Fast Vibration-Based Autonomous Fault Detection for Extraterrestrial Drills

Elsa Forberger*, Sarah Boelter*, Brian Glass† and Maria Gini*
* University of Minnesota Twin Cities, Minneapolis, Minnesota
†NASA Ames Research Center, Moffett Field, California

***Abstract*—Drilling is essential to characterize the subsurface of extraterrestrial bodies and to gather samples that enable further understanding and exploration of our solar system. However, extraterrestrial drills are prone to structural failure due to the extremities in environments beyond earth, and require robust autonomous operation to prevent structural damage and achieve mission objectives. This is challenging to accomplish with conventional autonomy methods in a low resource planetary environment. While previous work utilizing structural health management (SHM) techniques excelled at detecting most oncoming faults, it required creation and validation of complex dynamic models, operation of seven neural nets in parallel, nor did it address the percussive behavior seen in modern lunar drills or provide diagnostics quickly enough to catch faults that onset in under 20 seconds. Rather than seeking to create a new resource intensive and complex learner, we seek to create lightweight models to run alongside existing systems. This work enables the use of previous SHM research by providing an unsupervised ensemble-learning method to identify percussive beats in drill vibration applicable to computation and energy constrained space environments, and shows that the frequencies present during percussive strikes can be used to perform high-frequency health diagnostics.**

I. Introduction

The Regolith and Ice Drill for Exploring New Terrain (TRIDENT) [11] is a rotary percussive drill manufactured by Honeybee Robotics used to collect samples from up to 1 meter below the lunar surface. TRIDENT's goal is to characterize the water, minerals, and subsurface structure found in the lunar substrate [23]. Drills like TRIDENT are essential to study extraterrestrial bodies and there is increasing need for their autonomous operation as space exploration ventures beyond the moon to more distant extraterrestrial bodies. Drilling in extraterrestrial environments presents unique challenges of design and operation when compared to terrestrial drills, making them more prone to structural damage [20]. On Earth, fluid lubricants and gravity assist in the smooth penetration of rock and ejection of cuttings from the drill bit onto the surface, and prior knowledge of the subsurface structure can be obtained through past boreholes or seismic imaging [10][17]. In space, the low pressure and extreme environment requires dry drilling and consideration of how to move the cuttings to the surface, nor is prior knowledge of the substrate available.

Despite these challenges, TRIDENT must schedule tasks like percussion, cutting removal, and bit cooling while adjusting rotation speed and percussive power on the fly [23]. Identifying oncoming faults in time to react is crucial to avoid loss of equipment. But, the lightspeed delay makes human teleoperation from Earth impossible beyond the moon, and constant oversight of lunar robots is not feasible as the number of robots grows [18]. Therefore, it is necessary to develop autonomous systems that can identify when the drill is approaching a fault and take corrective action or retreat to a fail-safe state without human intervention, while operating in and adapting to unknown conditions [3][10]

But, computational resources and power are highly limited in extraterrestrial environments [20]. The NASA Space Technology Mission Directorate [7] lists reliable power generation and the development of high performance computing hardware within the top three civil space shortfalls. Without the infrastructure of earth, conventional AI/ML methods will strain the resources of a lander or rover, making them ineffective in a planetary environment. Computing power in space is further limited by radiation - radiation-hardened processors lag significantly behind their commercial counterparts in terms of processing power [19].

In this work we perform structural health monitoring of the TRIDENT drill to predict incoming faults. A lightweight ensemble learning system is used to find and segment the percussive beats from drill data, after which a neural net analyzes the top four frequency peaks and magnitudes of each beat to provide high-frequency diagnostics of drill health.

II. Project Background

Early work on autonomous space drilling comes from 2008 Drilling Automation for Mars Exploration (DAME) project [10]. DAME was a rotary, non-percussive drill that used three fault detection systems in parallel: SHM via a vibration based neural net, a telemetry based model, and a telemetry based rule system, which established a trend of using multiple fault detection systems in parallel. Data collected with DAME resulted in the categorization of faults into six types. Table I illustrates the types of faults encountered during drilling.

TABLE I
Fault Types for planetary drilling from Statham [20]

Fault Type	Description
Binding Fault	Increased torque to friction on drill string
Choking Fault	Cuttings caught in borehole increasing torque
Hard-Materials Fault	Stalled ROP with increased torque
Corkscrewing Fault	Flutes caught on protruding rock
Bit Inclusion	Gravel caught in drill flutes increasing torque

TRIDENT unlike DAME, utilizes rotary-percussive drilling which excels at pulverizing rock while minimizing weight on bit [12]. During 2022 field tests at the Honeybee Robotics facility, TRIDENT drilled into buckets of concrete or clay to find the percussive energy required to fracture samples of different materials. During these tests, personnel noted that a "hollow" percussive noise was found to precede a choking fault. The build-up of compacted dust in the bit flutes prevented the drill bit from moving up and down in the borehole, which prevented the percussion hammer from striking the top of the bit. Instead, the hammer hit the hard stop around the top of the bit, which had lower dampening and produced a hollow noise. The distinctness and reliability of this noise in indicating a fault led researchers to look at the vibration-based system used on DAME in 2008, and the Structural Health Monitoring (SHM) techniques it built on.

While autonomy development on DAME utilized models of nominal and off-nominal behavior, recent work on TRIDENT has focused on identifying changes anomalies in telemetry without using dynamic models [4]. Models can be complex, difficult to adapt, and require additional resources and data.

III. Related Work

A. Structural Health Management

Structural health monitoring (SHM) is an interdisciplinary field where structure stimuli responses are assessed to predict and quantify damage [9]. Use of SHM techniques in extraterrestrial drilling was first carried out by Statham [20] in the DAME project [10]. Statham performed online assessment of the ambient excitations observed from drill operations by passing the natural frequencies of the drill string into seven neural networks (one for nominal operation, and one each fault type) to get the probability that the drill was in that respective operational mode [21][20]. Each NN took about five minutes to train and consisted of $3x1$ input layer, two $20x1$ hidden layers, and a $2x1$ output layer. Data was recorded using multiple laser-doppler vibrometers [2][22] and harmonic excitations from the drill gearbox were filtered out before it was used for inference. The neural networks were trained on model parameters and associated natural frequencies generated from experimentally validated dynamic models of nominal and faulting operation. Without these models, relationships between drill health and frequencies would be difficult to determine, as the frequencies were also determined by auger depth, angular speed, etc. Statham's system, along with the model and rule-based systems enabled the drill to successfully operate during three field tests, each spanning at least 32 hours of autonomous operation [10],[20].

The inference rate of the SHM system was roughly three times a minute, which failed to detect instantaneous choking faults [20]. For the third field test of the SHM system, the team implemented a boundary condition (BC) system which flagged telemetry values exceeding set thresholds, reducing false positives and improving overall reliability by enabling identification of quick-onset faults.

B. Beat Identification

Beat identification falls under music analysis. The key challenges involve beats with soft onsets, non-percussive noise, and tempo changes. Physical properties like reverberations and recording artifacts also degrade identification accuracy [15]. Traditional model-based identification methods consist of two steps: Using input signal to calculate a novelty curve representing the changing state of the signal, and curve point picking [6] of believed valid beats. The accuracy of model-based methods depends heavily on the novelty curve type and peak picking scheme used. Rapid and very dynamic tempo changes can cause decreased accuracy even when using tempo-aware methods[14]. Some novelty curve methods struggle to capture beat onsets when beats do not have a hard start, or when non-percussive noise is loud enough to reduce beat prominence. The alternative to model-based methods is deep learning (DL), which excels with large training datasets [5].

IV. Methods

Although Statham's work proves SHM is a viable area for fault detection, the slow inference rate and subsequent utility of the BC system highlight the need for faster diagnostics [20]. To this date, no fault detection system for planetary drills has looked at percussion vibration, despite observations that it indicates faults and suitability for high-frequency diagnostics. TRIDENT percusses at 13.5 Hz, meaning as long as less than 270 percussive blows are required for each diagnostic, it will have a higher update rate than the DAME SHM system.

Our goal is to produce a computationally lightweight, robust method for identifying percussion from drill vibration data, and a system for identifying faults based on percussive beat spectral characteristics. This address the research gap around drill percussion in SHM and provides a high-frequency fault detection system to augment past work. The chosen methods account for the data sparsity in field robotics and the limited resources for development and operation. Available vibration data from the 2022 TRIDENT tests are 17 videos, each ranging from 3 to 20 minutes. The videos were not recorded with percussion vibration in mind, and no steps were taken to reduce background noise, characterize the microphone response curve, minimize reverberation, remove gearbox harmonics from recorded data, or record room tone.

A. Ensemble Learning for Beat Identification

Ensemble learning is an approach where results from multiple base learners are combined to produce a system more accurate than any component learners. Ensemble learning can use less data and computational requirements compared a single highly accurate model, as each base learner can be less accurate and still produce a comparable overall accuracy. Using ensemble learning for fault detection in space is not new [1], as a diverse spread of base learners means there is no single point of failure.

A lack of labeled data and the goal of minimizing deployment requirements limits us to unsupervised methods that require minimal training data. However, developing base

learners that reach a high level of accuracy under power and computational limits is challenging. While beat identification in music is well studied, identifying drill percussion is unexplored and challenging territory. Noise in the TRIDENT percussion dataset is far beyond what is seen in music and varies by file. Additionally, beat spectral qualities vary between nominal and faulting operation. TRIDENT's changing percussion tempo also informs the selection of base learners. TRIDENT typically starts percussion at 6 Hz and increases to 13.5 Hz within one second [23]. Therefore, base learners must track beats through rapid tempo changes.

TABLE II
BASE LEARNER METHODS REFERENCED IN THIS PAPER

	Novelty Curve	Peak Picking	Method Name
1	Spectral Flux	Threshold	$\Delta + \beta$
2	Spectral Flux	Clustering	$\Delta + GMM$
3	Predominant Local Pulse	PLPDP	$\Gamma + PLPDP$

After exploration, three were selected for noise robustness and adaptation to tempo changes: one basic non-ML method ($\Delta + \beta$), and two unsupervised methods ($\Delta + GMM$ and $\Gamma + PLPDP$), listed in Table II.

Each base method produces proposed beat timestamps. The ensemble learner determines which proposed beats from each base learner correspond to one another by greedily picking groups that minimize the duration between all beats in the group, with a maximum of one beat from each method. An upper limit of 0.0185 seconds is used as a cutoff, which is $0.25 * T_{fmax}$, where $T_{fmax} = 1.0/13.5Hz$. If beats are more than 0.0185 seconds away, they are assumed to be in separate groups even if they are the closest neighbors. Then, the weighted majority algorithm is applied to each group. If the weighted majority score of the group is over 0.5, the average timestamp of all group beats is taken as the final beat location.

1) Spectral Flux & Threshold - $\Delta + \beta$: The first base learner picks peaks by threshold from a spectral flux novelty curve. The spectral flux curve is foundational for other methods, and consists the change in spectral energy between consecutive frames of the signal. [6]. Spectral flux has no tempo concept, unlike more complex novelty curve methods, and is very reactive to highly dynamic signals and any noise within. To find the spectral flux, the audio signal is broken into overlapping frames, each N signals long and offset by a hop of length h. After each frame is weighted with a window to prevent spectral leakage (frequency artifacts resulting from frames not being an integer number of periods long), the representation of frame n in the frequency domain $X(n)$ is obtained with the Discrete Fourier Transform. Given the magnitude spectrogram $X(n,k)$, where n is the frame number and k is the frequency bin index, the spectral flux $\Delta(n)$ is found as follows:

$$\Delta(n) = median_{k=1}^{k=\frac{N}{2}} H(|X(n,k)| - |X(n-l,k)|) \quad (1)$$

where $H(x) = \frac{x+|x|}{2}$ is the half-wave rectifier function, zeroing all negative values and l is the lag in the number of frames. Conventionally $l = 1$. In most applications, the $H(|X(n,k)|-|X(n-l,k)|)$ terms are summed, but the median function was used so a nonzero value for $\Delta(n)$ only occurs when most frequency bins have a positive magnitude change, reducing the noise showing up as peaks in the spectral flux curve. Using a threshold is a basic form of peak picking [6], and consists of selecting all peaks in a novelty curve with local maxima above a threshold. The threshold β for the spectral flux curve was set to $\beta = 2.0$, and was chosen after inspecting results from various β dataset values.

2) Spectral Flux & Clustering - $\Delta + GMM$: The second base learner also used the spectral flux curve, but selected peaks with a Gaussian Mixture Model (GMM) fit to the results of the $\Delta + \beta$ method. Clustering is not an unexplored technique for peak picking [8][13], although we diverge from previous research by using a more complex metric and GMM instead of K-NN. Gouyon et al. evaluated K-NN clustering for beat identification on music using 172 spectral features and found features that measured energy variation across different frequency bands best produced distinct clusters that corresponded to valid and invalid beats [13]. However, using any single spectral feature for clustering performs poorly in this context due to the high amount of noise obscuring the spectral energy variation of true beats and creating large peaks in the spectral flux curve where no beat occurs. To address this problem, we cluster using a metric designed to produce a low score for noise-based peaks and a high score for true ones. It measures both the novelty curve at the peak and the strength of the immediately surrounding peaks. This method works, firstly, because the drill does not stop/start rapidly, and the average peak height around a valid peak is higher than the average peak height around an invalid one, even with noise. Secondly, despite noise, false peaks and obscured true ones are sparse compared to the presence (or absence of) genuine peaks. Within a window of 5 peaks, it can be assumed that less than half will be the product of noise.

Given a set of peak magnitudes $M = (m_1, m_2, \ldots, m_K)$ in a signal, a window length w, the peak score at index j is:

$$S_j = \frac{\sum_{k=j-w/2}^{k=j+w/2} m_k}{w} + m_j \quad (2)$$

By adding the height of the peak m_j to the average, the effect of invalid peaks on the score of surrounding valid peaks (and the reverse) is minimized, ensuring greater separation in scores between the two groups. We use a GMM over K-NN because the GMM can take advantage of the initial clusters provided by the results of $\Delta + \beta$. Applying the $\Delta + \beta$ method to the training dataset produces groups of valid" and "invalid" peaks, after which average scores for each group are used as initial cluster centers. The $\Delta + GMM$ improves the initial identification from $\Delta + \beta$, and produces two non-intersecting clusters corresponding to two peaks in score values. The statistics for the fitted GMM are in 1 with $w = 7$.

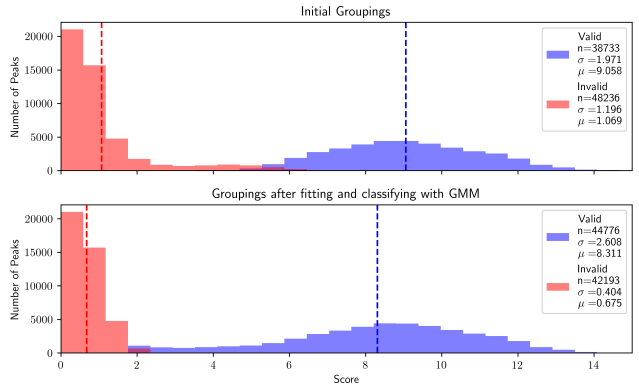

Fig. 1. Distribution of scores from $\Delta + \beta$ (top) and $\Delta + GMM$ (bottom)

3) Predominant Local Pulse & PLPDP - $\Gamma + PLPDP$: Predominant Local Pulse (PLP) is a novelty curve method measuring both strength and local periodicity of peaks in the spectral flux curve [14]. Sinusoidal kernels are applied each time position of an input spectral flux curve. The kernels are aggregated by summing, then half-wave rectification zeros negative sections. Periodic peaks in the spectral flux curve are enforced by sinusoidal kernels, whereas non-periodic peaks become subject to deconstructive interference, resulting in the PLP curve Γ. The ability of the PLP curve to "expect" beats based on beat periodicity in the local signal area allows it to correctly assume beats when the signal is weak or beats are missing, but means the curve takes time to react to sudden changes in tempo and beat strength [5].

Phase and periodicity of each kernel are computed using a tempogram derived from the Short Time Fourrier Transform $\mathcal{F}$ of the input novelty curve (in this case, spectral flux $\Delta(n)$).

$$\mathcal{F}(t,\omega) = \sum_{n\in\mathbb{Z}} \Delta(n) * W(n-t) * e^{-2\pi i\omega n} \tag{3}$$

where $W : \mathbb{Z} \to \mathbb{Z}$ is a windowing function size $2N+1$ for $N \in \mathbb{N}$. The tempogram $\mathcal{T} : [1:T] \times \Theta \to \mathbb{C}$ is defined as

$$\mathcal{T}(t,\tau) = \mathcal{F}(t,\tau/60) \tag{4}$$

where $t \in [1:T]$ is the timestamp, and $\tau \in \Theta$ is a BPM tempo, where $\Theta = Uniform[a:b]$. Dominant tempo $\tau_t \in \Theta$ is found by maximizing $\tau_t := argmax_{\tau\in\Theta}|\mathcal{T}(t,\tau)|$, after which the phase is found using:

$$\varphi_t := \frac{1}{2\pi}\arccos\left(\frac{Re(\mathcal{T}(t,\tau_t))}{|\mathcal{T}(t,\tau_t)|}\right) \tag{5}$$

The sinusoidal kernel κ is then defined as:

$$\kappa_t(n) := W(n-t)\cos(2\pi(n*\tau_t/60 - \varphi_t)) \tag{6}$$

The parameters N and Θ have a strong influence on how good the PLP curve is at tracking percussion. N determines T_W, the duration over which the window function is applied. If T_W is significantly longer than any changes occurring to the tempo, the resulting τ_t values lack resolution, and will not reflect tempo changes occurring within the window. If multiple peaks are not present in the window, the tempo measurement will also be incorrect. The ideal T_W will barely contain two beats from the slowest percussion of the drill.

$$T_W = \frac{1}{f_\Delta}(2*N+1) \tag{7}$$

where $f_\Delta = f_s/h$ is the spectral flux curve sampling rate, h is the hop length between frames and f_s is the original signal sampling rate. Given the slowest percussion rate is 6Hz, $T_W = 0.25$. Therefore, $N = 6$. Θ, the distribution from which the dominant tempo τ_t is found, heavily influences the ability to react to local tempo changes. If Θ does not include just the actual tempo frequencies, the PLP curve peaks will not align with beat onsets. The Θ distribution used is $Uniform[6, 13.5]$ Hz, chosen because the slowest drill percussion rate is 6Hz, and nominally 13.5Hz.

The dynamic programming algorithm "PLP with Dynamic Programming" (PLPDP) is used for peak picking. [5]. PLPDP seeks to identify a series of beats maximizing a cost function rewarding tempo consistency and novelty curve correspondence while preventing penalization for signal tempo changes. This requires conversion of the PLP curve $\Gamma : [1:N] \to [0,1]$ into two continuous piecewise functions expressing local tempo information: inter-beat-interval (IBI) $\hat{\delta}(n)$ and confidence $\lambda(n)$ of the locally detected period. $\hat{\delta}(n)$ and $\lambda(n)$ are calculated using reference beat timestamps. We deem the timestamp beats valid by applying a threshold of 0.15 to the PLP curve Γ. This threshold was obtained by visually analyzing the results of various values across the entire dataset. The PLP curve is then divided at the leftmost minimum of each reference beat to obtain a series of segments. For all n in a segment, $\hat{\delta}(n)$ is set to the IBI. Confidence $\lambda(n)$ is the average height of the segment's peaks. These two parameters are used to compute the score $D(n)$ at frame n, which consists of the PLP value for frame n, the accumulated previous frame scores, and a penalty value of tempo consistency penalties. Dynamic programming is then used to greedily select a sequence of beats maximizing the cumulative score.

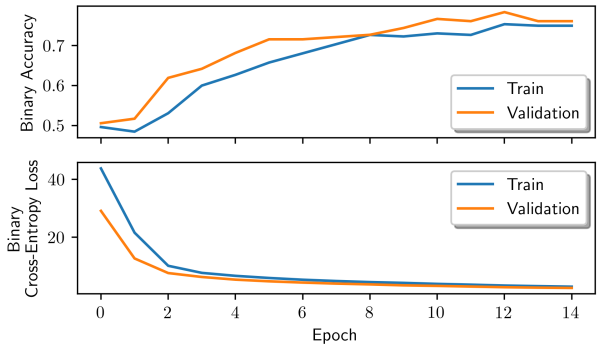

Fig. 2. Accuracy and loss values over time from one of the 23 training runs.

After the locations of beats are identified, each beat is segmented from the audio signal. We initially employed a method similar to Chiu et al. [5] for segmentation to find the start and end of each beat based on the local minima

of the spectral flux curve. We then noted that the average segmented beat duration (μ_d) varied greatly from file to file despite the true duration of beats being the same across files. Noisy audio had an $\mu_d = 0.05$ seconds, whereas quieter files had $\mu_d = 0.03$ seconds. To address this, a distribution-based segmentation method was applied in which the median beat duration and start offset are calculated from the initial novelty-curve segmentation, then used to segment all beats. A comparison can be seen in fig 3. Median is used instead of average because the median values are more consistent with the true beat duration and start offset unless there is extensive signal noise. It was noted the resulting segmentation did not include the beat tail, since the spectral flux curve quickly drops to 0 as the beat dissipates. Because the reverberation after initial impact is a key part of percussion [12] , the median duration was scaled by 1.5 to ensure complete segmentation of each beat.

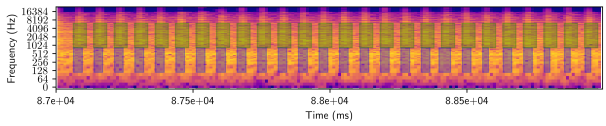

Fig. 3. Results of beat segmentation ontop of the audio spectrogram. Novelty curve segmentation is blue and distribution based segmentation scaled by 1.5 is green.

Each segmented beat is run through the Fast Fourier Transform (FFT). Incoherent averaging is applied using a window size of 5 beats to smooth out unique frequency features of individual beats and highlight common features indicative of nominal or faulting operation. The top four frequency peaks and the four corresponding magnitudes of the averaged signal are the input features for the NN. Figure 4 shows the clear difference in frequency peaks between faulting and non-faulting beats.

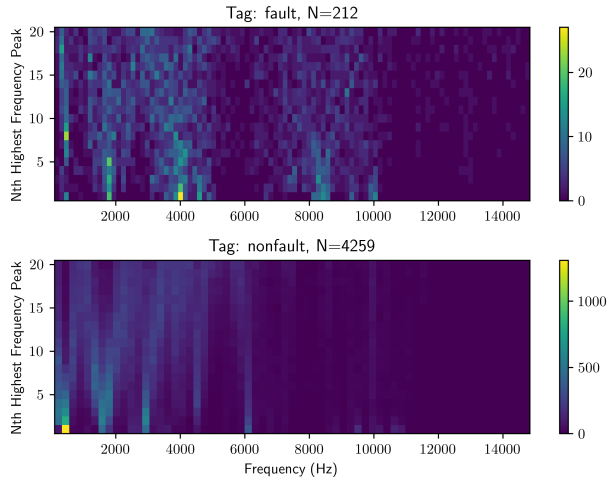

Fig. 4. Locations of top 20 frequency peaks for faulting and non-faulting data.

The NN consists of one $8x1$ input layer, one $4x1$ hidden layer using the ReLU function, and a $1x1$ output layer with a sigmoid activation function categorizing the beat as either "faulting" (0) or "non-faulting" (1). Binary cross entropy is used as the loss function, the optimizer is Adam [16], and binary accuracy is used as an evaluation metric. Data for training fault detection was limited, with only a single percussion file from a choking fault. To prevent the NN from getting stuck learning the faulting to non-faulting data ratio, an equal amount of each beat type were randomly selected from the data pool giving a datset of $N = 434$. 40% of data ($N = 174$) was used for validation, and the rest for training ($N = 260$). Training used a batch size of 5 over 15 epochs.

V. Results

A. Beat identification

The ensemble method accurately identified beats during fault onset and nominal operation, resulting in a more accurate system than any of the base learners. All methods dealt with tempo changes well, but noise degraded performance across the board. It is difficult to say if the worse performance on the faulting data is due to that audio containing more noise, or because of the difference frequency in characteristics. Despite this, the ensemble results are still viable for use in fault detection as there are no false positives, (no risk of acting on a false positive and inducing or exacerbating a fault) and the false negatives are sparse enough to have minimal effect on the neural net diagnostic rate (2.7 Hz in ideal conditions to 2.2 Hz with the actual reliability of the beat identification system). Figure 5 shows results of base and ensemble methods on different data sections.

$\Delta + \beta$ only does well when the percussion is clearly distinguishable from the background noise. It has few false positives, but many false negatives from beats masked by background noise. $\Delta + GMM$ outperforms $\Delta + \beta$ because it factors in strength of the surrounding beats and correctly identifies weak beats $\Delta + \beta$ assumes are noise, but introduces false positives. $\Gamma + PLPDP$ performed equivalently to $\Delta + \beta$, with variations in beat timing.

The beat identification system is extremely lightweight. Fitting the GMM was the most intensive part, and it took on average 7 seconds on a consumer laptop with minimal training data. Once fitted, beat detection can run in real time.

B. Fault Detection

The system was trained using the same parameters 23 times to accurately capture fault detection performance. Figure 2 shows the training process for one of the runs. Table III contains the mean and standard deviation (SD) of final accuracy and loss for 23 different training runs, as well as the mean and standard deviation the changes in those values during training. Since five beats are needed for incoherent averaging to produce input features for the neural net, the maximum diagnostic rate is 2.7 Hz. Like the beat identification system, this system extremely lightweight and can run in real time.

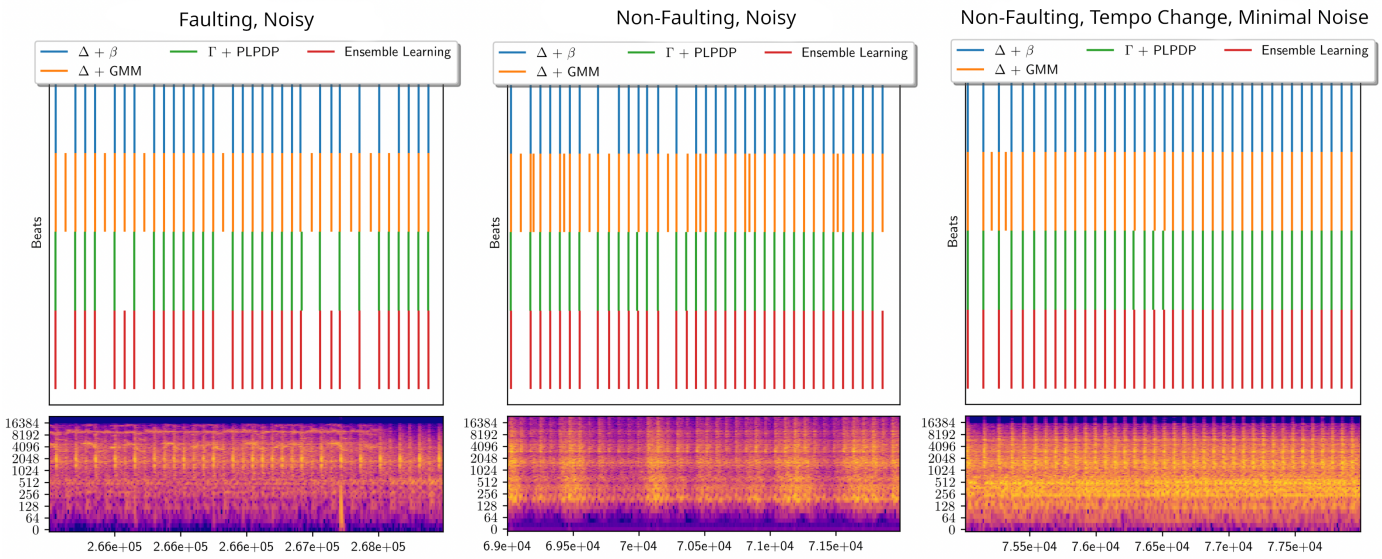

Fig. 5. Left: beat identification methods on faulting, noisy data. Middle: beat identification methods on non-faulting, noisy data. Right: beat identification methods on non-faulting data with minimal noise and a tempo change.

TABLE III
BINARY CROSS ENTROPY LOSS AND BINARY ACCURACY STATISTICS FROM THE TRAINING AND VALIDATION SETS AFTER TRAINING 23 INSTANCES OF THE NN.

Value	**Train**	**Train (Δ)**	**Validation**	**Validation (Δ)**
Loss μ	1.68	-53.04	1.67	-43.07
Loss σ	1.07	49.29	0.83	42.39
Accuracy μ	66.71%	+16.86%	67.03%	+17.58%
Accuracy σ	9.61%	9.43%	9.25%	8.99%

VI. Discussion

A. Beat Identification

The main takeaway from the beat identification efforts is that 1) existing methods in music theory need to be adapted for the noise of drilling and 2) cleaner data is needed. Comparing results on noisy data on the left and right versus the cleaner data in the center in Figure 5 highlights the marked effect noise has on accuracy. We believe noise is why $\Delta + \beta$ and $\Gamma + PLPDP$'s performance on TRIDENT did not match their performance on music. $\Delta + \beta$ sensitivity to noise resulted in higher threshold selection, leading to numerous false negatives. $\Gamma + PLPDP$'s demonstrated robustness to noise and tempo changes on music should have resulted in stellar performance on our noisy data. We believe the threshold used in selecting reference beats from the PLP curve was too high, resulting in few reference beats for $\Gamma + PLPDP$ and a sparse output.

Our custom method $\Delta + GMM$ was developed to deal with the noise in drill percussion, and performed well enough to be used in isolation if not for false positives. Improved scoring schemes, use of multiple parameters instead of a single score, and more accurate methods of determining initial clusters for clustering-based methods should be further researched.

Using basic methods like $\Delta + \beta$ to provide the reference data for methods like $\Gamma + PLPDP$ and $\Delta + GMM$ is a potential route to incorporate more complex base learners without requiring additional data collection and labeling.

B. Fault Detection

Although the fault detection system did not reach the desired accuracy, it demonstrated that percussive frequencies do indicate a jamming fault and should be further explored. Figure 4 shows clear differences in frequency peaks between faulting and non-faulting beats. But, only 14% of faulting beats and 23% of non-faulting have peaks in densest sections of their heatmaps. Incoherent averaging of additional beats may ensure indicative frequency peaks are are passed to the NN and increase fault detection accuracy. Other areas to explore include refining the selection of frequency peaks, and the use of other spectral parameters as inputs to the NN.

VII. Conclusions and Future Work

Fault detection using drill percussion fills a need for high-frequency fault detection to complement more comprehensive, but low-frequency methods. This work shows that the spectral features of percussion are promising for rapid fault detection of choking faults with a diagnostic rate of 2.7 Hz. While the NN accuracies are low, they are higher than expected given the extremely low amount and quality of training data. With cleaner data and an improved collection process, we expect to see significant increases in the accuracy of our fault detection, and to incorporate additional fault types beyond choking into the system. Overall, this system needs further work, but diagnoses faults in previously unexplored behavior at a significantly higher rate than existing systems and uses minimal computational resources, making it a viable companion alongside more intensive fault detection systems.

ACKNOWLEDGMENTS

Material based upon work supported by the Minnesota Space Grant Consortium, NASA PSTAR Mars Exploring through Analog-site Drilling (MEAD), and National Science Foundation CSGrad4US Fellowship Grant No. 2240197.

REFERENCES

[1] Madeline Anderson, Jeremy Muesing, Kerri Cahoy, and Kenneth Center. Ensemble Learning for Autonomous Onboard Satellite Fault Diagnosis With Validation Tool. In *Proceedings of the AIAA/USU Conference on Small Satellites, Automation - Enabling New Capabilites*, 2024. doi: 10.1109/ICASSP.2007.367318.

[2] Dean Bergman, Brian Glass, Kris Zacny, and Gale Paulsen. *Using Distributed Transfer Function Method (DTFM) for Autonomous Health Monitoring of Interplanetary Drills*. doi: 10.2514/6.2015-4482.

[3] James Blacic, D. Dreesen, and T. Mockler. The 3rd dimension of planetary exploration - Deep subsurface sampling. In *Proceedings of the AIAA Space 2000 Conference*. AIAA, September 2000.

[4] S. Boelter, L. Weber, R. Lenz, B. Glass, and M. Gini. Fault prediction in drilling using subspace analysis techniques. *Intelligent Autonomous Systems 19, Proc. 19th Int'l Conf IAS-19, accepted*, 2025.

[5] Ching-Yu Chiu, Meinard Müller, Matthew E. P. Davies, Alvin Wen-Yu Su, and Yi-Hsuan Yang. Local Periodicity-Based Beat Tracking for Expressive Classical Piano Music. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 31:2824–2835, 2023. doi: 10.1109/TASLP.2023.3297956.

[6] N. M. Collins. A Comparison of Sound Onset Detection Algorithms with Emphasis on Psychoacoustically Motivated Detection Functions. *Journal of The Audio Engineering Society*, 2005.

[7] NASA's Space Technology Mission Directorate. Stakeholder Webinar: Civil Space Shortfall Ranking, 2024.

[8] Simon Dixon. A Lightweight Multi-agent Musical Beat Tracking System. volume 1886, 10 2000. ISBN 978-3-540-67925-7. doi: 10.1007/3-540-44533-1_77.

[9] Vahidreza Gharehbaghi, Ehsan Farsangi, Mohammad Noori, Tony Yang, Shaofan Li, Andy Nguyen, Christian Málaga-Chuquitaype, Paolo Gardoni, and Seyedali Mirjalili. A Critical Review on Structural Health Monitoring: Definitions, Methods, and Perspectives. *Archives of Computational Methods in Engineering*, 29, 10 2021. doi: 10.1007/s11831-021-09665-9.

[10] B. Glass, H. Cannon, M. Branson, S. Hanagud, and G. Paulsen. DAME: Planetary-Prototype Drilling Automation. *Astrobiology*, 8(3):653–664, June 2008. doi: 10.1089/ast.2007.0148.

[11] B. Glass, D. Bergman, V. Parro, L. Kobayashi, C. Stoker, R. Quinn, A. Davila, P. Willis, W. Brinckerhoff, K. Warren-Rhodes, M.B. Wilhelm, L. Caceres, J. DiRuggiero, K. Zacny, M. Moreno-Paz, A. Dave, S. Seitz, A. Grubisic, M. Castillo, R. Bonaccorsi, and the ARADS Team. The Atacama Rover Astrobiology Drilling Studies (ARADS) Project. *Astrobiology*, 23: 1245–1258, December 2023.

[12] Samuel Goldman, Henryk Flashner, and Bingen Yang. An Approach to Modeling Percussive Drilling Systems. *Journal of Vibration and Acoustics*, 144(5):051015, 05 2022. ISSN 1048-9002. doi: 10.1115/1.4054472.

[13] Fabien Gouyon, Simon Dixon, and Gerhard Widmer. Evaluating Low-Level Features for Beat Classification and Tracking. In *IEEE Int'l Conf. on Acoustics, Speech and Signal Processing - ICASSP '07*, volume 4, pages IV–1309–IV–1312, 2007. doi: 10.1109/ICASSP.2007.367318.

[14] Peter Grosche and Meinard Müller. Extracting Predominant Local Pulse Information From Music Recordings. *IEEE Transactions on Audio, Speech and Language Processing*, 19:1688–1701, 2011. ISSN 15587924. doi: 10.1109/TASL.2010.2096216.

[15] Peter Grosche, Meinard Müller, and Craig Sapp. What Makes Beat Tracking Difficult? A Case Study on Chopin Mazurkas. pages 649–654, 01 2010.

[16] Diederik P. Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization. *CoRR*, abs/1412.6980, 2014.

[17] Raymond Leung, Andrew J. Hill, and Arman Melkumyan. Automation and Artificial Intelligence Technology in Surface Mining: A Brief Introduction to Open-Pit Operations in the Pilbara. *IEEE Robotics & Automation Magazine*, pages 2–21, 2023. ISSN 1558-223X. doi: 10.1109/MRA.2023.3328457.

[18] N. Muscettola, P. P. Nayak, B. Pell, and B. C. Williams. Remote Agent: to boldly go where no AI system has gone before. *Artificial Intelligence*, 103(1–2):5–47, 1998.

[19] Engineering National Academies of Sciences and Medicine. *Testing at the Speed of Light: The State of U.S. Electronic Parts Space Radiation Testing Infrastructure*. The National Academies Press, Washington, DC, 2018. ISBN 978-0-309-47079-7. doi: 10.17226/24993.

[20] Shannon Statham. *Autonomous Structural Health Monitoring Technique for Interplanetary Drilling Applications Using Laser Doppler Velocimeters*. PhD thesis, Georgia Institute of Technology, 2011.

[21] Shannon Statham, Sathya Hanagud, Massimo Ruzzene, Vinod Sharma, Brian J. Glass, and Howard Cannon. Design and validation of an ldv-based structural health monitoring in "drilling automation for Mars exploration" . In *48th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, volume 7, 2007. ISBN 1563478927. doi: 10.2514/6.2007-2304.

[22] Shannon Statham, Sathya Hanagud, Vinod Sharma, and Brian Glass. Three-Dimensional Structural Health Monitoring in Space Applications Using Laser Doppler Vibrometers. American Institute of Aeronautics and Astronautics (AIAA), 4 2008. doi: 10.2514/6.2008-2088.

[23] Kris Zacny et al. Development of TRIDENT Drill for Ice Mining on the Moon with NASA PRIME-1 and VIPER Missions .