, Heng Zhang, and Qiang Miao. 2023. Interactive channel attention for rotating component fault detection with strong noise and limited data. *Applied Soft Computing* 138 (2023), 110171.

[47] Angela Militi, Federica Sicari, Marco Portelli, Emanuele Maria Merlo, Antonella Terranova, Fabio Frisone, Riccardo Nucera, Angela Alibrandi, and Salvatore Settineri. 2021. Psychological and social effects of oral health and dental aesthetic in adolescence and early adulthood: An observational study. *International Journal of Environmental Research and Public Health* 18, 17 (2021), 9022.

[48] Ali Mirfazaelian and Abbas Monzavi. 2003. Use of a feeler gauge to measure the gap between adjacent teeth. *Journal of Prosthetic Dentistry* 90, 6 (2003), 613.

[49] Mostafa Mirshekari, Jonathon Fagert, Shijia Pan, Pei Zhang, and Hae Young Noh. 2020. Step-level occupant detection across different structures through footstep-induced floor vibration using model transfer. *Journal of Engineering Mechanics* 146, 3 (2020), 04019137.

[50] Srigitha.S. Nath, D Lingaraja, P Aravind, P Thangavel, P Arun Kumar, and G.Dinesh Ram. 2023. Comparison of Different Shapes MEMS based Piezo Electric Pressure Sensor for Smart Devices. In *2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT)*. 1–5. https://doi.org/10.1109/ICSSIT55814.2023.10061138 ISSN: 2832-3017.

[51] Shamima Easmin Nishi, Rehana Basri, Mohammad Khursheed Alam, Shinichi Komatsu, Atsuo Komori, Yoshihiko Sugita, and Hatsuhiko Maeda. 2017. Evaluation of masticatory muscles function in different malocclusion cases using surface electromyography. *Journal of Hard Tissue Biology* 26, 1 (2017), 23–28.

[52] National Institute of Dental and Craniofacial Research (US). 2000. *Oral health in America: a report of the Surgeon General*. US Public Health Service, Department of Health and Human Services.

[53] Institute of Medicine. 1980. *Public Policy Options for Better Dental Health: Report of a Study*. The National Academies Press, Washington, DC. https://doi.org/10.17226/9921

[54] Francisco Javier Ordóñez and Daniel Roggen. 2016. Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition. *Sensors* 16, 1 (2016), 115.

[55] Xiaomin Ouyang, Xian Shuai, Jiayu Zhou, Ivy Wang Shi, Zhiyuan Xie, Guoliang Xing, and Jianwei Huang. 2022. Cosmo: contrastive fusion learning with small data for multimodal human activity recognition. In *Proceedings of the 28th Annual International Conference on Mobile Computing And Networking*. 324–337.

[56] Emer O'Hare, John A Cogan, Frank Dillon, Madeleine Lowery, and Eoin D O'Cearbhaill. 2021. An Intraoral Non-Occlusal MEMS Sensor for Bruxism Detection. *IEEE Sensors Journal* 22, 1 (2021), 153–161.

[57] Emer O'Hare, John A. Cogan, Frank Dillon, Madeleine Lowery, and Eoin D. O'Cearbhaill. 2022. An Intraoral Non-Occlusal MEMS Sensor for Bruxism Detection. *IEEE Sensors Journal* 22, 1 (Jan. 2022), 153–161. https://doi.org/10.1109/JSEN.2021.3128246 Conference Name: IEEE Sensors Journal.

[58] Shijia Pan, Dong Yoon Lee, Jun Ho Lee, and Phuc Nguyen. 2021. TeethVib: Monitoring Teeth Functional Occlusion Through Retainer Vibration Sensing. In *2021 IEEE/ACM Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*. IEEE, 92–96.

[59] Shijia Pan, Dong Yoon Lee, Jun Ho Lee, and Phuc Nguyen. 2021. TeethVib: Monitoring Teeth Functional Occlusion Through Retainer Vibration Sensing. In *2021 IEEE/ACM Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*. 92–96. https://doi.org/10.1109/CHASE52844.2021.00018

[60] Shijia Pan, Tong Yu, Mostafa Mirshekari, Jonathon Fagert, Amelie Bonde, Ole J Mengshoel, Hae Young Noh, and Pei Zhang. 2017. Footprintid: Indoor pedestrian identification through ambient structural vibration sensing. *Proceedings of the ACM on Interactive,*

Mobile, Wearable and Ubiquitous Technologies 1, 3 (2017), 1–31.

[61] Jang-Ho Park, Dae-Geun Jang, Jung Wook Park, and Se-Kyoung Youm. 2015. Wearable sensing of in-ear pressure for heart rate monitoring with a piezoelectric sensor. *Sensors* 15, 9 (2015), 23402–23417.

[62] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. 2019. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems* 32 (2019).

[63] Pat Pataranutaporn, Abhinandan Jain, Casey M Johnson, Pratik Shah, and Pattie Maes. 2019. Wearable lab on body: combining sensing of biochemical and digital markers in a wearable device. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 3327–3332.

[64] Marco A Peres, Lorna MD Macpherson, Robert J Weyant, Blánaid Daly, Renato Venturelli, Manu R Mathur, Stefan Listl, Roger Keller Celeste, Carol C Guarnizo-Herreño, Cristin Kearns, et al. 2019. Oral diseases: a global public health challenge. *The Lancet* 394, 10194 (2019), 249–260.

[65] picotech. [n. d.]. 16 bit high resolution oscilloscope | Overview. https://www.picotech.com/oscilloscope/4262/picoscope-4262-overview

[66] Donald R Poulton and Sanford A Aaronson. 1961. The relationship between occlusion and periodontal status. *American Journal of Orthodontics* 47, 9 (1961), 690–699.

[67] William R Proffit, Henry W Fields, Brent Larson, and David M Sarver. 2018. *Contemporary orthodontics-e-book*. Elsevier Health Sciences.

[68] Valentin Radu, Catherine Tong, Sourav Bhattacharya, Nicholas D Lane, Cecilia Mascolo, Mahesh K Marina, and Fahim Kawsar. 2018. Multimodal deep learning for activity and context recognition. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 1, 4 (2018), 1–27.

[69] Samrudhdhi B Rangrej and James J Clark. 2021. A probabilistic hard attention model for sequentially observed scenes. *arXiv preprint arXiv:2111.07534* (2021).

[70] S Varalakshmi Reddy, M Praveen Kumar, D Sravanthi, Abdul Habeeb Bin Mohsin, and V Anuhya. 2014. Bruxism: a literature review. *Journal of international oral health: JIOH* 6, 6 (2014), 105.

[71] Maria Giulia Rezende Pucciarelli, Guilherme Hideki de Lima Toyoshima, Thais Marchini Oliveira, Heitor Marques Honório, Chiarella Sforza, and Simone Soares. 2020. Assessment of dental arch stability after orthodontic treatment and oral rehabilitation in complete unilateral cleft lip and palate and non-clefts patients using 3D stereophotogrammetry. *BMC Oral Health* 20, 1 (2020), 1–8.

[72] Tyto Robotics. [n. d.]. Series 1585: Drone Thrust Stand. https://www.tytorobotics.com/pages/series-1580-1585

[73] G Roopa Manjunatha, K Rajanna, D Roy Mahapatra, MM Nayak, Uma Maheswari Krishnaswamy, and R Srinivasa. 2013. Polyvinylidene fluoride film based nasal sensor to monitor human respiration pattern: An initial clinical study. *Journal of clinical monitoring and computing* 27 (2013), 647–657.

[74] Ronald H Roth. 1976. The maintenance system and occlusal dynamics. *Dental Clinics of North America* 20, 4 (1976), 761–788.

[75] Jose-Luis Ruiz. 2009. Seven signs and symptoms of occlusal disease: the key to an easy diagnosis. *Dentistry Today* 28, 8 (2009), 112–113.

[76] Jose-Luis Ruiz and Thomas A Coleman. 2008. Occlusal disease management system: the diagnosis process. *Compendium of Continuing Education in Dentistry (Jamesburg, NJ: 1995)* 29, 3 (2008), 148–52.

[77] Felix Joyonto Saha, Almut Pulla, Thomas Ostermann, Theresa Miller, Gustav Dobos, and Holger Cramer. 2019. Effects of occlusal splint therapy in patients with migraine or tension-type headache and comorbid temporomandibular disorder: A randomized controlled trial. *Medicine* 98, 33 (2019).

[78] Mohammed Saleh, Mohammad Y Hajeer, and Dieter Muessig. 2017. Acceptability comparison between Hawley retainers and vacuum-formed retainers in orthodontic adult patients: a single-centre, randomized controlled trial. *European Journal of Orthodontics* 39, 4 (2017), 453–461.

[79] Max W Seitz, Stefan Listl, Andreas Bartols, Ingrid Schubert, Katja Blaschke, Christian Haux, and Marieke M van der Zande. 2019. Current knowledge on correlations between highly prevalent dental conditions and chronic diseases: an umbrella review [dataset]. (2019).

[80] Mayank Shrivastava, Ricardo Battaglino, and Liang Ye. 2021. A comprehensive review on biomarkers associated with painful temporomandibular disorders. *International journal of oral science* 13, 1 (2021), 23.

[81] Roger A Solow. 2018. Clinical protocol for occlusal adjustment: Rationale and application. *CRANIO®* 36, 3 (2018), 195–206.

[82] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. 2014. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research* 15, 1 (2014), 1929–1958.

[83] Tanmay Srivastava, Prerna Khanna, Shijia Pan, Phuc Nguyen, and Shubham Jain. 2022. MuteIt: Jaw Motion Based Unvoiced Command Recognition Using Earable. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 3 (Sept. 2022), 1–26. https://doi.org/10.1145/3550281

[84] Saša Stanković, Slobodan Vlajković, Mirjana Bošković, Goran Radenković, Vladimir Antić, and Danimir Jevremović. 2013. Morphological and biomechanical features of the temporomandibular joint disc: an overview of recent findings. *Archives of oral biology* 58, 10 (2013), 1475–1482.

[85] Wei Sun, Franklin Mingzhe Li, Benjamin Steeper, Songlin Xu, Feng Tian, and Cheng Zhang. 2021. Teethtap: Recognizing discrete teeth gestures using motion and acoustic sensing on an earpiece. In *26th International Conference on Intelligent User Interfaces*. 161–169.

[86] Koji Toma, Keisuke Tomoto, Kumi Yokota, Nao Yasuda, Tatsuya Ishikawa, Takahiro Arakawa, and Kohji Mitsubayashi. 2018. Mouthguard controller for unconstrained control of external devices. *Sensors Mater* 30 (2018), 3053–3060.

[87] H Van Beek. 1979. The transfer of mesial drift potential along the dental arch in Macaca irus: an experimental study of tooth migration rate related to the horizontal vectors of occlusal forces. *The European Journal of Orthodontics* 1, 2 (1979), 125–129.

[88] Erik Vanegas, Raúl Igual, and Inmaculada Plaza. 2021. The effect of measurement trends in belt breathing sensors. *Engineering Proceedings* 6, 1 (2021), 84.

[89] Francis F Vernallis. 1953. *Teeth-grinding: Some relationships to anxiety, hostility, and hyperactivity.* The Pennsylvania State University.

[90] Chongyang Wang, Min Peng, Temitayo A Olugbade, Nicholas D Lane, Amanda C De C Williams, and Nadia Bianchi-Berthouze. 2019. Learning temporal and bodily attention in protective movement behavior detection. In *2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*. IEEE, 324–330.

[91] Qi Wang, Tao Ruan, Qingda Xu, Yuzhi Shi, Bin Yang, and Jingquan Liu. 2022. Piezoelectric Mems Unvoiced Speech-Recognition Sensor Based on Oral Airflow. In *2022 IEEE 35th International Conference on Micro Electro Mechanical Systems Conference (MEMS)*. IEEE, 63–66.

[92] Franz Josef Willeit, Francesca Cremonini, Paul Willeit, Fabio Ramina, Marta Cappelletti, Giorgio Alfredo Spedicato, and Luca Lombardo. 2022. Stability of transverse dental arch dimension with passive self-ligating brackets: a 6-year follow-up study. *Progress in orthodontics* 23, 1 (2022), 1–8.

[93] Julian B Woelfel and Rickne C Scheid. 1997. *Dental anatomy*. Williams & wilkins.

[94] Lyndia C. Wu, Calvin Kuo, Jesus Loza, Mehmet Kurt, Kaveh Laksari, Livia Z. Yanez, Daniel Senif, Scott C. Anderson, Logan E. Miller, Jillian E. Urban, Joel D. Stitzel, and David B. Camarillo. 2017. Detection of American Football Head Impacts Using Biomechanical Features and Support Vector Machine Classification. *Scientific Reports* 8, 1 (Dec. 2017), 855. https://doi.org/10.1038/s41598-017-17864-3 Number: 1 Publisher: Nature Publishing Group.

[95] Jianbo Yang, Minh Nhut Nguyen, Phyo Phyo San, Xiaoli Li, and Shonali Krishnaswamy. 2015. Deep convolutional neural networks on multichannel time series for human activity recognition.. In *Ijcai*, Vol. 15. Buenos Aires, Argentina, 3995–4001.

[96] Shuochao Yao, Shaohan Hu, Yiran Zhao, Aston Zhang, and Tarek Abdelzaher. 2017. Deepsense: A unified deep learning framework for time-series mobile sensing data processing. In *Proceedings of the 26th international conference on world wide web*. 351–360.

[97] Shuochao Yao, Yiran Zhao, Huajie Shao, Dongxin Liu, Shengzhong Liu, Yifan Hao, Ailing Piao, Shaohan Hu, Su Lu, and Tarek F Abdelzaher. 2019. Sadeepsense: Self-attention deep learning framework for heterogeneous on-device sensors in internet of things applications. In *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*. IEEE, 1243–1251.

[98] Zhiran Yi, Zhaoxu Liu, Wenbo Li, Tao Ruan, Xiang Chen, Jingquan Liu, Bin Yang, and Wenming Zhang. 2022. Piezoelectric dynamics of arterial pulse for wearable continuous blood pressure monitoring. *Advanced Materials* 34, 16 (2022), 2110291.

[99] Ming Zeng, Haoxiang Gao, Tong Yu, Ole J Mengshoel, Helge Langseth, Ian Lane, and Xiaobing Liu. 2018. Understanding and improving recurrent networks for human activity recognition by continuous attention. In *Proceedings of the 2018 ACM international symposium on wearable computers*. 56–63.

[100] Yue Zhang, Zhizhang Hu, Uri Berger, and Shijia Pan. 2023. CMA: Cross-Modal Association Between Wearable and Structural Vibration Signal Segments for Indoor Occupant Sensing. In *Proceedings of the 22nd International Conference on Information Processing in Sensor Networks*. 96–109.

[101] Yue Zhang, Zhizhang Hu, Uri Berger, and Shijia Pan. 2023. Integrating On-and Off-body Sensing for Young Adults Failure to Launch (FTL) Behavior Profiling. In *Proceedings of the 22nd International Conference on Information Processing in Sensor Networks*. 320–321.