Proceedings of the ASME 2021 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference
IDETC/CIE 2021
August 17-20, 2021, Virtual Conference, USA

DETC2021-70654

DRAFT: ANALYTICAL APPROACH TOWARDS OBJECTIVE ASSESSMENT OF ORTHOPEDIC SURGERY TRAINING: A PRELIMINARY ANALYSIS

Ronak R. Mohanty, Shantanu Vyas, Aman Nigam, Bruce L. Tai, Vinayak R. Krishnamurthy[§]
J.Mike Walker '66 Department of Mechanical Engineering
[§]Department of Computer Science and Engineering (by Affiliation)
Texas A&M University, College Station, TX, USA.
Emails: ronakmohanty@tamu.edu, svyas@tamu.edu, nigamaman21@tamu.edu, vinayak@tamu.edu btai@tamu.edu

ABSTRACT

Assessment techniques for orthopedics training are primarily subjective, and often based on qualitative metrics. In this paper, we propose an analytical approach for the quantitative assessment of orthopedic surgery training, specifically, bone drilling. Our goal in this paper is to help improve orthopedics training by providing a means to assess the resident training progress. To this end, we introduce a novel metric that assigns a unique signature to an individual's drilling activity based on their drilling trajectory, and we compare it with the signatures of experienced surgeons. We conduct a simple bone-drilling experiment with surgeons (experts) and non-expert users on a hybrid (physical - digital) setup consisting of $3D$ printed bone surrogates that emulate physical and perceptual properties of a human bone across the young and old age groups. Our preliminary analysis of drilling signatures across expert and non-expert users showcases a perceivable distinction between the two user groups. Our results indicate that the drilling signature captures the user's drilling behavior, thus, characterizing drilling performance between novice and experts.

1 Introduction & Background

Precise motor control is one of the fundamental skills acquired through years of practice, furnishing with the capability to perform skilled tasks such as sculpting, carving, painting, precision manufacturing, as well as, medical surgeries. Orthopedic residency programs are designed around a similar goal of facilitating resident surgeons with training and guidance to perform complex surgeries. Orthopedic surgeries are tasks requiring the ability to make fine and careful motor movements due to the inherent risk of damaging the bone, or any critical nerve or tissue [1], thereby endangering the patient. Therefore, proper training and evaluation of orthopedic residents is a critical task.

From our interviews with expert surgeons, it is clear that the evaluation of orthopedic resident training performance is not merely qualitative but completely **subjective**. The expert surgeon observes the residents visually and grades their performance based on that observation. Our main goal in this paper is to help improve orthopedic resident training by facilitating an objective, and preferably quantitative, means to evaluate training performance of orthopedic residents. For this, we designed a hardware setup to capture real-time bone-drilling data and developed a drilling signature to assess the quality of the drilling tasks. We demonstrate our evaluation approach by collecting data from an expert surgeon, computing the signature model of the expert, and subsequently evaluating the drilling performance of novice participants through a controlled lab experiment.

1.1 Orthopaedic Surgery Training

Many different techniques have been proposed to improve resident training to increase patient safety and reduce the time required for residents to become proficient with the necessary skills of bone drilling and screw placement. Recently, much research has been reported to improve bone-machining skills through surgical simulation in arthroscopy, orthopedic surgery, and craniomaxillofacial surgery. The majority of these studies are

focused on bone drilling [2]. Previous systems have employed cadaveric training, animal bones, and virtual reality with some levels of success.

Cadaveric training and animal model are the traditional orthopedic simulators used for surgical training. They serve as the best alternative to live surgery [3] as they offer realism, tactile feedback, and awareness about the anatomical construction [4]. However, human cadavers are expensive, and their limited availability restricts their widespread use. Further, cadavers require regular maintenance in special facilities and are also susceptible to disease transmission. To address the shortcomings of cadaveric training, VR-based orthopedic simulations have been widely researched. They are employed for skill acquisition, decision-making, pre-operative planning [5], and real surgery [6]. These simulators eliminate the requirement of cadavers or animal bones and reduce operative time to improve the performance of surgical trainees [7]. However, VR simulators are expensive and do not provide a realistic environment nor physical, tactile feedback.

By taking advantage of both physical and virtual-reality simulators, this study proposes a hybrid bone drilling simulator that will employ 3D printed-customizable bone simulant along with real-time visual information provided to the operator. The system collects motion data during drilling into a bone surrogate. Different bone conditions and geometries can be replicated with 3D printing and special polymer treatment. This versatility with the visual feedback will allow the system to inform both the instructor and the trainee about the drilling performance through the proposed evaluation technique.

1.2 Evaluating Bone-Drilling Performance

Past researches reveal that there exists a variable approach while judging bone drilling performance. Although teaching surgical technique is a significant residency task, traditional technical skills assessments are inconsistent and subjective [8]. As discussed in the previous section, different types of surgical simulators offer different kinds of drilling environment. Hence the performance assessment technique also varies. One common factor in assessment is the onset of osteonecrosis, a disease resulting from thermal damage to the bone tissue. [9].

The temperature generated during the bone drilling depends upon various parameters such as drill geometry, rotational speed, drilling forces, and cooling. Most of the investigations are related to rotational speed and drilling forces. There is a general agreement in the literature that the temperature increases with the drill speed [10]. The drill feed rate is another parameter in determining the heat generated during bone drilling. Generally, at higher feed rates, the drilling time decreases, and less heat is accumulated. However, high feed rates might also imply higher forces and higher heat generation [11].

Another approach in gauging drilling performance is the plunging distance. Plunging distance is the distance that a drill bit might travel after drilling through the second cortical region. Bone drilling requires precision in hand motion, and a greater plunging distance may cause soft tissue damage. [12] built a low-cost drilling simulator to train orthopedic residents in reducing the drill plunging depth. This study found that the plunging depths of the junior residents were significantly greater than orthopedic specialists (7.00 mm vs. 5.28 mm). However, no significant difference was observed between the senior residents and the orthopedic experts (6.33 mm vs. 5.28 mm).

Different researchers have suggested various approaches to judge drilling performance based on individual parameters. However, not much work has been reported in developing a single parameter to compare the drilling performance of a novice to that of a specialist. As seen, all the performance parameters are related to motion (feed rate, time in the bone, plunging). Therefore, this study aims to develop a metric based on the motion data to distinguish expert levels.

2 Bone Drilling Data Collection

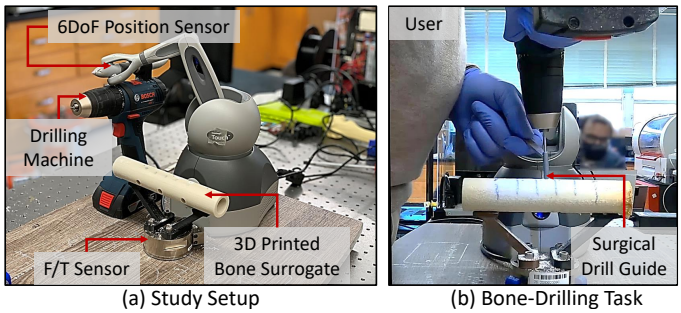
2.1 Physical Setup Design

The motivation behind designing a hybrid simulator was to develop a versatile device that provides the ability to utilize 3D printed-customizable cortical bones and furnishes performance feedback. As shown in Fig. (X), the physical set-up is focused on the 3D-printed bone-mimicking composite. The bone simulant is manufactured using a special 3D-printable plaster employing the binder jetting technology. 3D printing can create complex geometries, and thus, the testing sample is customizable and can be patient or bone-specific. In addition, the mechanical properties can be modified at the post-printing stage with an epoxy treatment. This treatment can produce various grades of hardness and toughness to simulate bones at different ages or conditions. In this study, two hardness grades were used, which replicated young (healthy) bone and osteoporotic bone. Based on the reviews from expert surgeons, the young bone's hardness was designed to be 95 while that of an osteoporotic bone was 45 on the Shore D scale.

In this study, the bone simulant was clamped to a 6-dof ATI gamma force sensor (ATI Industrial Automation, USA), which helped record the drilling force and torque in 3 axes. Further, a 1700 rpm Bosch hand drill (Bosch, USA) was connected to a Geomagic Touch (3D Systems, USA) haptic device, which rendered the position and orientation in 3 axes. Robust construction and ease of use were incorporated as the design criteria.

2.2 Data Recording and Processing

Bone drilling and fracture fixation require precision and accuracy in hand movement. As discussed in the previous section, the motion data was recorded using a 6-dof haptic device. To accurately judge a drilling performance, the position recorded at the stylus end needed to be offset to the drill tip. This was



(a) Study Setup (b) Bone-Drilling Task

FIGURE 1. (a) We designed a hybrid (physical-digital) setup for bone drilling training using 3*D*-printed bone surrogates; (b) Bone drilling task being performed on our setup using a surgical drill guide for better stability

achieved using a forward kinematic model, the details of which are discussed in the subsequent sections. Also, to compute the corresponding force at a particular position, the haptic device and F/T sensor data were synchronized during the postprocessing. Data synchronization also aided in accurately finding the drilling cycle's start and endpoint to eliminate redundant data.

2.3 Evaluation Methodology

The literature review and preliminary analysis found that more emphasis should be given to motion parameters (feed, speed, and plunging) as they directly affect temperature rise and overall drilling performance. To evaluate a particular drilling performance, initially, individual parameters were manually studied. A custom MATLAB function library was developed to extract these individual parameters from raw data. However, to make generalized conclusions and differentiate between expert and novice users, an extensive set of data may be required. Though manual analysis may have its benefits, it is time-consuming and tedious for larger data sets. This paper presents an automated approach to segregate the data into the corresponding sections in the bone and analyze the signature associated with them.

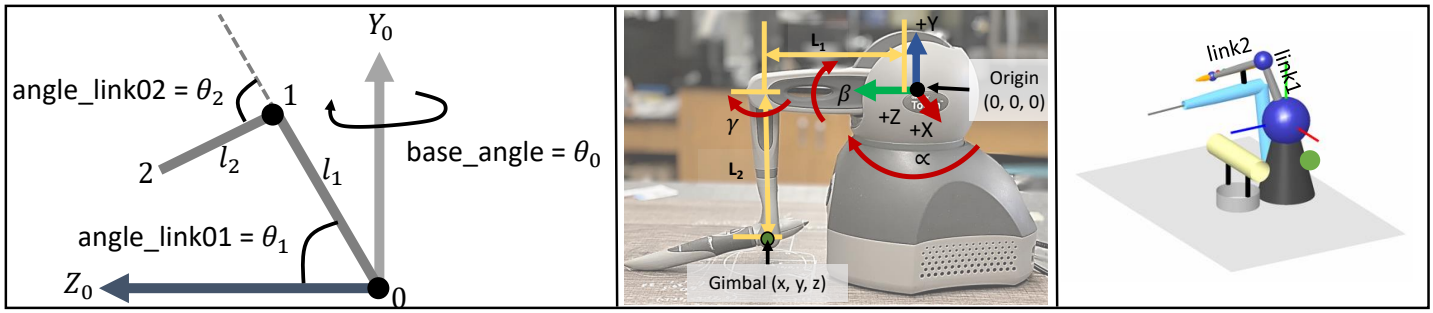
3 Technical Approach: Drilling Signature

Our goal is to characterize a given drilling task with a signature in order to facilitate experienced surgeons with means to assess the training progress of the resident surgeons over the duration of their residency. Surgical bone drilling is a highly constrained and precise task meaning that the 3*D* trajectory of the drill is more or less straight going through the bone cortices and coming out. In such a scenario, it stands to reason that getting a measurable difference between the drilling trajectory of two individuals would be inherently challenging. Our preliminary findings indicated that 3*D* position of points in the operator's trajectory, net force, and drilling speed are insufficient metrics to draw a clear distinction between a novice user and an expert user in terms of drilling performance. One of the primary reasons that attributes to this limitation is the constrained nature of the bone-drilling activity. As a case in point, the drill tip is surrounded by the bone-material for the entire drilling duration except for the hollow cavity between two bone cortices in a *to-and-fro* motion — the user enters and leaves the bone from the same position. Consequently, the trends for the aforementioned metrics across the novice user groups look similar when compared to an expert surgeon thereby making it difficult to draw comparisons for any type of performance assessment of the drilling activity. In this work, we introduce a curve signature metric that we call the **drilling signature**. The main idea for the drilling signature is to help evaluate a resident's drilling task by comparing their trajectory to a gold standard; which is an expert surgeon for orthopedics training.

3.1 Rationale behind Drilling Signature

The design of our drilling signature is based on our observations of the difference between how expert and novice individuals control the drill. After several years of practice, expert surgeons insert the drill into the first cortex at a relatively high speed, move in reasonably slowly to the second cortex and pay attention to overshooting once they pass through the second cortex (so as not to damage the tissue). On the other hand, novices begin cautiously right from the very beginning and therefore end up losing control at the end of the first cortex itself. This difference in expert and novice behavior has two implications on the geometry of the drilling trajectory. First, even though the nominal paths are "*straight*" (both going in and coming out), they are not sampled equally with time. In fact, from our experience, novice trajectories are generally densely sampled because they tend to maintain a slower speed in hope of getting better control. The second and more important observation is that the effect of noise generated through the drill's vibration, the bone's interaction with the drill, and the manual response are difference between experts and novices. It is these two observations — the sampling rate and the noise profile along the drilling trajectory — that inform the design a method to objectively (and preferably quantitatively) distinguish between novices and experts.

To characterize the sampling rate and noise profile along the drilling trajectories for a given operator, we draw from signal processing. There are many previous works noise characterization for applications such as fault analysis of electrical and mechanical components [13–15], cognitive neuroscience [16], biological spectral analysis [17], and image processing [18]. More recently, Cheyrev et al. [19] explain the use of path signatures as well as their use in machine learning. Signature-based machine learning models have also been used for distinguishing bipolar disorder and borderline personality disorder [20]. Few works discuss signature method from the point of pattern recognition in the cequel



(a) Schematic of the Kinematics Model

(b) Kinematics Model Overlayed on the Position Sensing Device

(c) Animated Representation using the Kinematics Models

FIGURE 2. We created a forward kinematics model based on raw data provided by the device; and computed the precise (σ = 3 mm) drill tip position.

clinical trial [21], also, make use of the path signatures to predict a diagnosis of Alzheimer's disease [22]. However, with our current focus on $3D$ drilling trajectory, we borrow from prior works using the Laplace-Beltrami operator to conduct $3D$ shape analysis [23–27].

3.2 Conceptual Framework

In our approach to characterize the noise in the drilling trajectory, we leverage the notion of Laplacian smoothing. Laplacian smoothing is a well known technique [28] that has been used extensively in computer graphics for curve and surface smoothing [29]. The basic idea is simple — for a given manifold (a curve or a surface) discretized in a piece-wise linear fashion (for curve this means a poly-line, for surface it is a polygonal mesh), we replace each vertex of the manifold with the weighted average of its neighbors. In the continuous case, this is essentially the application of the Laplace equation ($\nabla^2 f = 0$) for a harmonic function (f) defined on the manifold.

From a signal processing perspective, what Laplacian smoothing achieves is that it enforces the function f to become harmonic over a period of time thereby allowing it to reach it's steady state. In fact, the same principle is applied in heat diffusion problems. Now, applying the Laplace operator directly on the coordinates of a poly-line, which is essentially how our trajectories are represented, effectively diffuses the curvature on the trajectory. Another interpretation of the operator is that an eigen-decomposition of this operator is equivalent to removing high frequency noise in the manifold [30–32]. The equivalence between noise removal and diffusion to steady state provides a powerful clue toward developing a signature for our application.

We begin by first observing that for any poly-line approximating a given curve, a repetitive application of the Laplacian smoothing will ultimately lead to a completely straight line with uniformly sampled points. However, each point on the poly-line will take a different amount of time to reach the steady state. For instance, points whose neighborhoods are already straight (low-curvature and less noisy) will take less time to reach steady state than those that are noisy or highly irregularly sampled on the curve. Now we also note that the time for a given point to reach steady state can simply be described by the number of iterations it takes for this point to become static (i.e. there is negligible difference in the location of this point between two consecutive applications of smoothing).

Based on these observations, our idea for the drilling signature is rather simple. We repetitively apply Laplacian smoothing to the trajectory and record the number of iterations it takes for each point to reach steady state. Steady state is defined as a state where the euclidean norm a point before and after smoothing at a given iteration falls below a certain threshold. Once recirded, the number of iterations for each point is normalized and referred as the signature score (s) for a point on a given curve. For a curve P with k points on it, there will be ($\mathbf{S} = (s_1, s_2, ..., s_{k-1}, s_k)$) signature scores. This signature essentially characterizes the noise as well as sampling distribution in the drilling trajectory as we originally desired (Figure 4).

3.3 Algorithm

For a given smoothing iteration i, let us consider a trajectory curve $P_i = (\mathbf{p_i^1}, \mathbf{p_i^2}, \mathbf{p_i^3}, \ldots, \mathbf{p_i^{k-1}}, \mathbf{p_i^k})$, where $p_i^j \in \mathbb{R}^{\ltimes}$. The smoothed coordinate for every j^{th} point in P_i is given as:

$$\mathbf{p}_i^j = 0.5 \times (\mathbf{p}_{i-1}^{j-1} + \mathbf{p}_{i+1}^{j+1}) \tag{1}$$

The smoothing is applied successively to the smoothed iterations of the original trajectory P_0 until:

$$\|\mathbf{p}_i^j - \mathbf{p}_{i-1}^j\| < m \in \mathbb{R} \tag{2}$$

Here, m is the threshold for euclidean norm between corresponding points of two consecutive smoothed trajectory curves to

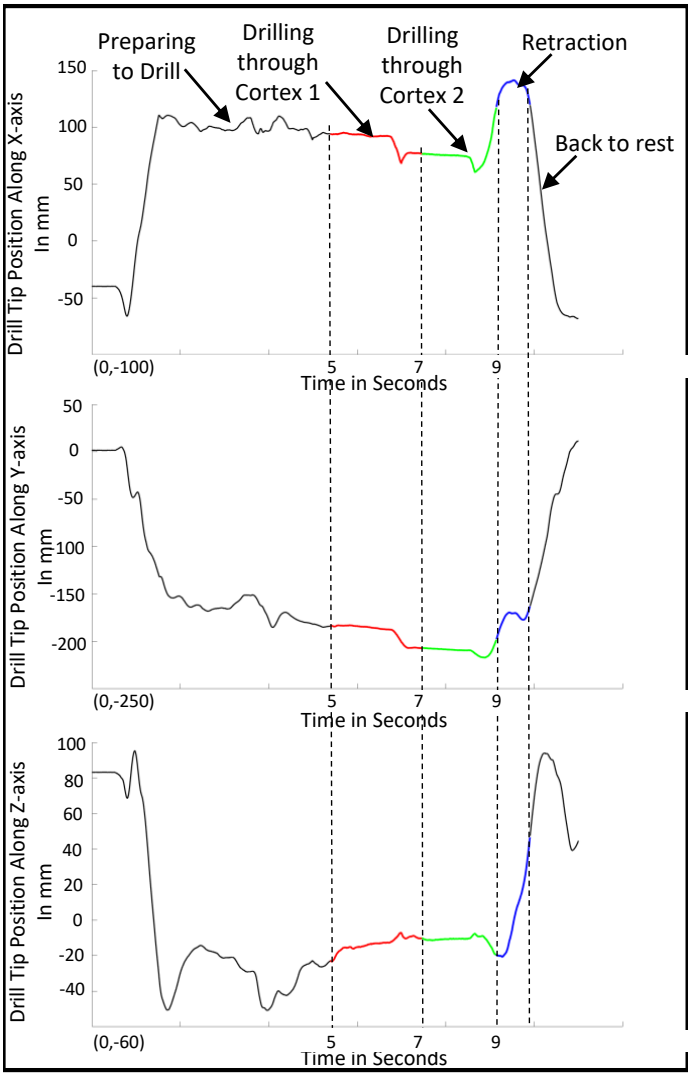


FIGURE 3. We automated the labelling process to identify different drilling regions for the raw data — *force* and *position*. The *drilling speed* is a derived metric computed from the position data.

reach steady state. We found $m = 0.005$ as a suitable threshold to verify the steady state of the drilling trajectories recorded in our data collection. The signature for each point on a given curve is computed as the normalized number of iterations to reach steady state as:

$$s_j = \frac{s_k - s_{min}}{s_{max} - s_{min}} \forall j \in [1,k] \tag{3}$$

Here, $s_k \in [0,1]$ and S_{max} and S_{min} are the maximum and minimum signature scores respectively along the entire drilling curve.

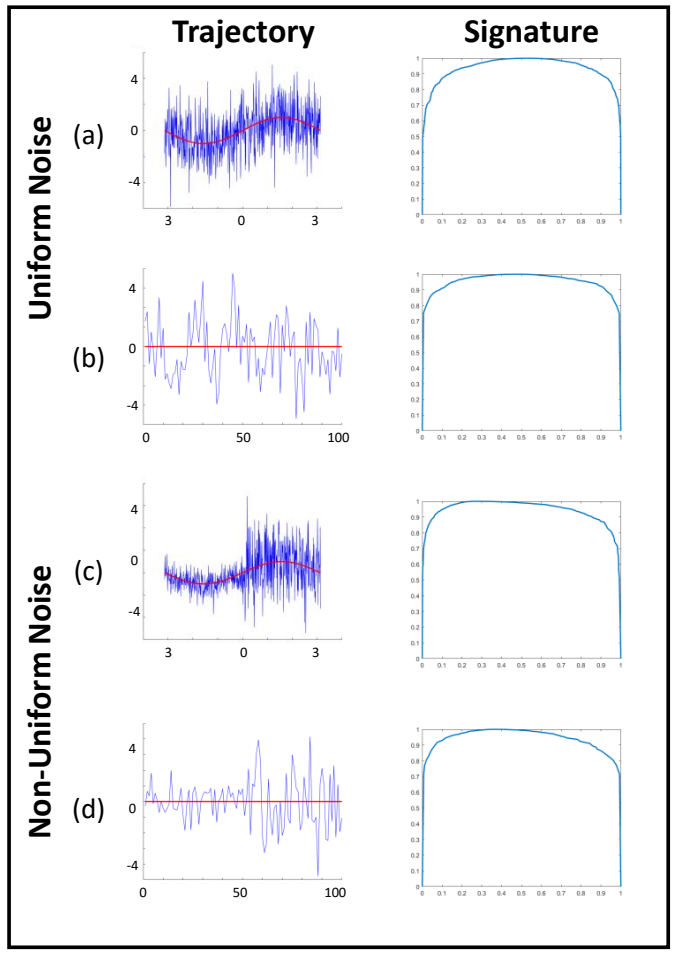


FIGURE 4. We automated the labelling process to identify different drilling regions for the raw data — *force* and *position*. The *drilling speed* is a derived metric computed from the position data.

4 Experiment

4.1 Participants

Our interviews with an expert orthopedic surgeon revealed that most first year resident surgeons haven't had prior experience with any type of bone-drilling tasks. Owing to the simplistic nature of our bone-drilling setup, it is safe to assume that first year residents fall under the category of novices akin to any non-medical student who also hasn't had the opportunity to conduct a bone-drilling task. We recruited 10 participants randomly sampled from undergraduate and graduate engineering students recruited through university advertisement[1]. The participants were within

[1]Due to the challenges presented by COVID-19, it was difficult to visit any medical center to collect data.

the age group of 18 to 30 years old. According to the information collected from the participants through a pre-study questionnaire, 8 participants had prior experience with manufacturing related drilling and 7 rated the expertise from an amateur to intermediate. Only one participant rated themselves as an expert in drilling.

While the 10 participants served as novice users for our bone-drilling experiments, we also recruited an expert user, who has been an orthopedic surgeon for over 30 years, and is responsible for training orthopedic residents in bone drilling in the Department of Orthopedic Surgery at the UT Health Science Center. Our goal is to use the expert surgeon's signature as a gold standard to evaluate the performance for the novice users.

4.2 Evaluation Tasks

Our evaluation task is a simple bone-drilling task using $3D$-printed bone surrogates and high precision tools for data recording. We designed the setup and the drilling task with three goals in mind: (a) first, we wanted to evaluate the efficacy of our setup in terms of bone material, and data recording hardware with the goal of standardizing the setup for orthopedics training, (b) second, we wanted to create a database of labelled drilling data that allows easy segmentation of position, force, speed, and other derived metrics for better analysis of a drilling task, (c) finally, we wanted to take a preliminary step towards objectifying the existing subjective and qualitative bone-drilling evaluation metrics, towards the possibility of quantifying drilling parameters in near future.

4.3 Procedure

The study involved drilling through a $3D$-printed bone surrogate for two bone variants of different hardness; emulating healthy and osteoporotic bones mechanically and perceptually. The study took approximately 30 minutes per participant with a minimum duration of 30 minutes between two consecutive participants. The drilling setup was sanitized after every experiment, also, the drill bit was ensured to be free of bone-surrogate debris from prior experiments.The study started with the participants filling up a demographic questionnaire and answering pre-screening questions inquiring their prior experience with drilling. The participants were then given a general introduction to the experimental setup and brief demonstration of the bone-drilling task was also shown. The participants and study investigators maintained a minimum distance of 6 feet during the study trials, masks worn by both parties at all times, also, there was a glass divider separating both. They were also given the option of donning a latex glove for protective reasons.

Practice: The participants began by making themselves comfortable with the setup from two perspectives. First, if they are donning a glove, does it allow them to perform the drilling task comfortably and with minimum distraction. Second, if the height of the setup is ergonomically feasible for a less constrained drilling activity. We tackled the latter by providing a platform to the participants ensuring the setup is around the waist level for them, for a comfortable drilling experience. They were asked to practice drilling using a drill guide on either bone variants as randomized by the study investigator to minimize learning bias. The practice sessions lasted for about 4 - 5 minutes as most participants had prior exposure to drilling.

Data Labeling: One of the key challenges for data labeling is the inability to identify the transition of the drill bit inside the bone-surrogate. Therefore, it is crucial that the study investigator collaborates with the participants for assistance with labeling. The participants were instructed to start drilling the moment they touched the outer surface of the bone surrogate. For transitions inside the bone, they were requested to shout "two", indicating the drill tip hitting the second cortex. Once drilled through both cortices, the participants were asked to pause the drilling momentarily before pulling the drill bit outside of the bone.

Trials: The participants were asked to drill 10 holes in all i.e. 5 per bone variant. The consecutive order of the bones was chosen randomly across participants and they completed drilling the 5 holes through their first bone before moving on to the next bone. We asked the participants to start the drilling process only when the drill bit was touching the top surface of the bone. The participants were then given a signal to start the task, after which they picked up the drilling machine, inserted the drill bit through the drilling guide and started drilling the hole. Once they had completed drilling the hole, they were asked to retract the drilling machine and keep it back on the work-desk, marking the end of the trial. They were provided with a wet wipe to clear any bone debris stuck in the flutes of the drill bit. The participants were given a short break (30-60s) after each trial, and a longer break (approx. 3 minutes) after the first 5 trials while the study investigator installed the second bone.

The same procedure was followed to collect data from the expert orthopedic surgeon as well. The expert conducted 10 drilling trials for a given bone variant.

Data: For each of the trials performed by the participants, we recorded the raw event log containing the (a) 3D position data of the drilling machine, (b) time taken to drill, (c) force, and (d) torque data for a given trial.

4.4 Expert Signature Model and Drilling Quality Metric

We use the signature scores s_j $(, j \in [1,k])$ on each point $\mathbf{p^j}$ of the drilling trajectory as our primary metric to evaluate drilling performance. The normalized signature scores per trial for the expert orthopedic surgeon are plotted with respect to the normalized arc length of the respective trial trajectory. We plot

the signature (**S**) vs. arc length (L) for all expert trials for a given bone type and compute an *average curve* for all plots. Following this, we uniformly re-sample the signature along the arc length. Finally, we compute the *average curve* for a collection of drilling signatures for multiple trials of the expert. We treat this average curve for each bone type as the expert model for the drilling performance assessment.

Drilling Quality Metric: We use the expert signature model to compute the root mean squared error (RMSE) between the model curve and a given user trial signature curve. This value is what we define as the drilling quality metric. For a given point u_i on a user trial $U = (u_1, u_2, u_3, ..., u_{k-1}, u_k)$, where $k \in \mathbb{R}^{\times}$. We compute the RMSE with reference to the corresponding point standardized model curve M as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{k}(S(m_i) - S(u_i))^2}{k}} \tag{4}$$

Here, $S(m_i)$ and $S(u_i)$ are signature scores at the ith point for the user trial curve and standardized model curve respectively.

5 Results

We discuss the results of our experiments from the point of evaluating the drilling performances for novice users. In the subsequent sections, we conduct a quantitative assessment using the drilling signature metric with the expert surgeon as our drilling performance benchmark.

5.1 Expert Models for Drilling

In order to standardize the expert surgeon data as our reference measure, we computed two expert models using drilling signatures across all expert trials for a given bone variant. However, we observed a difference in expert drilling signature scores between the two bone types, and decided to benchmark each bone variant with individual expert model based on the aforementioned approach (§4.4). First, we resampled 2000 equidistant values along the X-axis of the drilling signature plot; that represents the normalized arc length. The purpose for this larger distribution was to evenly capture the drilling signature across all expert trials for a bone variant. Further, we computed the drilling signature values for each expert trial for a bone variant at the resampled arc length values using piecewise linear approximation. Subsequently, we calculated the average of the new drilling signatures computed at the resampled points, across all the trials for a given bone variant, thus, resulting in two expert drilling signature Models across the two bone variants (Figure. 5). We use these expert models as our benchmark reference to objectively assess the quality and consistency of an user's drilling trials across both bone variants.

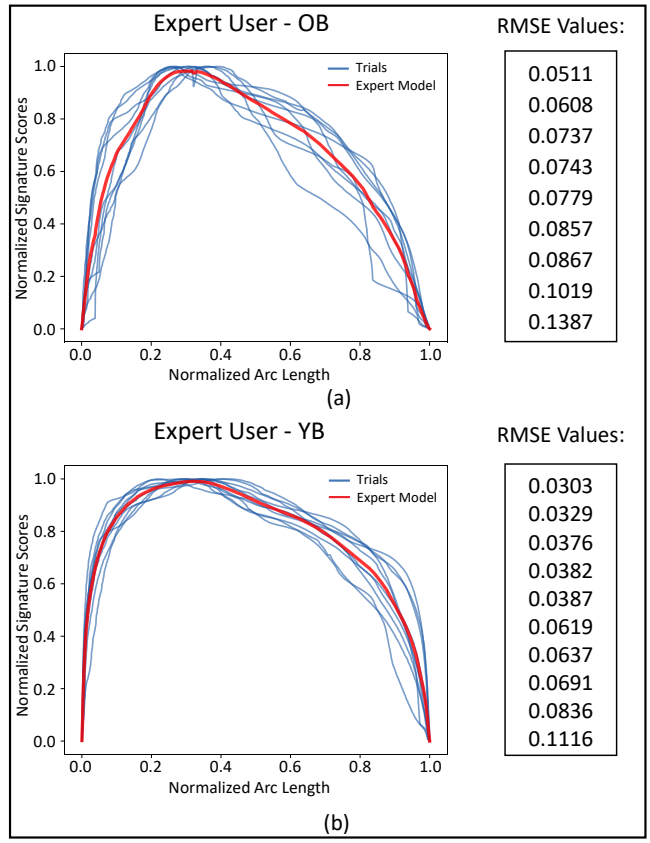


FIGURE 5. Plots showing the Signature Curves and Expert Models for the Orthopedic Surgeon's trials on (a) Osteoporotic Bone (OB) and (b) Young Bone (YB), along with the RMS Errors for each trial

5.2 Quality of Performance

Our primary goal for creating the expert models was to use them to evaluate the drilling performance of the users' trials on the two bone variants. We did this by following the aforementioned approach (§4.4) of calculating the RMSE between the expert model for the specific bone type and each of the user trials' signature curve, which we refer to as the quality of performance.

We observed that the users, in general, performed worse than the expert models across both the bone variants. While this observation was expected, the overall quality of performance of users on the Young Bone (Avg. RMSE = 0.248) was worse than that of their performance on the Osteoporotic Bone (Avg. RMSE = 0.153). This is in contrast to the expert's performance on the two bone variants, where the expert performed better on the Young Bone (Avg. RMSE = 0.056) compared to the Osteoporotic Bone (Avg. RMSE = 0.075). A reason for this could be the difference in hardness levels of the bone variants, as the Young Bones are harder and thus provide more resistance to the drilling action compared to the Osteoporotic Bone. This is further seen when we compare each of the users performances between the two bones

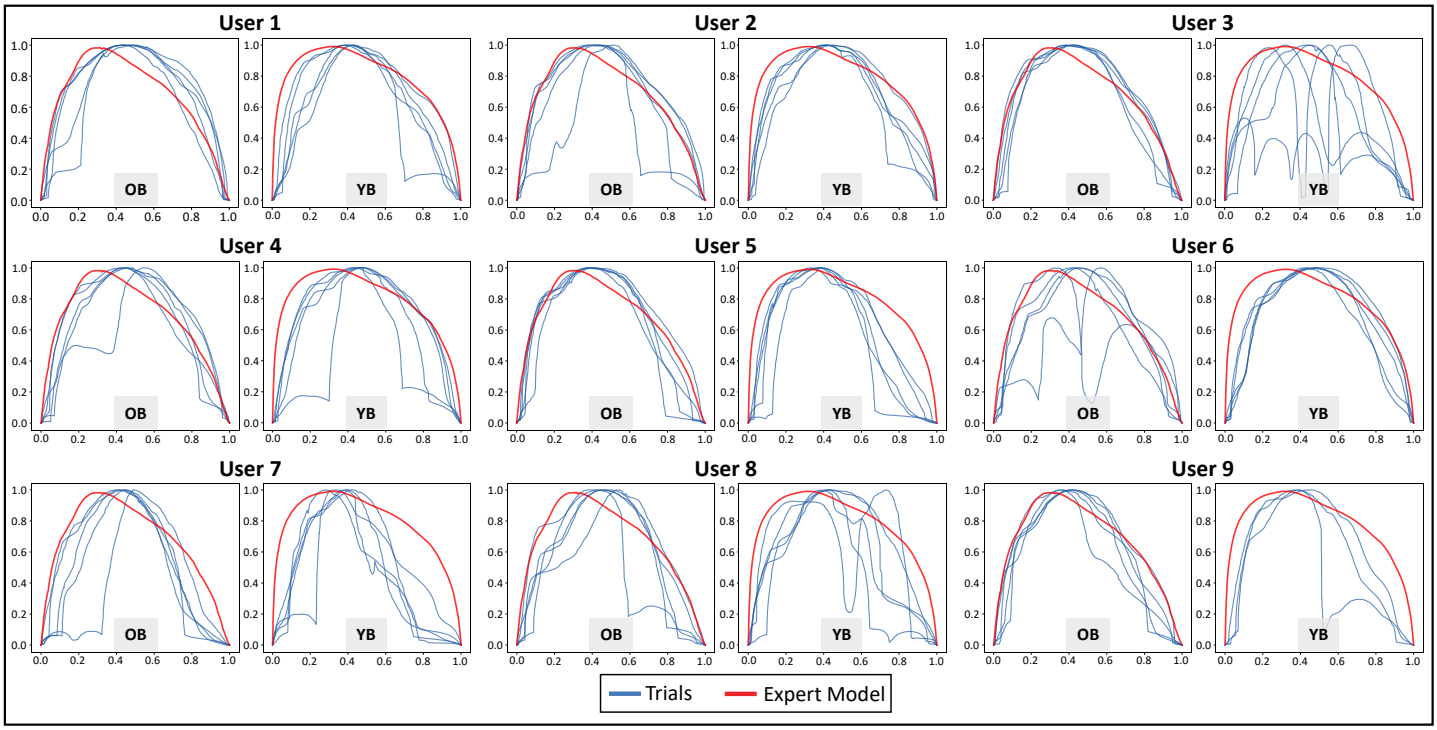


FIGURE 6. Plots showing the Signature Curves of all the users' trials across both the bones: Osteoporotic Bone (OB), Young Bone (YB) compared with the bone specific Expert Models. The x-axis for the plots correspond to the Normalized Arc Length and the y-axis corresponds to the Normalized Signature Scores.

(Table 1), as every single user performs better on the Osteoporotic Bone.

We made further observations on the nature of the users' signature curves when compared with the expert models (Figure. 6. The expert models reached their peak signature values at smaller values of the normalized arc length than all of the users trials. We observed the peak shifting to the right for almost all of users' trials, irrespective of the bone variant. One reason for this shift in peak could be due to the users drilling the bones at a slower speed compared to the expert. The expert generally completed their drilling trials much quicker than most of the users. Their approach to the different regions of the bones was also quicker, which can be explained by the steep initial slope of the expert models when compared to the gradual slope of the users' signature curves. Most users performed better on the Osteoporotic bone with respect to the initial steepness of the slope, which can again be explained by the hardness of the bones.

5.3 Consistency of Performance

We define consistency as the repeatability of a user's drilling behaviour across consecutive trials for a given bone type, as well as, across both bone variants. On analyzing the user trials with respect to the expert model, we observe consistently poor drilling behavior for the Osteoporotic Bone variant (Figure. 6) when compared with the the respective expert model. On the other hand, the drilling behaviour for the Young Bone variant was poor, but highly inconsistent as observed for the varying RMSE scores across drilling trials for all participants (Table. 1). This indicates that users were relatively comfortable drilling through the Osteoporotic bone variant due to relatively softer material hardness, also, faced minimal drilling resistance unlike the Young Bone variant.

We also observed that a higher consistency across one bone might not necessarily mean the user is experienced. This can be clearly seen for *User* 3 (*User* 3 YB plot in Figure. 6), whose RMSE for the Osteoporotic Bone was the lowest of all the users, but second highest for the Young Bone. We observe the opposite behavior for *User* 6, whose consistency across trials was better for the Young Bone (range of RMSE = 0.09), when compared with the trials on Osteoporotic Bone (range of RMSE = 0.22). The expert on the other hand, consistently performed better across both the bone types (range of RMSE for Young Bone = 0.081; Osteoporotic Bone = 0.087).

5.4 Participant Specific Observations

While we observe the general trends in quality and consistency of performance for the users in the above sections, there are some interesting things to note about some of the specific signature curves of the users. Firstly, almost all users have at least one trial, where the signature curves dip and then rise again to form a second peak (See trials from *User* 3 *YB, User* 4 *OB, User* 6 *OB, User* 7 *YB, User* 8 *OB/YB and User* 9 *YB* in Figure. 6). This dip is usually seen towards the middle and later half of the sections of the curve and it signifies a lower number of iterations required to straighten the drilling path curve, which might be caused by the users stopping their drilling motion for a small amount of time, before proceeding through the remaining regions of the bones. This is better explained by the sudden rise and forming of the second peak, which shows the continuation of the drilling motion. Another interesting observation is the sudden change in the slope of the signature curves for some of the participants (See trials from *User* 1 *OB, User* 2 *OB/YB, User* 4 *YB, User* 5 *OB/YB, User 7 OB/YB, User 9 OB* in Figure. 6). These sudden changes in the slopes are usually seen in the initial and final sections of the curves and may correspond to a sudden change in the drilling action by the user. The flatter slopes signify very small changes in the signature scores, which may be due to a slow and gradual drilling motion. The steeper slopes signify a big change in the signature score, which may be a result of a faster drilling approach. This may be explained by a slow-start and slow-end approach that novice users may use as a means to be careful while completing the task.

Osteoporotic Bone						**Average**
User 1	0.11	0.13	0.16	0.17	0.27	**0.17**
User 2	0.06	0.08	0.13	0.24	0.30	**0.16**
User 3	0.07	0.08	0.11	0.13	0.14	**0.11**
User 4	0.11	0.11	0.15	0.16	0.28	**0.16**
User 5	0.09	0.11	0.13	0.14	0.19	**0.13**
User 6	0.10	0.15	0.16	0.24	0.32	**0.19**
User 7	0.17	0.20	0.24	0.27	0.44	**0.26**
User 8	0.14	0.15	0.22	0.22	0.28	**0.20**
User 9	0.09	0.14	0.15	0.17	0.18	**0.15**
Young Bone						**Average**
User 1	0.10	0.21	0.23	0.26	0.37	**0.23**
User 2	0.11	0.16	0.18	0.23	0.23	**0.18**
User 3	0.30	0.30	0.32	0.35	0.44	**0.34**
User 4	0.15	0.16	0.27	0.29	0.42	**0.26**
User 5	0.20	0.30	0.31	0.32	0.37	**0.30**
User 6	0.16	0.20	0.22	0.22	0.25	**0.21**
User 7	0.35	0.38	0.38	0.39	0.40	**0.38**
User 8	0.27	0.27	0.27	0.28	0.38	**0.29**
User 9	0.19	0.28	0.40	N/A	N/A	**0.29**

TABLE 1. RMS Error Values for each of the users' trials for both bone variants, when compared to the bone specific Expert Models

6 Discussion

In this section, we highlight the key limitations of our work and discuss the challenges faced in this paper.

6.1 Limitations

Our drilling signature successfully helps distinguish between the overall bone-drilling behaviour of a non-expert user, with respect to an expert surgeon. However, one of the fundamental limitations of our metric is the lack of a ground truth reference to evaluate the efficacy of our metric. In this paper, we benchmark the drilling performance of an expert surgeon as our reference comparison, although, the presence of a ground truth would have helped form a fundamental basis towards the development of novel user performance metrics. Furthermore, we only had one expert surgeon as our user performance benchmark. This limitation is brought by the varying level of expertise across orthopedic surgeons, also, how many years of bone-drilling experience is sufficient to help compute our expert model. Recruiting more expert surgeons would be one of our immediate future goals to strengthen our evaluation assessment. On similar lines, there is a need to collect more user trials from novice users, as well as, resident surgeons across different years of training. This will not only help us understand the distinction between a novice (firs year resident) and expert user, also, compare training progress with orthopedic residents in their second and later years of residency..

We also experienced long drilling signature computation times (approx 15 - 20 minutes) for a given user trial. One of the primary reasons for this is the relatively longer drilling duration, resulting in larger trajectory points (¿ $10,000$). This limitation can also be attributed to our threshold parameter m that ensures a stead state for a given drilling trajectory. There is a need for efficient recording of user drilling data and a threshold parameter to speed up the signature computation process. In addition to signature plots, we observed the histogram plots for the drilling signature also to show distinctive distributions for different user expertise and bone variant. One of our near goals is to explore the histogram plots and compare its efficacy with respect to the expert model based on signature plots.

6.2 Multi-modal Data

We record and process different types of data such as drilling force, drill tip position, and drilling speed for a bone-drilling activity. While the focus of this work primarily emanates from the drilling trajectory, our preliminary analysis of other parameters have shown observable differences in the force and speed profiles of expert and novice users. We believe that an individual's drilling behaviour is not dependent on their physical motion, but also, in the way they apply force while drilling, also, the rate at which they drill through a solid object based on the material resistance; which in our context is a $3D$-printed bone surrogate. Since force and speed data emulate signals across a time duration, we could extend our signature approach to characterize drilling behaviour based on data other than the drilling trajectory. The collective analysis of different physical parameters of a bone drilling task could help us better understand parameters such as force, which isn't as comprehensible as spatial motion or drilling speed.

6.3 Bone Materials

In this paper, we conducted our user evaluation based on two bone variants — *Young* and *Osteoporotic*, emulating perceptual and physical properties of bones across young and old age groups. $3D$-printing the bone surrogates makes it easier to design, manufacture, and improvise upon the material properties. Similar to our metric, the current bone models are experimental and constantly improved based on expert surgeon feedback to match the material properties of an actual bone. Our comparison of novice vs. expert users shows inconsistent and erratic signature plots for the Young Bone variant trials. We believe that user performance was affected by material hardness of the Young Bone surrogate, and designing a bone with varying intermediate hardness between the Young and Osteoporotic bone hardness could provide a new way to train orthopedic residents. The idea is that novice users learn to drill from softer to harder bones or the opposite way, so as to develop fine motor control crucial for patient safety in orthopedic surgery.

7 Conclusion & Future Directions

In this paper, we showcase the possibility to objectively evaluate and asses drilling performances across user groups of varied expertise. The preliminary success of the drilling signature lays down numerous possibilities of exploring and rethinking different bone-drilling parameters towards categorization and classification of user behavior across expertise, as well as, bone-variants. This highlights the possibility of quantifying drilling performance from the point of improving orthopedic resident training. Our goal in the future is to continue collecting and analyzing drilling data across users, and design robust metrics to distinguish drilling behavior of a novice from an expert surgeon.

ACKNOWLEDGMENT

This work was partially supported by Project No. 0014566A funded by the American Board of Orthopaedic Surgeons (ABOS) and the Orthopaedic Research and Education Foundation (OREF) funded by ABOS. We also thank Dr. James F. Kellam and Dr. Catherine G. Ambrose from the Department of Orthopedic Surgery at The UT Health Science Center for their continued collaboration, and support towards the progress of our work.

REFERENCES

[1] Lee, J., Chavez, C. L., and Park, J., 2018. "Parameters affecting mechanical and thermal responses in bone drilling: A review". *Journal of Biomechanics,* ***71***, pp. 4–21.

[2] Lin, Y., Wang, X., Wu, F., Chen, X., Wang, C., and Shen, G., 2014. "Development and validation of a surgical training simulator with haptic feedback for learning bone-sawing skill". *Journal of biomedical informatics,* ***48***, pp. 122–129.

[3] Akhtar, K., Chen, A., Standfield, N., and Gupte, C., 2014. "The role of simulation in developing surgical skills". *Current reviews in musculoskeletal medicine,* ***7***(2), pp. 155–160.

[4] Madan, S. S., and Pai, D. R., 2014. "Role of simulation in arthroscopy training". *Simulation in Healthcare,* ***9***(2), pp. 127–135.

[5] Chou, Y.-J., Sun, S.-P., and Liu, H.-H., 2011. "Calcaneal osteotomy preoperative planning system with 3d full-sized computer-assisted technology". *Journal of medical systems,* ***35***(5), pp. 755–763.

[6] Wagner, A., Ploder, O., Enislidis, G., Truppe, M., and Ewers, R., 1996. "Image-guided surgery". *International journal of oral and maxillofacial surgery,* ***25***(2), pp. 147–151.

[7] Seymour, N. E., 2008. "Vr to or: a review of the evidence that virtual reality simulation improves operating room performance". *World journal of surgery,* ***32***(2), pp. 182–188.

[8] Zirkle, M., Taplin, M. A., Anthony, R., and Dubrowski, A., 2007. "Objective assessment of temporal bone drilling skills". *Annals of Otology, Rhinology & Laryngology,* ***116***(11), pp. 793–798.

[9] Pandey, R. K., and Panda, S., 2013. "Drilling of bone: A comprehensive review". *Journal of clinical orthopaedics and trauma,* ***4***(1), pp. 15–30.

[10] Timon, C., and Keady, C., 2019. "Thermal osteonecrosis caused by bone drilling in orthopedic surgery: A literature review". *Cureus,* ***11***(7).

[11] Bachus, K. N., Rondina, M. T., and Hutchinson, D. T., 2000. "The effects of drilling force on cortical temperatures and their duration: an in vitro study". *Medical engineering & physics,* ***22***(10), pp. 685–691.

[12] Kazum, E., Dolkart, O., Rosenthal, Y., Sherman, H., Amar, E., Salai, M., Maman, E., and Chechik, O., 2019. "A simple and low-cost drilling simulator for training plunging distance

among orthopedic surgery residents". *Journal of surgical education,* ***76***(1), pp. 281–285.

[13] Hassan, O. E., Amer, M., Abdelsalam, A. K., and Williams, B. W., 2018. "Induction motor broken rotor bar fault detection techniques based on fault signature analysis – a review". *IET Electric Power Applications,* ***12***(7), pp. 895–907.

[14] Miljković, D., 2015. "Brief review of motor current signature analysis". *CrSNDT Journal,* ***5***, 06, pp. 14–26.

[15] El Hachemi Benbouzid, M., 2000. "A review of induction motors signature analysis as a medium for faults detection". *IEEE Transactions on Industrial Electronics,* ***47***(5), pp. 984–993.

[16] Sergent, C., Corazzol, M., Labouret, G., Stockart, F., Wexler, M., King, J.-R., Meyniel, F., and Pressnitzer, D., 2021. "Bifurcation in brain dynamics reveals a signature of conscious processing independent of report". *Nature communications,* ***12***(1), February, p. 1149.

[17] Alleon, J., Montagnac, G., Reynard, B., Brulé, T., Thoury, M., and Gueriau, P. "Pushing raman spectroscopy over the edge: purported signatures of organic molecules in fossil animals are instrumental artefacts". *BioEssays,* ***n/a***(n/a), p. 2000295.

[18] El-ghazal, A., Basir, O., and Belkasim, S., 2009. "Farthest point distance: A new shape signature for fourier descriptors". *Signal Processing: Image Communication,* ***24***(7), pp. 572–586.

[19] Chevyrev, I., and Kormilitzin, A., 2016. "A primer on the signature method in machine learning". *arXiv preprint arXiv:1603.03788*.

[20] Arribas, I. P., Goodwin, G. M., Geddes, J. R., Lyons, T., and Saunders, K. E., 2018. "A signature-based machine learning model for distinguishing bipolar disorder and borderline personality disorder". *Translational psychiatry,* ***8***(1), pp. 1–7.

[21] Kormilitzin, A., Saunders, K., Harrison, P., Geddes, J., and Lyons, T., 2016. "Application of the signature method to pattern recognition in the cequel clinical trial". *arXiv preprint arXiv:1606.02074*.

[22] Moore, P., Lyons, T., Gallacher, J., and Initiative, A. D. N., 2019. "Using path signatures to predict a diagnosis of alzheimer's disease". *PloS one,* ***14***(9), p. e0222212.

[23] Lai, R., Shi, Y., Dinov, I., Chan, T. F., and Toga, A. W., 2009. "Laplace-beltrami nodal counts: A new signature for 3d shape analysis". In 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 694–697.

[24] Aubry, M., Schlickewei, U., and Cremers, D., 2011. "The wave kernel signature: A quantum mechanical approach to shape analysis". In 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops), pp. 1626–1633.

[25] Reuter, M., Wolter, F.-E., and Peinecke, N., 2006. "Laplace–beltrami spectra as 'shape-dna' of surfaces and solids". *Computer-Aided Design,* ***38***(4), pp. 342–366. Symposium on Solid and Physical Modeling 2005.

[26] Andreux, M., Rodolà, E., Aubry, M., and Cremers, D., 2015. "Anisotropic laplace-beltrami operators for shape analysis". In Computer Vision - ECCV 2014 Workshops, L. Agapito, M. M. Bronstein, and C. Rother, eds., Springer International Publishing, pp. 299–312.

[27] Zhang, D., wu, z., Wang, X., Lv, C., and Zhou, M., 2021. "3d non-rigid shape similarity measure based on fréchet distance between spectral distance distribution curve". *Multimedia Tools and Applications,* ***80***, 01, pp. 1–26.

[28] Field, D. A., 1988. "Laplacian smoothing and delaunay triangulations". *Communications in applied numerical methods,* ***4***(6), pp. 709–712.

[29] Vollmer, J., Mencl, R., and Mueller, H., 1999. "Improved laplacian smoothing of noisy surface meshes". In Computer graphics forum, Vol. 18, Wiley Online Library, pp. 131–138.

[30] Hansen, G., Douglass, R., and Zardecki, A., 2005. *Mesh Enhancement: Selected Elliptic Methods, Foundations And Applications*. 03.

[31] Sorkine, O., Cohen-Or, D., Lipman, Y., Alexa, M., Rössl, C., and Seidel, H.-P., 2004. "Laplacian surface editing". In Proceedings of the 2004 Eurographics/ACM SIGGRAPH Symposium on Geometry Processing, SGP '04, Association for Computing Machinery, p. 175–184.

[32] Herrmann, L. R., 1976. "Laplacian-isoparametric grid generation scheme". *Journal of Engineering Mechanics-asce,* ***102***, pp. 749–907.