Domain-Adaptive ML for Surface Roughness Predictions in Nuclear Fusion

Shashank Galla[1] **Antonios Alexos**[2] **Jay Phil Yoo**[2] **Junze Liu**[2]
Kshitij Bhardwaj[3] **Sean Hayes**[3] **Monika Biener**[3]
Pierre Baldi[2] **Satish Bukkapatnam**[1] **Suhas Bhandarkar**[3]
[1]Texas A&M University
[2]University of California, Irvine
[3]Lawrence Livermore National Lab

Abstract

In Inertial Confinement Fusion (ICF) experiments, achieving high surface quality for hydrogen fuel-filled capsules is critical, requiring a meticulous and time-intensive polishing process. Surface roughness measurements, however, is labor-intensive, time-consuming, and reliant on human operators. To automate this evaluation process, we developed domain-adaptive machine learning model that address the variability in polishing conditions and resulting data distributions. Domain classification techniques were first employed to identify distinct polishing domains, which were then used to create a domain adaptation model for accurate surface roughness prediction across varying conditions. This model enables real-time generation of surface roughness predictions, allowing operators to make adjustments during polishing to achieve optimal results. Our methodology has demonstrated its effectiveness in adapting to different data domains while maintaining consistent performance. Furthermore, we provide physics-based explanations for the emergence of specific domains, enhancing the interpretability of the process.

1 Introduction

Fusion energy offers the highly desired potential of sustainable and virtually limitless source of energy, thus providing a clean and safe alternative to fossil fuels. A significant breakthrough in this field occurred in December 2022 at the National Ignition Facility (NIF), where an energy gain of 1.5 was achieved in an Inertial Confinement Fusion (ICF) experiment [9]. This milestone, which produced 3.15 megajoules (MJ) of fusion energy from 2.05 MJ of laser energy, has been successfully replicated with even greater energy gains [1]. These advancements represent a critical step toward the prospect of development of fusion-powered generators capable of meeting future energy demands sustainably, without producing harmful waste or emissions [11]. In ICF experiments, the quality of the spherical shells, made of high-density carbon (HDC) [3], is crucial for successful fusion reactions [15]. These shells, which contain hydrogen isotopes used as the fusion fuel, must be flawlessly smooth and round to ensure uniform and symmetric implosion during the fusion process [5][12]. Defects on the shell's surface can cause significant instabilities, disrupting the reaction and lowering energy yields [7, 4, 13]. Therefore, the shells undergo an ultra-precision polishing process to achieve surface roughness below 10 nanometers, a requirement that presents significant manufacturing challenges.

The polishing process, however, is highly stochastic, involving the simultaneous processing of up to 20 shells, which increases the risk of surface-damaging anomalies [10]. Traditional methods of measuring surface roughness are time-consuming and labor-intensive, underscoring the need for automatic monitoring techniques that can detect potential defects early and determine the optimal time to terminate the polishing process. Recent advancements in micro-electro-mechanical systems

Machine Learning and the Physical Sciences Workshop, NeurIPS 2024.

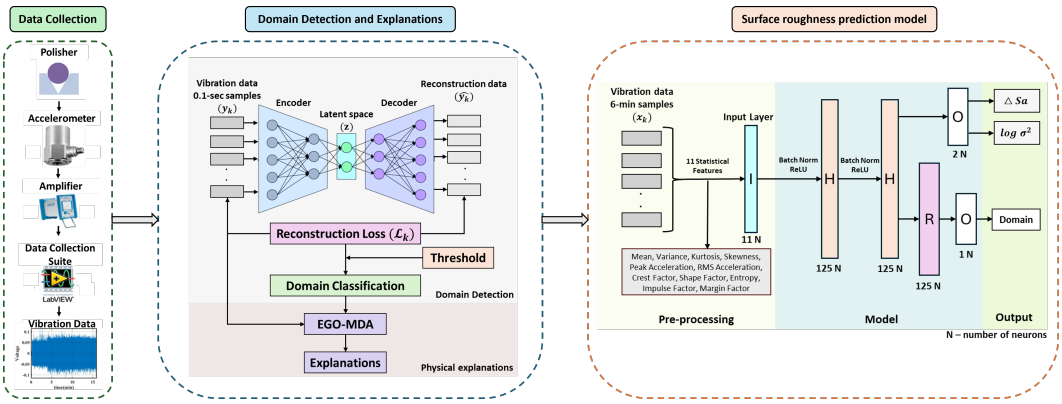


Figure 1: Methodology overview

(MEMs) sensor technology and artificial intelligence have enabled the use of vibration sensors for monitoring, which is the primary focus of this study [8].

This study aims to develop a automated surface roughness prediction model for HDC shells using machine learning algorithms. The key challenges are the inherent variability in polishing conditions, which leads to diverse data sets that complicate model generalization, and the difficulty in maintaining consistent performance across different surface roughness profiles due to their variability. Additionally, the limited availability of data for training deep learning models presents a significant hurdle, as these models typically require large data sets for optimal performance. To overcome these challenges, this study makes the following key contributions:

Domain Classification and Physical Explanations: The study identifies distinct domains within the data, establishing a robust foundation for the prediction model. In addition, it provides frequency-based physical explanations for these domains, deepening the understanding of the polishing process.

Enhanced Domain Adaptation Method: To ensure the generalizability and precision of the model under varying polishing conditions, we introduce an enhanced domain adaptation method. This method is designed to maintain consistent performance with variability in surface roughness profiles.

The core idea is to classify distinct vibration domains and leverage these classification results to enhance the domain adaptation model for surface roughness predictions. This approach reduces the manual workload, improves quality control, and increases process efficiency in automated surface roughness prediction for HDC shells. By advancing the precision of ICF experiments, this study contributes to the broader goal of achieving sustainable fusion energy generation.

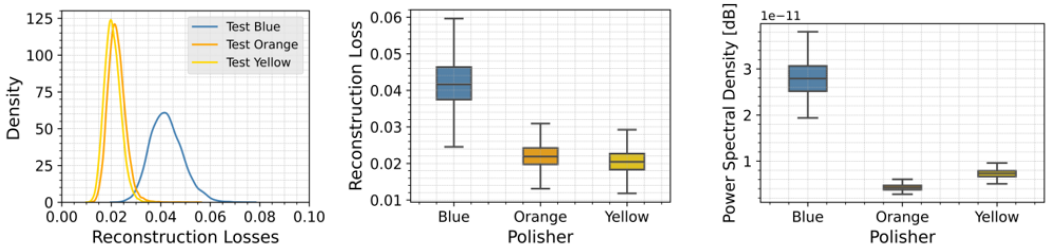
2 Methodology

2.1 Data Collection

We started this study by mounting vibration sensors to a polishing machine and collecting data at 10,000 Hz. Data collection is structured into two stages:

Baseline Experiments: The experiments were designed to classify domains and gain insights into the polishing process. Conducted on three polishers—blue, orange, and yellow—under identical conditions, each polishing session lasted one hour, during which vibration data was collected. This data was then segmented into 0.1-second intervals, resulting in time series data points with 1,000 samples each (y_k). These segments were used to identify distinct vibration domains, facilitating a better understanding of the underlying processes.

Surface Roughness Prediction: We conducted five 24-hour polishing experiments to develop a robust and generalizable model for predicting surface roughness. In each experiment, we collected vibration data and surface roughness measurements, which were segmented into 6-minute intervals. The surface roughness values (Sa) for each segment were assigned using linear-log interpolation [2]. The datasets from these experiments were labeled S173, S179, S211, S233, and S238. Each 6-minute segment (x_k) resulted in data points sized 360 10,000, enabling surface roughness predictions every six minutes throughout the polishing process. This segmentation approach ensures continuous and timely predictions, enhancing the model's applicability in real-time polishing operations.



(a) KDE (loss distribution) (b) Box plots (loss distribution) (c) Energy box plots [825, 3000] Hz

Figure 2: (a) Density plots of unseen testing data highlight the separation between Blue and Yellow/Orange domains using a 0.03 threshold. (b) Box plots display reconstruction loss distributions, with higher losses for the Blue polisher. (c) The EGO-MDA identified frequency band [825, 3000] Hz reveals the energy difference between the Blue and Yellow/Orange polishers.

We first identify distinct vibration domains using baseline data, which are then leveraged as a pre-classification step on experimental data, to improve the surface roughness model predictions.

2.2 Autoencoders for Domain Classification

An autoencoder with three hidden layers in both the encoder and decoder, and an embedding dimension of ($z = 32$), was trained using Mean Absolute Error (MAE) as the loss function. A reconstruction loss ($\mathcal{L}_k$) threshold was set to differentiate domains within the data.

For physical interpretation, our aim is to identify frequency bands that capture energy differences between domains. To achieve this, we adapted the Bayesian unsupervised EGO-MDA algorithm [14, 6], which identifies optimal frequency bands for given spectral data. EGO-MDA employs a gradient-free search strategy, first applying Mixture Discriminant Analysis (MDA) to determine the best classifier with model parameters M^*. Then, Efficient Global Optimization (EGO) iteratively searches for the frequency bands α^* that best distinguish classes by minimizing the deviation:

$$D(\alpha, M) = -2 \sum_{r \in \{A,B\}} \sum \log \hat{p}(r \mid E^{(r)}(\alpha)) + \eta \log(|\alpha|) \tag{1}$$

where $E^{(r)}(\alpha)$ represents the frequency band energy for class r, $\hat{p}(r \mid E^{(r)}(\alpha))$ is the class probability, and $|\alpha|$ denotes the total span of frequency bands with η as the penalty coefficient.

2.3 Surface Roughness Prediction Experiments

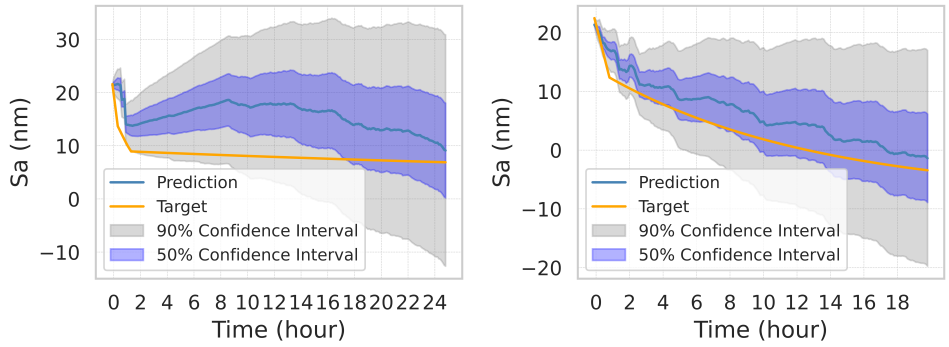
Following domain Classification, a neural network was designed to predict changes in arithmetic surface roughness (ΔSa) from the vibration data, reducing the need for continuous surface roughness measurements. The model minimizes the prediction error for the final surface roughness value (Sa_{final}) and incorporates a confusion metric into the loss function to improve the prediction confidence. To quantify prediction uncertainty, the model assumes a normal confidence distribution and applies Kullback-Leibler Divergence Loss. The baseline loss for a discrete probability distribution and fixed scalar target is:

$$-\log Q(x) = \frac{1}{2}\log(2\pi\sigma^2) + \frac{(x-\hat{x})^2}{2\sigma^2} \tag{2}$$

The custom loss function builds on this by incorporating mean absolute error (MAE) and handling small target values through exponentiation. A linear regularization term for σ prevents excessive uncertainty, with scaling factors γ_1 and γ_2 that balance variance loss. The final KL loss is:

$$L(\hat{y}, \log(\sigma^2), y) = \gamma_1 \frac{e^{|y-\hat{y}|}}{\sigma} + \gamma_2 \left(\sigma + \log(\sigma\sqrt{2\pi})\right) \tag{3}$$

Here, $\hat{y}$ is the predicted ΔSa, y the target value, and σ^2 the prediction variance. γ_1 and γ_2 balance accuracy and variance, optimizing for accurate and confident predictions, particularly in domain adaptation.



(a) Validation result for S238.

(b) Test result for S211.

Figure 3: ΔSa and confidence intervals (CI) are predicted using 6-minute interval data. The graph cumulates predictions, projecting Sa values over time. The 50% and 90% CIs are shown with blue and grey shading, respectively. (a) ΔSa prediction with CI for validation data. (b) ΔSa prediction with CI for test data.

3 Results and Discussion

3.1 Domain Classification and Explanations

We trained the autoencoder on 80% of the samples, reserving 20% for testing, with performance measured by reconstruction loss. Our analysis identified two distinct domains: one for the Blue polisher and another combining the Orange and Yellow polishers. The Blue polisher exhibited significantly higher reconstruction losses, as shown in the loss distribution (Figure 2a) and box plots (Figure 2b). Across five experimental runs (S173, S179, S211, S233, S238), the autoencoder grouped the Blue polisher runs (S173, S179, S233) into one domain, while the Orange and Yellow polishers (S211, S238) formed another.

To investigate these domain differences, we applied EGO-MDA to identify frequency bands distinguishing between the Blue polisher and the combined Orange/Yellow spectra. In the 825-3000 Hz range, the Blue polisher's energy levels were 72% and 81% higher than the Yellow and Orange polishers, respectively (Figure 2c), explaining the observed differences between the polishers.

3.2 Automated Surface Roughness Prediction Model

The primary objective of our automated surface roughness prediction model is to enhance the polishing process by providing accurate and reliable predictions. We focused on capturing surface roughness trends with minimal error, using the custom KL loss as the primary metric for evaluating model accuracy and loss. Additionally, we assessed the model's performance within the distinct domains identified by the autoencoder.

The prediction model was trained using 11 statistical features extracted from raw vibration data: mean, variance, kurtosis, skewness, peak acceleration, RMS acceleration, crest factor, shape factor, entropy, impulse factor, and margin factor. Training was conducted on data from the blue domain or domain A (S173, S179, and S233) and subsequently validated and tested on the orange and yellow domains or domain B (S238 and S211). The following plots illustrate the results of these experimental runs.

The test results for S211 and S238 [Figure 3] demonstrate the model's ability to accurately predict changes in surface roughness (ΔSa), which were then translated into actual surface roughness values (Sa) throughout the polishing process. The plots visually depict the model's predictions on unseen test data, highlighting its generalization capability across different domains.

The findings indicate that the model effectively predicts surface roughness trends, especially in the early stages of the polishing process, which is critical to predict to determine if polishing is headed in the right direction. However, the increasing prediction error and widening confidence intervals in later stages underscore the challenges of real-time prediction in complex, stochastic environments like polishing.

4 Discussion and Future Work

4.1 Discussion

Our primary objective is to automate the prediction of surface roughness during the polishing process. To achieve this, we developed a Domain Adaptation Prediction Model that leverages autoencoders for domain identification. The methodology begins with domain identification, where we cluster the distribution of reconstruction losses to identify distinct domains. Frequency-based explanations using EGO-MDA further validate these domains, allowing us to detect similar characteristics in the data without relying on predefined labels. Once domains are identified, the model is trained on data from one domain A (e.g., Blue) and tested on domain B (e.g., Yellow or Orange). This cross-domain evaluation assesses the model's generalizability and its ability to adapt to different data distributions resulting from variations in machine vibrations and environmental conditions.

Domain adaptation is particularly crucial in our context because the polishing process is conducted on multiple machines. Although the polishing conditions and sensor placements are identical, variations in internal vibrations and environmental factors can alter data distributions, resulting in distinct domains. Our methodology ensures that the model remains effective despite these variations.

In our surface roughness prediction experiments, we conducted five runs: three with the Blue polisher (S173, S179, S233), one with the Yellow polisher (S211), and one with the Orange polisher (S238). This setup mirrors real-world dynamic scenarios where data originates from different machines. We trained the model using data from the Blue domain and validated and tested it on data from the Yellow and Orange domains. Figure 3 illustrates the testing and validation results for Domain B, demonstrating the model's capacity to adapt and maintain accuracy across different domains.

4.2 Future Work

Our next step is to implement the developed model for real-time surface roughness prediction with six-minute intervals, facilitating timely decision making in operational settings. Prior to deployment, we will conduct extensive training and testing on diverse datasets and perform inference studies to ensure the model's reliability and scalability in real-time environments.

Recognizing that real-world applications may involve additional domains, we will enhance the model's domain adaptation capabilities by developing methods to automatically identify and incorporate new domains without compromising prediction performance. This will ensure that the model remains generalizable and robust as more domains are introduced.

We also plan to expand our dataset by including more runs and polishers to cover a broader range of operational conditions. This expansion will involve thorough tests in various machine configurations and environmental settings to validate the consistency of the model accuracy in an enlarged data set.

Acknowledgements

This work was performed under the auspices of the U.S. Department of Energy by LLNL under contract DE-AC52 -07NA27344 and was supported by the LLNL laboratory-directed research and development (LDRD) program under project 23-ERD-014.

This material is partially based on work supported while serving at the National Science Foundation. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

During the preparation of this work, the author(s) used ChatGPT, developed by Open AI, to improve readability and overall language. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content. No AI tool was used for problem formulation or analysis or to draw insights in the research process.

References

[1] H Abu-Shawareb, R Acree, P Adams, J Adams, B Addis, R Aden, P Adrian, BB Afeyan, M Aggleton, L Aghaian, et al. Achievement of target gain larger than unity in an inertial fusion experiment. *Physical Review Letters*, 132(6):065102, 2024.

[2] Antonios Alexos, Junze Liu, Shashank Galla, Sean Hayes, Kshitij Bhardwaj, Alexander Schwartz, Monika Biener, Pierre Baldi, Satish Bukkapatnam, and Suhas Bhandarkar. Nuclear fusion diamond polishing dataset. In *The Thirty-eight Conference on Neural Information Processing Systems Datasets and Benchmarks Track*, 2024.

[3] J Biener, DD Ho, C Wild, E Woerner, MM Biener, BS El-Dasher, DG Hicks, JH Eggert, PM Celliers, GW Collins, et al. Diamond spheres for inertial confinement fusion. *Nuclear Fusion*, 49(11):112001, 2009.

[4] DT Casey, JL Milovich, VA Smalyuk, DS Clark, HF Robey, A Pak, AG MacPhee, KL Baker, CR Weber, T Ma, et al. Improved performance of high areal density indirect drive implosions at the national ignition facility using a four-shock adiabat shaped drive. *Physical review letters*, 115(10):105001, 2015.

[5] DS Clark, AL Kritcher, SA Yi, AB Zylstra, SW Haan, and CR Weber. Capsule physics comparison of national ignition facility implosion designs using plastic, high density carbon, and beryllium ablators. *Physics of Plasmas*, 25(3), 2018.

[6] Shashank Galla, Akash Tiwari, Saikiran Chary Nalband, Sean Michael Hayes, Suhas Bhandarkar, and Satish Bukkapatnam. Detecting anomalous motions in ultraprecision shell-polishing process combining unsupervised spectral-band identification and explainable-ai. *Journal of Manufacturing Systems*, 2024.

[7] OA Hurricane, PK Patel, R Betti, DH Froula, SP Regan, SA Slutz, MR Gomez, and MA Sweeney. Physics principles of inertial confinement fusion and us program overview. *Reviews of Modern Physics*, 95(2):025005, 2023.

[8] Shilan Jin, Rui Tuo, Akash Tiwari, Satish Bukkapatnam, Chantel Aracne-Ruddle, Ariel Lighty, Haley Hamza, and Yu Ding. Hypothesis tests with functional data for surface quality change detection in surface finishing processes. *IISE transactions*, 55(9):940–956, 2023.

[9] Lawrence Livermore National Laboratory. A shot for the ages: Fusion ignition breakthrough hailed as 'one of the most impressive scientific feats of the 21st century'. `https://rb.gy/oczaib`, 2022. Accessed: [January 2023].

[10] Lawrence Livermore National Laboratory. Target evolution: Key to llnl's continued success, 2024.

[11] Nature. Nuclear fusion breakthrough: this physicist helped to achieve the first-ever energy gain. *Nature*, 625:11–12, 2024.

[12] JS Ross, D Ho, J Milovich, T Döppner, J McNaney, AG MacPhee, A Hamza, J Biener, HF Robey, EL Dewald, et al. High-density carbon capsule experiments on the national ignition facility. *Physical Review E*, 91(2):021101, 2015.

[13] Mark J Schmitt, Paul A Bradley, James A Cobble, Scott C Hsu, Natalia S Krasheninnikova, George A Kyrala, Glenn R Magelssen, Thomas J Murphy, Kimberly A Obrey, Ian L Tregillis, et al. Defect-induced mix experiment for nif. In *EPJ Web of Conferences*, volume 59, page 04005. EDP Sciences, 2013.

[14] Akash Tiwari and Satish Bukkapatnam. Unsupervised spectral-band feature identification for optimal process discrimination. *arXiv preprint arXiv:2212.03800*, 2022.

[15] AB Zylstra, AL Kritcher, OA Hurricane, DA Callahan, JE Ralph, DT Casey, A Pak, OL Landen, B Bachmann, KL Baker, et al. Experimental achievement and signatures of ignition at the national ignition facility. *Physical Review E*, 106(2):025202, 2022.

A Supplementary Material

A.1 Explanation on Custom Kullback–Leibler Divergence Loss

In this section, we provide a detailed explanation of the custom Kullback-Leibler (KL) Divergence Loss function.

We begin with the definition of the KL divergence for discrete probability distributions P and Q defined on the same sample space $\mathcal{X}$:

$$D_{\mathrm{KL}}(P \parallel Q) = \sum_{x \in \mathcal{X}} P(x) \log \left(\frac{P(x)}{Q(x)} \right)$$

For our specific case, where the target y is a scalar value with zero variance, the KL divergence simplifies to:

$$D_{\mathrm{KL}}(P \parallel Q) = \log \left(\frac{1}{Q(x)} \right)$$

Assuming that the distribution $Q(x)$ follows a normal distribution with mean μ and variance σ^2, the probability density function is given by:

$$Q(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left(-\frac{(x-\mu)^2}{2\sigma^2} \right)$$

Taking the negative logarithm, we obtain the KL loss as:

$$-\log Q(x \mid \mu, \sigma^2) = \frac{1}{2} \log(2\pi\sigma^2) + \frac{(x-\mu)^2}{2\sigma^2}$$

Next, we introduce the customized loss function, which incorporates the observed benefits of the mean absolute error (MAE) and accounts for very small target values through exponentiation. Additionally, to prevent the model from increasing uncertainty indefinitely, a linear regularization term for σ is included. The custom loss function is defined as:

$$L(\hat{y}, \log(\sigma^2), y) = \gamma_1 \frac{e^{|y-\hat{y}|}}{\sigma} + \gamma_2 \left(\sigma + \log(\sigma\sqrt{2\pi}) \right)$$

This function combines the benefits of MAE, adjusted for small target values, with a regularization term to control the model's uncertainty. The scaling factors γ_1 and γ_2 are introduced to balance the contribution of each component to the overall loss.

Thus, the final form of the custom KL loss function is obtained.