Using Knee Acoustical Emissions for Sensing Joint Health in Patients with Juvenile Idiopathic Arthritis: A Pilot Study

Beren Semiz*, Sinan Hersek*, Daniel C. Whittingslow*, Lori Ponder, Sampath Prahalad, and Omer T. Inan, *Senior Member, IEEE.*

***Abstract*—In this paper, we present a pilot study evaluating novel methods for assessing joint health in patients with Juvenile Idiopathic Arthritis (JIA) using wearable acoustical emission measurements from the knees. Measurements were taken from four control subjects with no known knee injuries, and from four subjects with JIA, before and after treatment. Time and frequency domain features were extracted from the acoustical emission signals and used to compute a knee audio score. The score was used to separate out the two groups of subjects based solely on the sounds their joints produce. It was created using a soft classifier based on gradient boosting trees. The knee audio scores ranged from 0-1 with 0 being a healthy knee and 1 being an involved joint with arthritis. Leave-one-subject-out cross-validation (LOSO-CV) was used to validate the algorithm. The average of the right and left knee audio scores was 0.085±0.099 and 0.89±0.012 for the control group and group with JIA, respectively ($p<0.05$). The average knee audio score for the subjects with JIA decreased from 0.89±0.012 to 0.25±0.20 following successful treatment ($p<0.05$). The knee audio score metric successfully distinguished between the control subjects and subjects with JIA. The scores calculated before and after treatment accurately reflected the observed clinical course of the subjects with JIA. After successful treatment, the subjects with JIA were classified as healthy by the algorithm. Knee acoustical emissions provide a novel and cost-effective method for monitoring JIA, and can be used as an objective guide for assessing treatment efficacy.**

***Keywords*—Wearable technologies, knee joint health, acoustics.**

*Indicates that the authors contributed equally to the manuscript and should be considered first author.

This material is based upon work supported in part by the Pediatric Technology Center (Georgia Tech and Children's Healthcare of Atlanta), the Marcus Foundation, Inc., Atlanta, GA, the National Institutes of Health, National Institute of Biomedical Imaging and Bioengineering, Grant No. 1R01EB023808, as part of the NSF/NIH Smart and Connected Health Program, and the National Science Foundation CAREER Award, Grant No. 1749677.

B.Semiz, S.Hersek, and O. T. Inan are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, 30332 USA. Corresponding Author e-mail: inan@gatech.edu.

D.C. Whittingslow is with the Emory University School of Medicine and Georgia Institute of Technology Coulter Department of Biomedical Engineering under MD/PhD program.

L. Ponder is with the Department of Pediatrics, Emory University School of Medicine, and Children's Healthcare of Atlanta, Atlanta, GA.

S. Prahalad is with the Departments of Pediatrics and Human Genetics, Emory University School of Medicine, and Children's Healthcare of Atlanta, Atlanta, GA.

I. Introduction

JUVENILE idiopathic arthritis (JIA) describes a clinically heterogeneous group of arthritides. It is the most common rheumatic condition in children and one of the more common chronic illnesses of childhood affecting more than 50,000 children in the United States [1],[2]. The cause and pathogenesis of JIA are still poorly understood, but associations with various genetic and environmental factors have been made. There are seven defined categories of JIA, each with their own distinct presentation, clinical signs, symptoms, and clinical course. Identifying these categories coupled with recent studies into the genetic contributors of JIA have allowed for more precise treatment protocols; however, there is still significant variability within each of these categories. Each patient's clinical course is unique. Often, patients with JIA experience cyclical periods of active disease and remission. These unpredictable flare-ups, coupled with the highly variable causes and presentations have made predicting the long-term prognosis of JIA difficult. This difficulty in predicting prognosis exacerbates the already difficult process of selecting an ideal treatment regimen [3],[4].

The pathophysiologic changes to the joints caused by JIA can lead to progressive joint destruction. The long-term sequela of JIA is severe and includes chronic pain, joint immobility, unstable articulation, and even disability. Fortunately, if JIA is detected and treated properly early in its presentation, the long-term consequences can be largely prevented [5],[6]. One of the joints most commonly affected by JIA is the knee [6]. Unfortunately, there are very few quantitative means for readily assessing the disease status in affected knees.

There are both invasive and non-invasive procedures for knee-health evaluation. Various imaging technologies such as computed tomography (CT), musculoskeletal ultrasound, magnetic resonance imaging (MRI), and fMRI are non-invasive evaluation methods, however they fail to provide early diagnosis, are prohibitively expensive for continuous monitoring, and are inconvenient to perform [7]. One of the more common invasive procedures for assessing knee health is arthroscopy. Arthroscopy provides detailed information, but it is a small surgical procedure, cannot be used on highly degenerated knees and is both cost and time intensive [8],[9]. In a chronic condition as variable as JIA, there is a compelling need for a quantitative, unobtrusive, and affordable method for assessing joint health.

The use of acoustics—recording the sounds that the joints make during movement—may provide such a method [10]. In the case of the knee, vibrations are emitted from the mid-patellar region during active movements such as flexion and extension [11]. These vibrations can be measured on the surface of the skin as sound. These so called vibroarthographic signals have been proposed as a possible diagnostic tool for early diagnosis of joint disorders [12]. Since Blodgett pioneered the technique in 1902, a number of advances have been made in the field of vibroarthography [13]. Vibration arthrometry was used to show substantial differences in injured and healthy joints [14]. Later, the power profile was found to be significantly different between joints that were healthy, had rheumatoid arthritis and degenerative arthritis. It was also discovered that the majority of the signals occurred in the range of 20-20,000 Hz [15]. The field of vibroarthography made a large leap forward with the development and application of piezoelectric accelerometers. This type of sensor is sensitive to physical vibrations (such as those seen on the skin during joint movement) and has less chance of registering external noises as compared to electret or other microphones sensing airborne pressure waves [16],[17].

Accurately recording the joint's acoustical emissions is only part of what is required for the development of a successful joint health monitoring suite. Those signals must also be analyzed and given physiologic context, such that this technique can be applied on patients. To analyze the recorded signals, different signal processing and machine learning techniques have been used. The techniques used thus far include wavelet decomposition [18], time-frequency analysis [19], Fourier analysis [20], autoregressive modelling [21], statistical parameter investigation and neural networks [22]. There is a significant need for, and lack of understanding about, which signal analysis technique is ideal for a given acoustical emission signal, and more particularly for a given disorder, such as JIA.

In order to contribute to the field of acoustical emission signal analysis and non-invasive joint health monitoring, our team has developed a novel hardware setup and signal analysis algorithm. In our hardware setup, a small piezoelectric accelerometer is attached to the medial side of the patella to acquire the acoustical emissions from the knee joint. The recorded signals are then analyzed using our novel algorithm that computes a knee audio score, which places the recorded joint along a gradient from healthy to an involved joint with arthritis. In this manuscript, we detail both our hardware setup for recording knee acoustical emissions and our developed machine learning algorithm for classifying the knees. We validate our methods via a human subject study involving four healthy control subjects and four subjects with JIA. Finally, to investigate the ability of this technique for quantifying disease progression and treatment response—essential components for the monitoring of JIA—we present data acquired from the subjects with JIA before and after treatment.

II. Methods

A. Human Subject Protocol and Subject Demographics

The study was conducted under a protocol approved by the Georgia Institute of Technology and Emory University Institutional Review Boards. Ten subjects participated in this study. However, due to a microphone failure, two control subjects had to be excluded from the data analysis and thus data is presented from only eight subjects. Although the dataset is small, the number of instances was increased using appropriate window and step sizes, which made it possible to train different machine learning algorithms as previously done in [23], [24],[25]. Four of the subjects were diagnosed with JIA by a pediatric rheumatologist and four of the subjects were healthy controls with no history of JIA or acute knee injuries. The group with JIA consisted of three females (height=157.1$\pm$8.8 cm, weight=48.9$\pm$12.3 kg, and age=14.7$\pm$2.1 years old), and one male (height=175.7 cm, weight=65.3 kg, age=17 years old). The healthy control group consisted of five females (height=141.7$\pm$10 cm, weight=34.1$\pm$3.6 kg, and age=9.6$\pm$1.8 years old) and one male (height=167.6 cm weight=54.5 kg, age=11 years old). In order to monitor the changes in knee acoustical emissions during the course of treatment, data were acquired from the subjects with JIA a second time, 3-6 months after initial measurements.

The data acquisition set up for each subject is shown in Fig.1(a). To record the sounds produced by the joints, a uniaxial analog accelerometer (3225F7, Dytran Instruments Inc. Chatsworth, CA) was attached 2 cm medial to the patellar tendon using KinesioTex tape (Kinesio Tex Gold, Kinesio, Albuquerque, NM). This accelerometer has a broad bandwidth (2Hz-10kHz), high sensitivity (100 mV/g), low noise floor (700 μg_{rms}), miniature size and low weight (1 gram). To ensure strong contact between the accelerometer and the subject's knee, the accelerometer was additionally wrapped in MEDca adhesive tape. The medial patellar location was selected due to the relatively unimpeded route (only a thin layer of muscle, tendon, and fat) to the articulating surface of the knee (where inter-joint friction is theorized to produce the recorded vibrations)[26],[27].

To record the knee acoustical emissions, the subjects performed ten unloaded knee flexion/extension exercises, while seated on a height-adjustable stool to prevent foot contact with the ground. The signals from the accelerometer were sampled at 108 kHz and recorded using a data acquisition module (USB-4432, National Instruments Corporation, Austin, TX). The exercise and recording protocol was repeated for both knees for all subjects. The recorded signals were analyzed using Matlab (MathWorks, Natick, MA) and Python (Python Software Foundation, Beaverton, OR).

B. Signal Processing and Feature Extraction

The joint acoustical emissions were analyzed in the time and frequency domains. Fig.1(a) shows a representative plot of the time domain signal from one subject with JIA. The acoustical emissions from knee joints have high energy and a short duration (between 10-20 ms). One unique characteristic of these signals is the audible "clicks" that have a spike-like

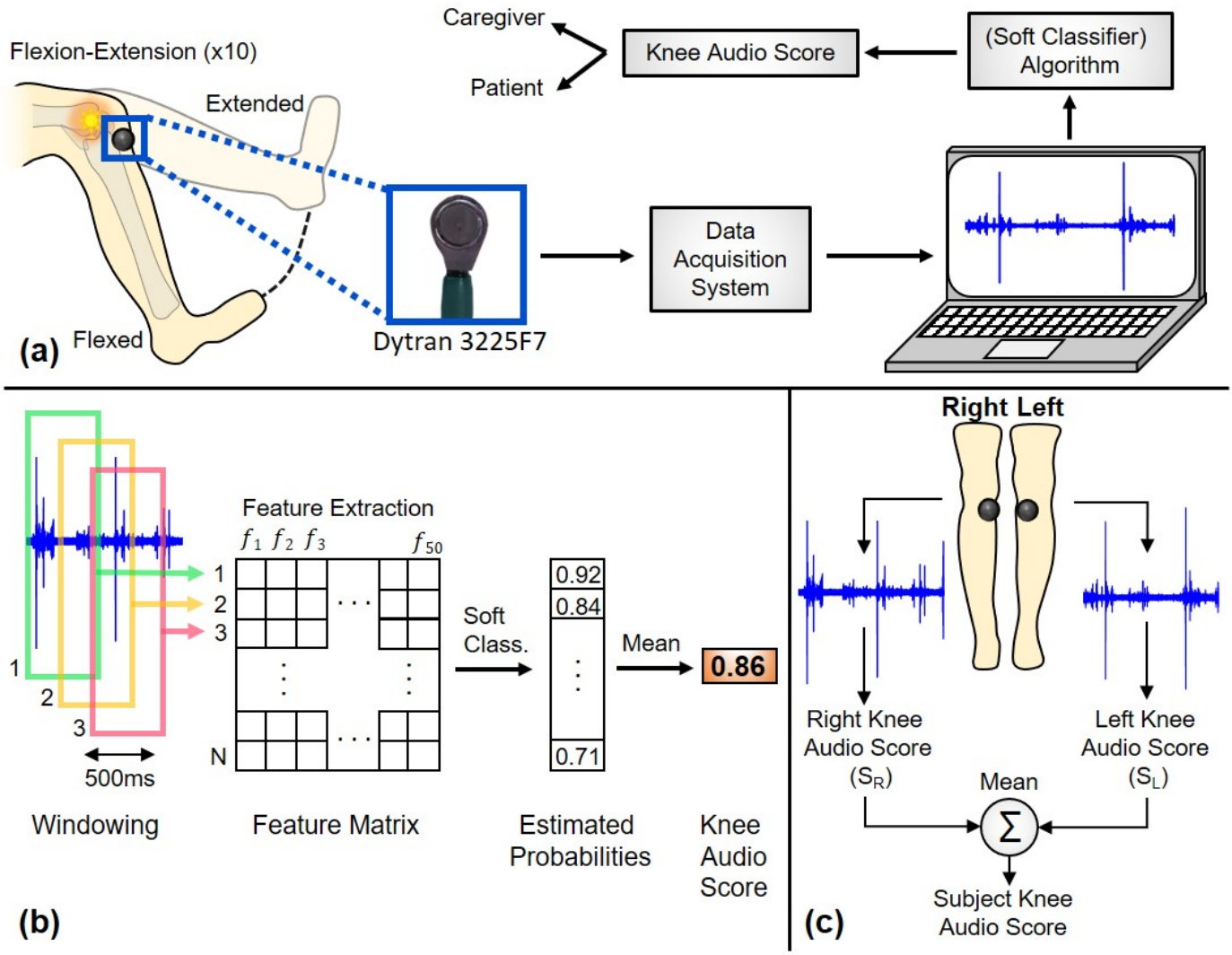

Fig. 1. An overview of the methods used to acquire acoustical emissions from the knee and the algorithms used to analyze the signals. (a) The hardware and sensor placement setup used to acquire knee acoustical emissions. An accelerometer is affixed at the medial aspect of the patella using athletic tape. The signals from the accelerometer are sampled and recorded using a data acquisition module. Signal processing and data mining algorithms are used to extract a knee audio score from the acoustical emissions. This score can be used to monitor knee arthritis continuously during the course of treatment in order to titrate care. (b) The signal analysis workflow for the knee joint acoustical emissions. The signals were filtered and separated into frames (frame length of 500ms and 50% overlap). The probability that a given frame belongs to an arthritic knee was estimated. This was done by extracting 50 audio features from the frame and feeding these to a pre-trained soft classifier. Finally, the knee audio score was calculated by averaging the estimated probabilities for all frames. (c) The knee audio score was calculated for both knees of each subject (S_R and S_L). The right and left knee scores were averaged to give the subject knee audio score.

appearance in the time domain plot. Additionally, these signals have high bandwidth frequency content reaching up to 20 kHz, which is fully expected for acoustical emissions [10],[15],[28].

Fig.1(b) details the signal analysis workflow for knee acoustical emissions. The signals are pre-processed using a digital finite impulse response (FIR) band-pass filter with 250Hz-20kHz bandwidth. Artifacts related to sensor placement and repositioning often observed at the the beginning and end of the recordings are removed manually. These signals are then separated into 500ms-long frames with 50% overlap, resulting in N=102 frames. The frame length of 500ms ensures that a single frame comprises an adequate sample of the variable signal, including both a combination of silent segments of the signal and those where clicks are observed. Additionally, for this dataset, using 500ms-long frames allows for multiple joint sound signatures to be present within a given frame [28]. The 50% overlap is selected to increase the number of instances used in processing. An overlap greater than 50% could decrease the processing speed. Fifty signal features are extracted from each frame and stored in the matrices $\boldsymbol{X}_R$ and $\boldsymbol{X}_L$ for the right and left knees, respectively. The rows of these matrices represent a single signal frame, and the columns represent the 50 features extracted (see Feature Matrix in Fig.1(b)).

The features extracted are categorized as either “time domain”, “spectral”, “MFCC” or “bandpower” features. The time

domain features are the zero crossing rate (ZCR), energy, and energy entropy (f_1-f_3). The most important characteristic of these signals is the audible "clicks" that have a spike-like appearance in the time domain plot, and this unique, consistent 'time domain' pattern of the signals result in distinctive time domain features. On the other hand, the high bandwidth frequency characteristics of the joint sounds result in them having distinctive "spectral" features, which are the spectral centroid, spectral spread, spectral flux, spectral entropy and spectral roll-off (f_4-f_8). The "MFCC" features are composed of the 13 mel-frequency cepstrum coefficients (f_9-f_{21}) which have the ability to separate joint sound signatures from background noise as previously shown in [28]. The "bandpower" features consist of the signal power in 29 distinct frequency bands, between 30 logarithmically spaced frequencies in the range of 250Hz-20kHz (f_{22}-f_{50}). These extracted features are detailed in [28],[29].

C. t-Distributed Stochastic Neighbor Embedding (t-SNE)

The features extracted from the 500ms signal frames from all subjects are analyzed using machine learning techniques. For visualizing the ability of our feature set to distinguish between the control group and group with JIA, dimensionality reduction is performed using t-Distributed Stochastic Neighbor Embedding (t-SNE) [30]. Each feature of our data set represents one dimension. For this analysis, we reduce the dimensionality of our data from fifty to two dimensions for ease of visualization. We then construct a scatter plot of the data with the two axes representing the two t-SNE dimensions and each point representing one 500ms signal frame. Two colors are used to categorically label the data points—JIA in red and healthy controls in blue. If a particular feature set has the ability to distinguish between control subjects and subjects with JIA, the groups would form two separate clusters in the scatter plot.

The dimensionality reduction technique that t-SNE employs attempts to maintain the distances of points based on their probabilities of being neighboring data points. Assuming that x_i is a point in the high dimensional space, x_i chooses x_j as its neighbor with the conditional probability $p_{i|j}$; likewise, in the low dimensional space, this probability can be represented as $q_{i|j}$. The t-SNE method aims to find the best low-dimensional data representation for minimizing the mismatch between the probabilities that two points are neighbors in high dimensional space ($p_{i|j}$) and low dimensional space ($q_{i|j}$) [30]. These probabilities represent a similarity metric between the two points. The t-SNE method is our preferred dimensionality reduction technique because, as shown in the literature, it minimizes local distortions and preserves the details within the data structure better that competing techniques, such as principal components analysis (PCA) or isometric feature mapping (ISOMAP) [31].

In order to visualize the dataset, we first concatenate the matrices $\boldsymbol{X}_R$ and $\boldsymbol{X}_L$ from the right and left knees of all subjects to form a matrix $\boldsymbol{X}_{\text{all}}$. As this matrix has features with different physical units, we standardize all of the columns to zero mean and unity variance. This allows all features to be weighted equally during dimensionality reduction. We store the labels corresponding to the rows of $\boldsymbol{X}_{\text{all}}$ in a vector $\boldsymbol{y}$. A given entry of $\boldsymbol{y}$ is labeled 0 if the corresponding frame belongs to a healthy subject and 1 if it belongs to a subject with JIA. Using all features, a scatter plot is plotted. The points of the scatter plot are colored according to the vector $\boldsymbol{y}$ to investigate if there is separation between the two different groups.

D. Knee Audio Score Calculation

A knee audio score is calculated to place the knees on a gradient ranging from 0 to 1 where a score of 0 represents a healthy, unaffected knee and 1 represents an involved knee with arthritis. To calculate the knee audio score from the features extracted from the acoustical emissions, we train a classification model. Signals acquired from the subjects in the training set are separated into frames and fifty features are extracted from each frame as explained previously. These features are stored in the matrix $\boldsymbol{Z}_{\text{train}}$ and their corresponding labels are stored in the vector $\boldsymbol{w}_{\text{train}}$. $\boldsymbol{Z}_{\text{train}}$ and $\boldsymbol{w}_{\text{train}}$ can then be used to train the classification model with the relationship between $\boldsymbol{Z}_{\text{train}}$ and $\boldsymbol{w}_{\text{train}}$. After training, the model can be used to predict the label (healthy or JIA) of an incoming frame. We selected a classification method that supports "soft classification," or the ability to estimate the *probability* that a given frame belongs to a subject with JIA [32]. Those probabilities are used to calculate the knee audio score from a recording. The knee audio score is the mean of the probabilities across all the frames from a given subject's recording. We expect this score to be higher for subjects with JIA compared to control subjects. We also hypothesize that this score will decrease with treatment.

The classification model is a necessary component of our knee sound analysis algorithm. The relationships observed in this type of data are non-linear and high-dimensional making more traditional classification techniques (which often apply linear methods) ineffective. We use the Extreme Gradient Boosting (XGBoost) classification which is a relatively new machine learning algorithm [33]. XGBoost is an implementation of the gradient boosting machine learning algorithm [34] which falls under a category of learning algorithms called ensemble methods. In this type of algorithm, to predict a variable, multiple estimators are simultaneously used rather than the typical use of a single estimator [35]. In the XGBoost algorithm, many decision trees are iteratively trained. This iterative training allows the model to predict the residual errors from the previous iteration and improve over time.

The assessment of the performance of the algorithm for classifying healthy vs. involved joints with arthritis is performed using leave-one-subject-out cross-validation (LOSO-CV) [36]. In each fold of the cross-validation, one subject is omitted and an XGBoost classifier is trained using the data from both knees of the remaining seven subjects. The trained model is then used to classify the signal frames of the excluded subject's knee acoustical emissions. The classified frames are stored in the vectors $\boldsymbol{y}_{R,i}$ and $\boldsymbol{y}_{L,i}$, for the right and left knees of the ith subject. The same classification model is used to estimate the probability of JIA for each frame. These probabilities are

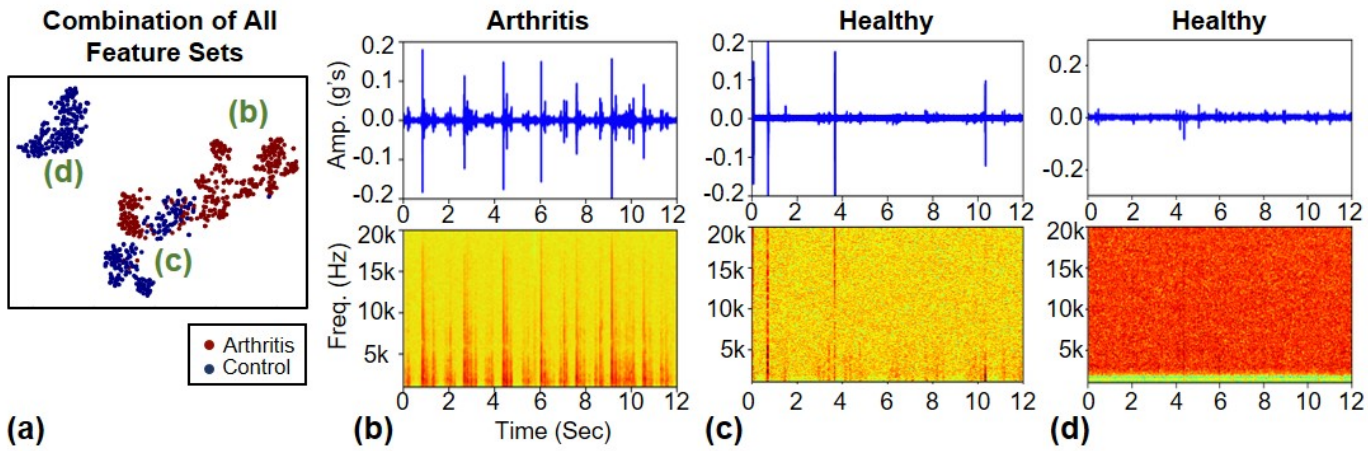

Fig. 2. (a) The t-SNE visualizations of knee acoustical emission data from four control subjects and four subjects with JIA using all features. We plotted representative signals and spectrograms from one knee of a subject with JIA (b), and two control subjects (c) and (d). These signals belonged to different clusters on the t-SNE plots, which are labeled on (a). (b) and (d) fall into clearly separated clusters in the t-SNE plot (a) as (b) has consistent "clicking" sounds throughout, while (d) is mostly silent. (c) is silent with some clicking sounds which places these signals to a cluster close to (b) on the t-SNE plot.

stored in the vectors $\boldsymbol{p}_{R,i}$ and $\boldsymbol{p}_{L,i}$ for the right and left knees of the ith subject.

The audio scores for the subject's knees are calculated by averaging the contents of the $\boldsymbol{p}_{R,i}$ and $\boldsymbol{p}_{L,i}$ for the right and left knees separately. An average knee audio score is also calculated for each subject by averaging the audio scores of the right and left knees (Fig.1(c)). A given subject is classified into the JIA group if the average knee audio score is greater than 0.5. For subjects with JIA, this process is repeated to calculate the knee audio scores for the post-treatment recordings. The cross-validation is completed by calculating knee audio scores for all eight subjects, excluding one subject per fold. Note that the post-treatment data for the JIA group is not used for training, as the ground truth labels for this data is not known certainly. The generalizability of our model is assessed by calculating the accuracy of our algorithm in labelling each frame. We also calculate the accuracy of our algorithm in predicting whether a subject belongs to the JIA or control groups.

E. Feature Importance Ranking

As was shown in Section II-D the knee audio score was calculated using the XGBoost classifier using fifty audio features. The relative weighting of each of those features in the model needs to be calculated to better understand which features are most relevant for the classification. All gradient boosting trees, including the XGBoost classifier train decision trees which can be used to rank the features according to their relative importance in the generated classification algorithm. Typically, the nodes of a tree divide using less important features while the initial node divides on the most important feature. The importance of features obtained from all the trees in the model are averaged resulting in the final relative feature importance scores [37]. These scores are applied to this particular data set helps to discern which features are most important for properly classifying joints with JIA.

We apply this XGBoost classifier to evaluate which audio features were most relevant for distinguishing between the control subjects and subjects with JIA. The data from every subject with JIA (excluding the follow-up data due to it lacking a ground truth classification) is used to train the classifier and the resulting model is used to generate relative feature importance scores. No testing set is required to quantify feature importance as we are not evaluating how well our model generalizes. After ranking the fifty features, we analyze how the top six features, which contributed the majority of the classification strength, differed between the control subjects and subjects with JIA as well as the manner in which they changed post-treatment.

F. Effect of Model Type

In order to evaluate the sensitivity of the results to the type of classifier used, we also train a neural network using the same audio features. We use the estimated class probability output of the neural network to score each audio frame as described in Section II-D. We try varying the number of hidden layers (1 and 2), number of neurons in the hidden layers (16, 32, 64), and number of epochs (20, 50, 100, 200) to classify the recorded acoustic emission data. All the activation functions in the hidden layers are chosen to be rectified linear unit (ReLU) activation functions. The final layer's activation function is chosen to be a sigmoid to get probability estimates as outputs. The network is trained using a binary cross-entropy loss function via a RMSprop optimizer [38]. We then find the best combination and compare the cross-validated accuracy values of the proposed approach and the neural network approach.

III. Results and Discussion

A. t-SNE Visualization of the Knee Acoustical Emissions from Controls and Subjects with JIA

The data from the four controls and the four subjects with JIA was visualized using t-SNE as described in Section II-C.

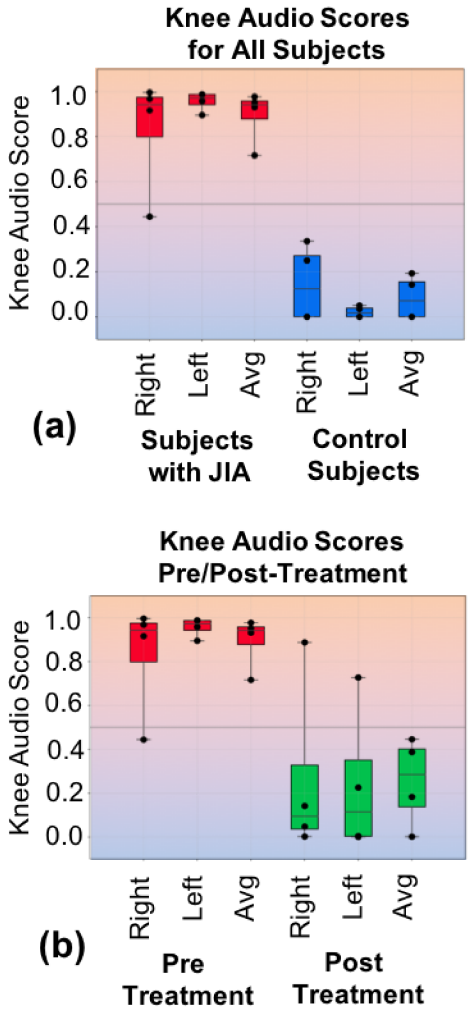

Fig. 3. (a) Knee audio scores calculated using LOSO-CV and shown with black markers. Red and blue box graphs show the data statistics for the right, left and average knee audio scores of the subjects with JIA and control subjects, respectively. Score 1.0 corresponds to the arthritis ground truth and 0.0 corresponds to the control ground truth. (b) Knee audio scores for the subjects with JIA before and after treatment. Red and green box graphs show the data statistics for the right, left and average knee audio scores pre and post treatment, respectively. The knee audio scores decreased following treatment for the subjects, consistent with the clinical observations and symptoms.

The t-SNE output using all features is plotted in Fig.2(a). In this plot, we observe that the healthy subjects form two different clusters. We plotted representative signals in the time domain and their spectrograms from each of the clusters in Fig.2(a). To determine the cause of the two distinct clusters formed by the healthy controls knee sounds, we compared the time-domain plots and spectrograms of a representative subject from these two clusters (Fig.2(c) vs. 2(d)). The plot in Fig.2(c) is of one of the healthy controls from the cluster that overlapped with the JIA data clusters in the t-SNE plots. It can be seen that the signal shown in this plot has a number of clicks. Fig.2(d) is from a healthy control from the cluster with no overlap of the JIA data clusters in the t-SNE. This plot has no sizeable clicks. It can be inferred from the difference in these sound profiles that the healthy data with some clicks led to the overlap in the t-SNE analysis with the arthritis data that included much more prevalent and louder clicks. On the other hand the healthy subjects with no clicks form the t-SNE clusters that are distinct from the JIA group. This is simply a minor issue that will not really have an impact on the classification process, but it is an interesting scientific observation worth mention.

The subjects with JIA (Fig.2(b)) show periodic high energy clicks in each flexion-extension cycle. These clicks have frequency components that reach up to 20 kHz. The most important characteristic of these signals is the audible "clicks" that have a spike-like appearance in the time domain plot. This unique, consistent 'time domain' pattern of the signals belonging to the group with JIA results in them having distinctive time-domain features. On the other hand, the high bandwidth 'frequency characteristics' of these signals result in distinct spectral features which increases the separation ability of our feature set. More work is needed to determine the origin of these high frequency clicks, but we hypothesize that they occur due to the degraded articulating surface creating increased friction in the joints of patients with JIA [39],[40]. The degraded articulating surface of the knees develops from the chronic inflammation of the synovial membrane and reduced joint space [41].

B. Knee Audio Score Assessment

Fig.3(a) shows the knee audio scores of the right legs, left legs and the average of both for all subjects. A knee audio score of 1.0 corresponds to an involved joint with arthritis; a score of 0.0 corresponds to a healthy knee. A subjects score was calculated by training the model on all of the other subjects (LOSO-CV). The average left knee audio score was 0.022±0.025 and 0.96±0.044 for the control group and group with JIA respectively ($p<0.05$ using Mann-Whitney U Test). The average right knee audio score was 0.147±0.17 and 0.83±0.26 for the control group and group with JIA, respectively ($p<0.05$ using Mann-Whitney U Test). The average knee audio score was 0.085±0.099 and 0.89±0.012 for the control group and group with JIA respectively ($p<0.05$ using Mann-Whitney U Test). The calculated knee audio scores were closer to 1.0 for the subjects with JIA while the scores were closer to 0.0 for the control subjects as expected. It can be observed that the right and left knee audio scores for both JIA and control subjects differed by 0.2 (20% of the score's range). Therefore, using this method, we can possibly detect and monitor asymmetries in symptoms as well. This capability would be useful in tracking compensation when one of the knees is more symptomatic than the other by comparing the calculated knee audio scores.

We classified the subjects with average audio scores higher than 0.5 as JIA and lower than 0.5 as healthy. This led to a cross-validated accuracy of 100% for our dataset. Additionally, we classified each signal frame as JIA or healthy (as explained in Section II-D) which led to a cross-validated accuracy of 92.3%.

Next, in an attempt to quantify the response of the subjects with JIA to treatment, the knee audio scores were calculated

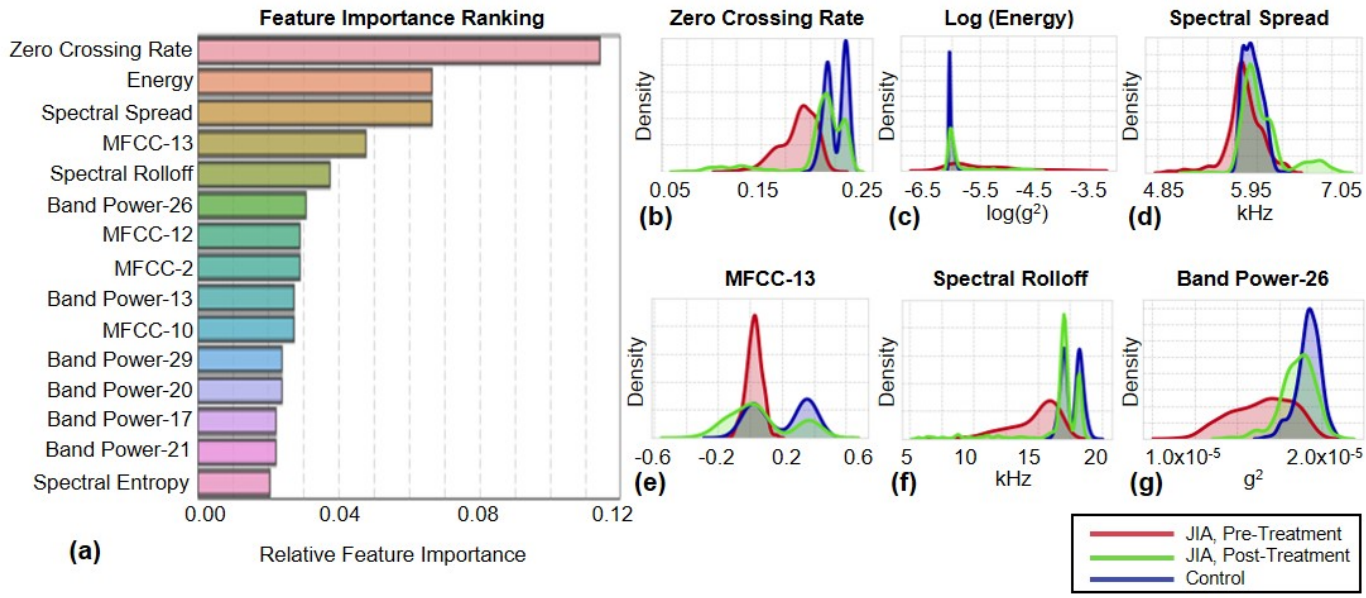

Fig. 4. (a) Feature importance ranking showing the top fifteen features in terms of salient information provided for classification. (b)-(g) Distributions of the top six features for the control subjects and pre- and post-treatment subjects with JIA. The distributions were calculated using kernel estimation.

for the data recorded pre-treatment and 3-6 months post-treatment (Fig.3(b)). Again, the data statistics for the right leg, left leg and averaged knee audio scores are plotted. The right knee audio score for these subjects decreased from 0.83±0.26 to 0.27±0.42 ($p>0.05$ using Wilcoxon Test). The left knee audio score for these subjects decreased from 0.96±0.044 to 0.24±0.34 ($p<0.05$ using Wilcoxon Test). The average knee audio score for these subjects decreased from 0.89±0.012 to 0.25±0.20 ($p<0.05$ using Wilcoxon Test). Of note, each of the post-treatment audio scores fell below our selected 0.5 classification threshold and thus were classified as healthy. The before and after treatment scores were statistically significant except the right knee scores, where statistical significance was not present as one subject's score increased post-treatment.

C. Feature Importance Ranking

The fifteen most important features from the feature importance ranking are shown on Fig.4(a). Features from all sets (time, spectral, MFCC and band power) can be seen within the top fifteen features, showing the importance of a diverse feature set. The Kernel Density Estimated distributions of the top six features are shown for the control, pre-treatment JIA and post-treatment JIA data [37]. It can be seen in Fig.4(b) that control subjects have higher Zero Crossing Rate (ZCR) than the subjects with JIA as the signals from the control group are closer to the noise floor. Fig.4(c) shows that the energy distribution of the signal frames is slightly narrower for the control subjects compared to the subjects with JIA. This can be attributed to the fact that the subjects with JIA have a more variable set of signal frames ranging from silent to ones that contain clicks. The healthy subjects mainly contain low energy (silent) signal frames. A similar observation can be made for the spectral spread in Fig.4(d) for the same reasons. In Fig.4(f), we see that the subjects with arthritis have lower spectral rolloff on average than the controls. This might be a similar phenomenon as white noise having a high bandwidth since the healthy subjects have signals that are closer to the noise floor and thus a higher spectral rolloff. The differences in MFCC-13 and band power-26 seen in Fig.4(e) and (g) are due to the differences in the frequency content of signals from the two groups. The differences seen in the feature distributions (Fig.4(b)-(g)) between the pre-treatment JIA and control groups demonstrates the ability of separating these groups based on those features. The post-treatment data is distributed similarly to the healthy control subjects indicating that the treatment was successful and the patients with JIA returned to their healthy baseline.

D. Effect of Model Type

The best result we obtained using a neural network was 72.9% cross-validated accuracy with 2 layers, 32 neurons and 100 epochs. As seen, we observed a great discrepancy between the result of our proposed algorithm (92.3%) and the one we calculated using neural networks (72.9%). Although using neural networks is an appropriate approach for audio and speech processing applications, it is not the best fit for this study because of the limited size of our dataset. Nevertheless, as we recruit more subjects for the future studies, we will have a chance to apply deep neural networks properly as well.

IV. Conclusion

In this paper, we described the recording and analysis of acoustical emissions from the knee joints of healthy subjects and subjects with JIA. We demonstrated that acoustical emissions acquired from the knee joint can potentially be used as a quantitative metric in the assessment and monitoring of JIA.

We first visualized knee acoustical emission data acquired from control subjects and subjects with JIA using t-SNE plots. We explored the ability of different feature sets to distinguish between the two groups. Following this analysis, we designed an algorithm to compute a knee audio score. This algorithm utilizes a soft classification model based on gradient boosted trees. We showed that our algorithm can accurately distinguish control subjects and subjects with JIA for our dataset using LOSO-CV. We also demonstrated how the knee audio score of four subjects with JIA changes before and after treatment. Finally, to understand the contribution of individual audio features to our algorithm, we performed feature ranking. This analysis revealed that elements from all the feature sets that we calculated contributed to some degree in the knee audio score calculation.

This paper provides, for the first time to the best of our knowledge, a framework for extracting a clinically-relevant and actionable joint health score from acoustical emission signals measured with miniature sensors that can be embedded into a wearable system for home use. The approach leverages several key innovative concepts, including the adaptation of a gradient boost regression algorithm as a soft classifier, and the use of t-SNE as a means of discovering features of relevance from the high-dimensional acoustical data. While the studies performed were with a small group of patients, the results are already statistically significant and suggest that the approach holds merit and should be tested and validated further in larger populations. The concept of delivering a joint health score to the physician or patient to assist in clinical decisions can be extended beyond JIA to assistive rehabilitation following musculoskeletal injury; such a joint health score can, for example, be used to determine when an athlete rehabilitating an acute injury can resume certain activities or intensities of activities.

In future studies we will investigate how different methods of interfacing the sensors to the skin affects the acoustical emission signals. Furthermore, we will explore the feasibility of integrating the sensors to a knee brace or sleeve for continuous monitoring. We will also study the underlying physiology and source of the knee acoustical emissions using cadaver models. Most importantly, we will validate the statistically significant, but still preliminary, results from this paper in a larger dataset of patients and controls. Upon completion of rigorous testing and validation in a larger population of subjects, the proposed approach can potentially be used to deliver a joint health score to the physician or patient to assist in clinical decisions.

References

[1] P. J. Manners and C. Bower, "Worldwide prevalence of juvenile arthritis why does it vary so much?" *The Journal of Rheumatology*, vol. 29, no. 7, pp. 1520–1530, 2002.

[2] K. Oen, "Comparative epidemiology of the rheumatic diseases in children," *Current Opinion in Rheumatology*, vol. 12, no. 5, pp. 410–414, 2000.

[3] B. Prakken, S. Albani, and A. Martini, "Juvenile idiopathic arthritis," *The Lancet*, vol. 377, no. 9783, pp. 2138–2149, 2011.

[4] A. Ravelli and A. Martini, "Juvenile idiopathic arthritis," *The Lancet*, vol. 369, no. 9563, pp. 767–778, 2007.

[5] J. Packham and M. Hall, "Long-term follow-up of 246 adults with juvenile idiopathic arthritis: functional outcome," *Rheumatology*, vol. 41, no. 12, pp. 1428–1435, 2002.

[6] A. D. Woolf and B. Pfleger, "Burden of major musculoskeletal conditions," *Bulletin of the World Health Organization*, vol. 81, no. 9, pp. 646–656, 2003.

[7] B. Manaster, C. Andrews, J. Crim *et al.*, *Diagnostic and surgical imaging anatomy: Knee, Ankle,Foot*. Amirsys, 2007.

[8] G. F. McCoy, J. D. McCrea, D. E. Beverland, W. G. Kernohan, and R. Mollan, "Vibration arthrography as a diagnostic aid in diseases of the knee. a preliminary report," *Bone & Joint Journal*, vol. 69, no. 2, pp. 288–293, 1987.

[9] S. Cai, S. Yang, F. Zheng, M. Lu, Y. Wu, and S. Krishnan, "Knee joint vibration signal analysis with matching pursuit decomposition and dynamic weighted classifier fusion," *Computational and Mathematical Methods in Medicine*, vol. 2013, 2013.

[10] C. N. Teague, S. Hersek, H. Töreyin, M. L. Millard-Stafford, M. L. Jones, G. F. Kogler, M. N. Sawka, and O. T. Inan, "Novel methods for sensing acoustical emissions from the knee for wearable joint health assessment," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 8, pp. 1581–1590, 2016.

[11] Y. Wu, *Knee joint vibroarthrographic signal processing and analysis*. Springer, 2015.

[12] S. C. Abbott and M. D. Cole, "Vibration arthrometry: a critical review," *Critical Reviews in Biomedical Engineering*, vol. 41, no. 3, 2013.

[13] W. E. Blodgett, "Auscultation of the knee joint," *The Boston Medical and Surgical Journal*, vol. 146, no. 3, pp. 63–66, 1902.

[14] A. Steindler, "Auscultation of joints." *JBJS*, vol. 19, no. 1, pp. 121–136, 1937.

[15] M. L. Chu, I. Gradisar, M. Railey, and G. Bowling, "An electro-acoustical technique for the detection of knee joint noise." *Medical Research Engineering*, vol. 12, no. 1, pp. 18–20, 1976.

[16] R. A. B. Mollan, "Vibration emission in bone and joints," Ph.D. dissertation, Queen's University, 1981.

[17] D. A. Winter, *Biomechanics and motor control of human movement*. John Wiley & Sons, 2009.

[18] K. Umapathy and S. Krishnan, "Modified local discriminant bases algorithm and its application in analysis of human knee joint vibration signals," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 3, pp. 517–523, 2006.

[19] S. Krishnan, R. M. Rangayyan, G. D. Bell, and C. B. Frank, "Adaptive time-frequency analysis of knee joint vibroarthrographic signals for noninvasive screening of articular cartilage pathology," *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 6, pp. 773–783, 2000.

[20] T.-F. Lee, W.-C. Lin, L.-F. Wu, and H.-Y. Wang, "Analysis of vibroarthrographic signals for knee osteoarthritis diagnosis," in *Genetic and Evolutionary Computing (ICGEC), 2012 Sixth International Conference on*. IEEE, 2012, pp. 223–228.

[21] J.-H. Lee, C.-C. Jiang, and T.-T. Yuan, "Vibration arthrometry in patients with knee joint disorders," *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 8, pp. 1131–1133, 2000.

[22] R. M. Rangayyan and Y. Wu, "Screening of knee-joint vibroarthrographic signals using statistical parameters and radial basis functions," *Medical & Biological Engineering & Computing*, vol. 46, no. 3, pp. 223–232, 2008.

[23] K.-R. Müller, M. Tangermann, G. Dornhege, M. Krauledat, G. Curio, and B. Blankertz, "Machine learning for real-time single-trial eeg-analysis: from brain–computer interfacing to mental state monitoring," *Journal of neuroscience methods*, vol. 167, no. 1, pp. 82–90, 2008.

[24] U. Hoffmann, J.-M. Vesin, T. Ebrahimi, and K. Diserens, "An efficient p300-based brain–computer interface for disabled subjects," *Journal of Neuroscience methods*, vol. 167, no. 1, pp. 115–125, 2008.

[25] B. Hamadicharef, H. Zhang, C. Guan, C. Wang, K. S. Phua, K. P. Tee, and K. K. Ang, "Learning eeg-based spectral-spatial patterns for

attention level measurement," in *Circuits and Systems, 2009. ISCAS 2009. IEEE International Symposium on*. IEEE, 2009, pp. 1465–1468.

[26] W.-C. Lin, T.-F. Lee, S.-Y. Lin, L.-F. Wu, H.-Y. Wang, L. Chang, J.-M. Wu, J.-C. Jiang, C.-C. Tuan, M.-F. Horng *et al.*, "Non-invasive knee osteoarthritis diagnosis via vibroarthrographic signal analysis," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 5, no. 3, pp. 497–507, 2014.

[27] L.-K. Shark, H. Chen, and J. Goodacre, "Knee acoustic emission: a potential biomarker for quantitative assessment of joint ageing and degeneration," *Medical Engineering & Physics*, vol. 33, no. 5, pp. 534–545, 2011.

[28] S. Hersek, M. B. Pouyan, C. N. Teague, M. N. Sawka, M. L. Millard-Stafford, G. F. Kogler, P. Wolkoff, and O. T. Inan, "Acoustical emission analysis by unsupervised graph mining: A novel biomarker of knee health status," *IEEE Transactions on Biomedical Engineering*, 2017.

[29] T. Giannakopoulos and A. Pikrakis, *Introduction to Audio Analysis: A MATLAB® Approach*. Academic Press, 2014.

[30] L. v. d. Maaten and G. Hinton, "Visualizing data using t-sne," *Journal of Machine Learning Research*, vol. 9, no. Nov, pp. 2579–2605, 2008.

[31] G. J. Berman, D. M. Choi, W. Bialek, and J. W. Shaevitz, "Mapping the stereotyped behaviour of freely moving fruit flies," *Journal of The Royal Society Interface*, vol. 11, no. 99, p. 20140672, 2014.

[32] M. West, C. Blanchette, H. Dressman, E. Huang, S. Ishida, R. Spang, H. Zuzan, J. A. Olson, J. R. Marks, and J. R. Nevins, "Predicting the clinical status of human breast cancer by using gene expression profiles," *Proceedings of the National Academy of Sciences*, vol. 98, no. 20, pp. 11 462–11 467, 2001.

[33] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2016, pp. 785–794.

[34] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of Statistics*, pp. 1189–1232, 2001.

[35] T. G. Dietterich, "Ensemble learning," *The Handbook of Brain Theory and Neural Networks*, vol. 2, pp. 110–125, 2002.

[36] S. Hersek, H. Töreyin, C. N. Teague, M. L. Millard-Stafford, H.-K. Jeong, M. M. Bavare, P. Wolkoff, M. N. Sawka, and O. T. Inan, "Wearable vector electrical bioimpedance system to assess knee joint health," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 10, pp. 2353–2360, 2017.

[37] J. Friedman, T. Hastie, and R. Tibshirani, *The elements of statistical learning*. Springer series in statistics New York, 2001, vol. 1.

[38] S. Ruder, "An overview of gradient descent optimization algorithms," *arXiv preprint arXiv:1609.04747*, 2016.

[39] S. Nalband, A. Sundar, A. A. Prince, and A. Agarwal, "Feature selection and classification methodology for the detection of knee-joint disorders," *Computer Methods and Programs in Biomedicine*, vol. 127, pp. 94–104, 2016.

[40] I. Nakamura, K. Michishita, M. Tanno, and K. Ito, "Synovial impingement after posterior cruciate-retaining total knee arthroplasty for rheumatoid arthritis," *Journal of Orthopaedic Science*, vol. 11, no. 3, pp. 303–307, 2006.

[41] A. Ravelli *et al.*, "Toward an understanding of the long-term outcome of juvenile idiopathic arthritis," *Clinical and Experimental Rheumatology*, vol. 22, no. 3, pp. 271–275, 2004.