HUMS2025 Data Challenge Result Summary

Team Name: LMCO

Team Members: Nathaniel Rigoni, Gabriel Ruiz

Institutions: Lockheed Martin Corporation, © 2024

Publishable: Yes, Publication Permission Granted

1. Summary of Findings

We employed an autoencoder neural network, optimized for multiplexed time series compression and reconstruction, to analyze the raw data. By examining the latent space of the autoencoder, we uncovered differences between sequence points and characteristic frequencies across the system's test. Notably, the variations in the latent space groupings reflected distinct behavioral patterns within the system. Through comparisons of characteristic frequencies in these groups, we identified and defined a spectra shift, which we propose is indicative of the associated fault observed during testing.

Each shift in the latent space corresponds to an underlying change in characteristic frequencies detected by the model, attributable to system changes. By identifying the most significant statistical change in the spectra, based on grouped areas of the latent space, we pinpointed the earliest statistically detectable change in the data. Our key findings are summarized in the following table.

Table 1: Summary of Findings

#	Detection & Trending	Data file name/number	Comments
1	Detection of behavioral shift into first statistically relevant group (group 3)	Day007_20240409_115526_100%TT	First sign of behavioral shift
2	Detection of sustained shift into statistically relevant group (group 3)	Day011_20240424_101221_100%TT	Sustained behavioral shift
3	Detection of behavioral shift into second statistically relevant group (Group 5)	Day014_20240507_133354_126%TT	First sign of behavioral shift
4	Confirmed trend of accelerated fault	Day014_20240507_133354_126%TT	Sustained behavioral shift

2. Description of Analysis Methods

We propose a data-driven approach, leveraging state-of-the-art neural network techniques, to analyze statistically significant information in the dataset related to fatigue and crack propagation in the gear box housing. Our method combines behavioral analysis with clustering to identify shifts in system behavior over time. These shifts result from events like crack propagation, which alter the system's characteristic geometry, mass, and mechanics from one cycle to another.

To capture these changes, we utilize an autoencoder to encode the characteristic frequencies of sequences of readings. This allows us to examine how these frequencies change over time and pinpoint times where there are larger shifts in these frequencies.

To train a behavioral model for this data challenge, we considered two key pieces of information: the sequence length representative of a behavior and the removal of gear-related frequencies. We evaluated the sample rate and dataset length to determine an optimal sequence length. Data was sampled at 65573.77049180328 Hz in 30 second bursts every 4 minutes. This gives us 1,967,213 points per file. Since the data was not continuous, we chose a sequence length (85531 points) that is a multiple of the total points in each file and larger than the sample rate. This enables us to relate behaviors to a period greater than one second and gives us 23 sequences per file.

To eliminate gear-related characteristics, we applied the Fourier Transform to break down sequences into their characteristic frequencies. We then removed frequencies above the mean magnitude, which eliminated gear mesh frequencies and larger magnitude frequencies introduced by our sequence length selection. This step isolates the frequencies most likely to exhibit behaviors related to changes in the gear box housing.

The augmented frequency domain data is used to train the autoencoder, which compresses and decompresses the frequency information to learn a latent space representation. This latent space is the core of our analysis, where points with similar characteristic frequencies at similar magnitudes are closer together. The latent space is defined by data similarity in the context of the problem, enabling us to identify patterns and shifts in system behavior.

From this latent space we will identify characteristic frequency shifts. Then identify files associated with statistically significant shifts in those characteristic frequencies. These statistical shifts will be considered our early detection method.

3. Key Fault Characteristics for Early Detection

In our test case, we aim to define the P-value of a P-F Curve, which represents the point before the failure state F, allowing us to intervene with maintenance and prevent catastrophic incidents. Since our case involves fatigue, we must examine the causal effects of fatigue on the system. Specifically, we are interested in the behavioral shifts that occur as a system fatigues. The vibrational behavior of a system changes as it fatigues, and our latent space analysis enables us to detect these subtle shifts in behavior.

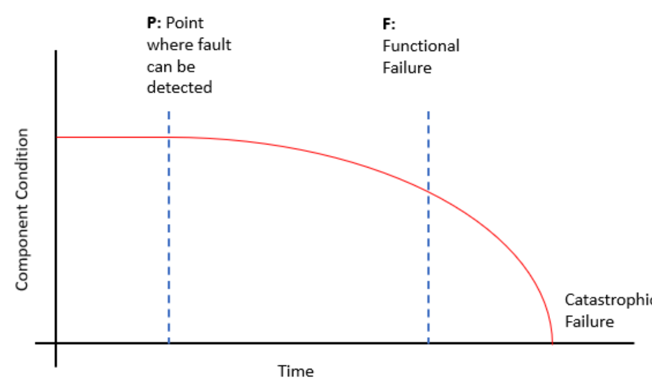

Figure 1: Notional PF Curve.

The latent space representation learned from the model (Figure 2) reveals at least two distinct groups. The two we will focus on at first will be points that occur below x=-1 and points that occur above x=-1. The coloration of this plot shows that the behaviors learned by the model are aligned with the time domain, despite the model not receiving any temporal information during training. This indicates that the behaviors learned are progressive and continuous along the time domain, as shown by the progression of color from purple (early points in time) to gold (later points in time).

A notable gap between x=-1 and x=-2 (Figure 2 left, red line), containing only a few points, suggests a rapid shift in behavior in the system. Given that points in proximity represent similar behaviors, this gap likely indicates the functional failure event. This also suggests that shifts prior to this point are smaller and more indicative of fatigue leading to this rapid shift. We define the area between -1>=x>=-2 as our transitional event.

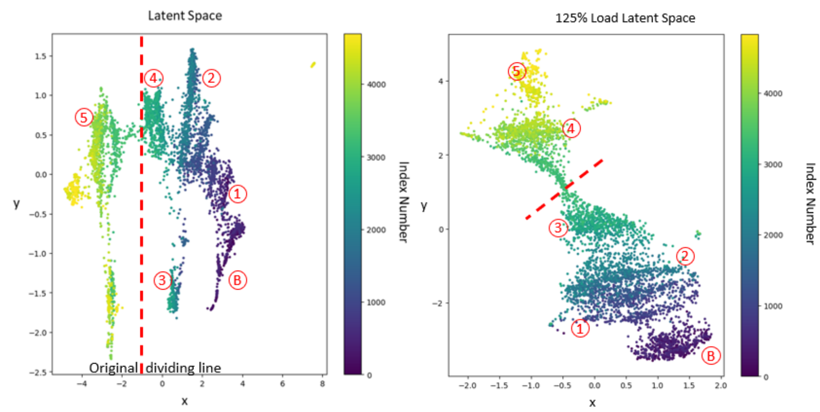

Figure 2: Annotated Latent Spaces. 100% load (left) and >100% load (right).

To analyze different areas of the latent space, we divide the points into groups, based location x<-1 and x>-1. We then compare the characteristic frequencies of these points by averaging the magnitude at each frequency for all points in a group. By subtracting the averaged spectra of one group from the other (Figure 3), we can identify differences in frequency behavior. Our results show that the characteristic change between the two groups occurs in the frequency range of 13kHz to 23kHz. This is evident in Figure 3, which we will use as a reference to compare groups from the latent space.

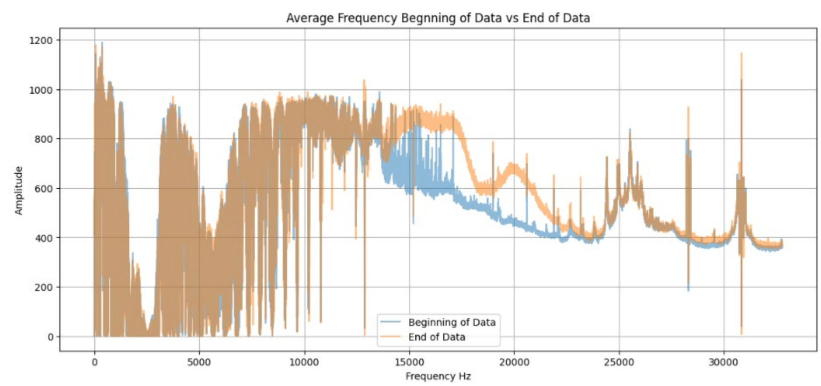

Figure 3: Difference in average characteristic frequencies, early data point vs later data points.

This difference represents the major distinction in characteristic frequency behavior between the beginning and end of the data. We will focus on this frequency range as an area of interest in further analysis. Notably, our analysis of the 100% load file indicates that the largest sustained transitional point occurs on Day014_20240507_091907_100%TT.

To validate our findings from the 100% load data, we applied the same analysis process to the 125% raw data (Figure 2 right, red line). Our results show that the model successfully segregated this data into two main groups, below y=1 and above y=1, like the 100% load data. Furthermore, the temporal relationship of the data is preserved, with behavioral shifts occurring progressively from the first file to the last data point.

The Fourier plot for the 125% data reveals the same relationship at the same frequencies as the 100% load data, providing further evidence of the consistency of our findings. Notably, a sustained behavioral shift is observed starting at file Day014_20240507_133354_126%TT, which corroborates our previous conclusion that the event occurred on Day 14. This validation exercise demonstrates that the observed behaviors are not dependent on the load applied to the box, increasing our confidence in the accuracy of our results.

4. Fault Progression Trending Curve

To analyze the progression of fatigue during the test, we will utilize the difference of average spectra plot to identify differences between points from specific areas on the latent space plot. This involves selecting two x, y coordinates and sampling 100 examples nearest to those points in the latent space to plot their average spectra difference.

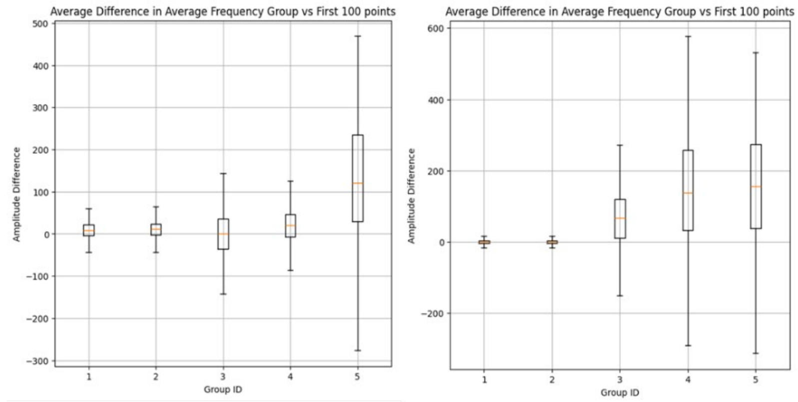

Figure 4: Statistical trending of differences from baseline. 100% load data (left) and >100% load data (right).

The boxplots in Figure 4 illustrates the difference in average spectra between 13kHz and 23kHz when comparing spectra between the numbered point in the latent space (Figure 2) and the baseline point (the first 100 points in the data). The plots reveals that the initial groups (1 and 2) exhibit minimal difference, with only a small deviation in median. However, as we progress to group 3, we observe a significant increase in noise, indicated by the wider whiskers on the box, despite a reduction in median deviation. Groups 4 and 5 display both increased noise and a growing deviation from the median.

Notably, the behavior of the groups in the 100% load data, Figure 4 (left), is consistent with the 125% load data, as shown in Figure 4 (right). The sustained deviation from the baseline begins at group 3, which corresponds to groups 3 and 4 in the 100% data. This validation reinforces our understanding of the fault progression trend.

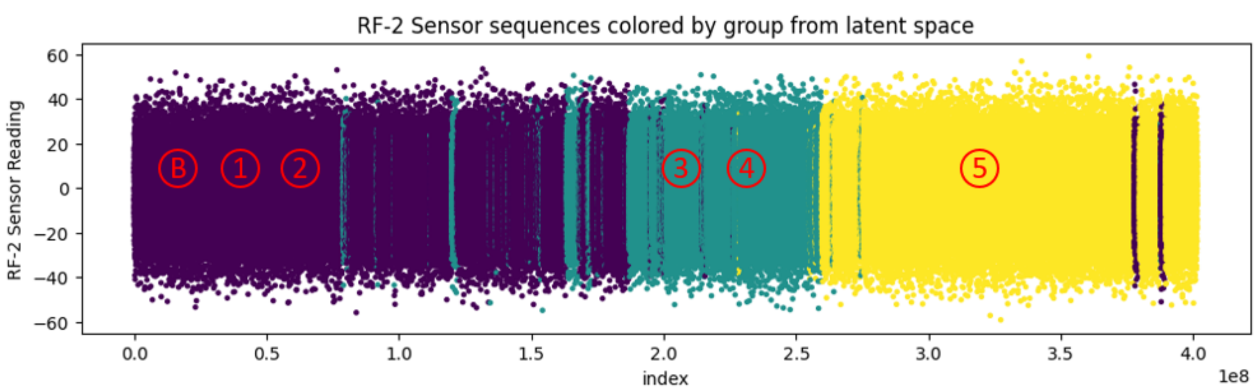

Figure 5: Colorized progression of identified groups in captured readings of sensor RF-2

Our analysis reveals that the earliest indications of behavioral group 3 and 4 can be detected as early as Day 7, specifically in file 'Day007_20240409_115526_100%TT'. This behavior is sustained across multiple files on Day 11, including 'Day011_20240424_101221_100%TT', 'Day011_20240424_102004_100%TT', 'Day011_20240424_113340_101%TT', and 'Day011_20240424_142333_100%TT'. Notably, the machine learning model detects subtle differences in behavior even earlier, as evidenced by the points shifting from the baseline to group 1 in the latent space.

While the shift from baseline to group 1 indicates a distinct difference, we cannot conclusively attribute it to a specific mechanism that is not inherent to the system's operation. However, using statistical box plotting and average spectra comparison, we can confidently assert that the differences between the baseline and groups 3 and above are due to a shift in the box's response in the 13kHz to 23kHz range. This forms the basis of our earliest detection.

By applying the P to F interval analogy, we estimate that the earliest possible detection occurs on Day 7, while the functional failure point is reached on Day 14. Our analysis demonstrates the effectiveness of machine learning and statistical methods in detecting early signs of failure, enabling proactive maintenance, and preventing catastrophic incidents.