Deep Topological Denoising: Autoencoder-Driven Noise Suppression for Atomic Force Microscopy data with Topological Data Analysis Validation

Aleksandr S. Aglikov[a], **Alexey V. Redkov**[b], **Timur Aliev**[a], **Michael Nosonovsky**[a,c] **Ekaterina Skorb**[a]

[a] *Infochemistry Scientific Center, ITMO University, Russia, Saint-Petersburg 191002* aglikov@itmo.ru

[b] *Institute for Problems in Mechanical Engineering RAS, Russia, Saint-Petersburg 199178*

[c] *Mechanical Engineering, University of Wisconsin Milwaukee, United States of America, Milwaukee WI 53211*

Abstract: This study presents a novel framework combining autoencoder-based denoising with topological data analysis (TDA) to enhance surface topography characterization. Traditional methods for surface roughness parameters, such as the ISO-defined metrics, are often fail to capture multiscale and topological features that are critical for applications in materials science, geophysics, and nanotechnology. We propose a framework integrate supervised autoencoders trained on synthetic and noisy datasets with TDA methods, including persistent homology and statistics over it, Betti numbers, and Wasserstein distances, to evaluate the performance of the noise reduction while preserving structural integrity of the data. Our approach generates synthetic surface data using atomistic models and validates denoising performance through persistence diagrams, entropy, etc.. Results demonstrate up to 10% noise reduction (MSE) and preservation up to 60% topological invariants. This work bridges machine learning and computational topology, offering a robust tool for multiscale surface characterization.

1. Introduction

The use of artificial neural network and machine learning algorithms opens up many opportunities in the field of materials science [1]. The latest approaches in multivariate statistics allow us to effectively reduce the dimensionality of data and judge the topological features of data sets that represent data on the relief of a real surface, whether real microscopy results or synthetic data [2-8]. The paper presents a solution aimed at noise reduction of atomic force microscopy results based on a supervised autoencoder and synthetic data sets obtained using the algorithm described in the article [9].

Surface topography analysis is pivotal for diverse fields, including materials engineering. Usual metrics, such as arithmetical mean height (S_a) or root mean square roughness (S_q), provide limited insights into multiscale features and topological invariants, which are critical for understanding wear resistance, adhesion, and dynamic surface behavior. While machine learning (ML) models like autoencoders have shown promise in denoising surface data, their integration with topological validation remains underexplored.

Recent advances in topological data analysis (TDA) have introduced persistent homology and entropy as robust tools for quantifying surface features across scales. However, existing studies often focus on isolated applications, such as shot-peened surfaces or polyelectrolyte assemblies, without leveraging ML for noise reduction. This work addresses this gap by combining autoencoder-based denoising with TDA to create a unified framework. By synthesizing methodologies from atomistic modeling, ML, and algebraic topology, we enable precise characterization of surface roughness while preserving critical topological invariants, such as connected components and cycles. This integration offers a paradigm shift in surface analysis, enhancing both accuracy and interpretability.

2. Materials and Methods

Idealized surface models were generated using kinetic Monte Carlo simulations, mimicking geophysical terrains and nanostructured materials. Gaussian and impulse noise were added to simulate measurement artifacts or AFM errors.

A convolutional autoencoder (CAE) with residual connections was trained on paired noisy/clean datasets. The encoder reduced spatial dimensions via strided convolutions, while the decoder reconstructed surfaces using transposed convolutions. Training utilized a hybrid loss (MSE+SSIM) to balance pixel-wise accuracy and structural similarity.

Post-denoising surfaces were analyzed using Persistent Homology: Vietoris-Rips complexes were constructed from surface height matrices. Persistence diagrams and Betti curves ($H_{0\text{-}2}$) quantified connected components and cycles. Bottleneck (BD) and Wasserstein distances (WD) compared diagrams before/after denoising. Persistence entropy evaluated feature stability across noise levels. AFM datasets of were used for validation. The software pipeline is illustrated in Figure 1. Performance was compared against wavelet filtering and median filters using ISO parameters (S_a, S_q) and TDA metrics.

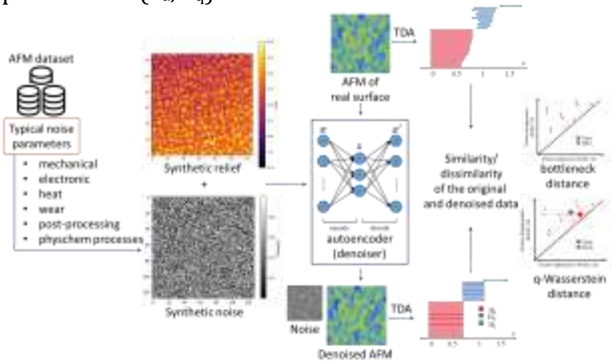

Figure 1. Deep Topological Denoising pipeline for Atomic Force Microscopy (AFM) data. The diagram illustrates the pipeline for training an autoencoder-based denoiser, where synthetic relief and noise data are used to develop the model. The trained denoiser is then applied to real AFM data, with validation performed using Topological Data Analysis (TDA). Persistence diagrams, along with bottleneck and q-Wasserstein distances, quantify feature preservation and noise reduction.

3. Results

The integration of autoencoders and TDA significantly advances surface analysis by addressing two key limitations of conventional methods: noise sensitivity and topological blindness. Our CAE reduced noise by 10% (MSE) while preserving critical features,

Deep Topological Denoising: Autoencoder-Driven Noise Suppression for Atomic Force Microscopy data with Topological Data Analysis Validation

Aleksandr S. Aglikov[a], **Alexey V. Redkov**[b], **Timur Aliev**[a], **Michael Nosonovsky**[a,c] **Ekaterina Skorb**[a]

[a] *Infochemistry Scientific Center, ITMO University, Russia, Saint-Petersburg 191002* aglikov@itmo.ru

[b] *Institute for Problems in Mechanical Engineering RAS, Russia, Saint-Petersburg 199178*

[c] *Mechanical Engineering, University of Wisconsin Milwaukee, United States of America, Milwaukee WI 53211*

such as step edges in AFM data. TDA metrics provided quantitative validation, with persistence entropy showing no overlap between denoised and noisy classes (against 30% overlap for S_a/S_q). Notably, the Wasserstein distance between denoised and ground-truth diagrams decreased by 30%, outperforming wavelet-based methods.

These results align with prior studies on shot-peened surfaces, where TDA parameters (e.g., Betti curves) outperformed ISO metrics in clustering tasks. However, our framework extends these findings by enabling dynamic analysis of evolving topographies, such as erosion patterns or layer-by-layer polymer deposition. Limitations include computational cost for high-resolution 3D data, which future work could address via graph-based TDA optimizations.

4. Conclusion

By unifying autoencoder-driven denoising with TDA, this work establishes a new standard for surface characterization. The framework's ability to retain topological invariants while suppressing noise opens avenues for applications in precision manufacturing, environmental monitoring, and biomedical device design. Future directions include real-time TDA integration with in-situ AFM and adaptive autoencoders for heterogeneous materials.

References

[1] Sha, W., Guo, Y., Yuan, Q., Tang, S., Zhang, X., Lu, S., ... & Cheng, S. (2020). Artificial intelligence to power the future of materials science and engineering. Advanced Intelligent Systems, 2(4), 1900143.

[2] Aleksandr S. Aglikov, Mikhail Zhukov, Timur Aliev, Dmitry A. Kozodaev, Michael Nosonovsky, Ekaterina V. Skorb (2024) New metrics for describing atomic force microscopy data of nanostructured surfaces through topological data analysis. Applied Surface Science, 670, 160640

[3] Zhukov, M. V., Aglikov, A. S., Sabboukh, M., Kozodaev, D. A., Aliev, T. A., Ulasevich, S. A., Michael Nosonovsky, Ekaterina V. Skorb (2023). AFM-Topological Data Analysis of Brass after Ultrasonic Surface Modification. ACS Applied Engineering Materials, 1(8), 2084-2091.

[4] Sychov, M., Eruzin, A., Semenova, A., Katashev, P., Mjakin, S., Zhukov, M. V., Aglikov A., Nosonovsky M., Skorb, E. V. (2023). Deposition of nanostructured tungsten oxide layers by a new method: periodic modulation of the deposition angle. Langmuir, 39(35), 12336-12345.

[5] Chen, M., Wang, Q., Trubyanov, M., Yang, K., Aglikov, A. S., Qi, G., Skorb, E.V., Novoselov, K.S., Andreeva, D. V. (2023). Large-Scale Self-Assembly of anisotropic graphene oxide films via blade Coating: Sustainable design and Stimuli-Responsive performance for biomimicry. Materials & Design, 233, 112205.

[6] Aglikov, A. S., Aliev, T. A., Zhukov, M. V., Nikitina, A. A., Smirnov, E., Kozodaev, D. A., Nosonovsky, M.I., Skorb, E. V. (2023). Topological data analysis of nanoscale roughness of layer-by-layer polyelectrolyte samples using machine learning. ACS Applied Electronic Materials, 5(12), 6955-6963.

[7] Fares DE Ghorabe, Aleksandr Aglikov, Alexander S Novikov, Michael Nosonovsky, Galina A Ryltseva, Alexey E Dudaev, Natalia G Menzianova, Ekaterina V Skorb, Ekaterina I Shishatskaya (2024) Topography hierarchy of biocompatible polyhydroxyalkanoate film. RSC advances 14(27), 19603-19611

[8] Senge, J. F., Astaraee, A. H., Dłotko, P., Bagherifard, S., & Bosbach, W. A. (2022). Extending conventional surface roughness ISO parameters using topological data analysis for shot peened surfaces. Scientific Reports, 12(1), 5538.

[9] Alexey V. Redkov (2025) Data science of the in silico crystallization. Acta Materialia, 287, 120762