

Kuo Zhan

Peilin Xin, Yingqi Zhao, Han Gu, Enock Adjei Agyekum, Zhou Chen, Shuai Li, Jian Ye, Lu Cheng, Jian-an Huang*

Research Unit of Health Sciences and Technology, Faculty of Medicine, University of Oulu, Aapistie 5 A, 90220 Oulu, Finland.

Kuo.Zhan@oulu.fi, Jianan.Huang@oulu.fi

Physics-Aligned Deep Learning Enables SM-SERS DNA Oligomers decoding

Single-molecule surface-enhanced Raman spectroscopy (SM-SERS) captures dynamic molecular behavior with ultrahigh sensitivity⁽¹⁾, but its biopolymer analysis is hindered by strong spectral heterogeneity, transient hotspot sampling, and background interference^(2, 3). Here, we develop a physics-aligned deep learning framework integrating contrastive attention-based multiple-instance learning (CAMIL), a tri-channel multi-kernel CNN classifier, and neighbor-transition-guided sequence reconstruction to decode single-molecule DNA oligomer dynamics in a plasmonic nanocavity. In this work, CAMIL mines informative spectra enriched with chain-embedded nucleotide and dinucleotide signatures from contrastive DNA trajectory bags, generating a SM-SERS DNA-segment library. This library trains a 12-class CNN to assign query time-resolved DNA SM-SERS frames to high-confidence DNA-segment states, enabling trajectory-level analysis of transition matrix. The transition matrix is further converted into neighbor-state-transition evidence for candidate sequence scoring, enabling asymmetric DNA sequences inference from confined stochastic sampling dynamics. This framework transforms SM-SERS heterogeneity into quantitative analytical information, advancing dynamic single-molecule decoding.

References

- [1]. M. M. Schmidt, A. G. Brolo, N. C. ACS Nano, acsnano.4c09483 (2024).
- [2]. D. P. Dos Santos, M. L. A. Temperini, A. G. Brolo, Acc. Chem. Res. 52, 456–464 (2019).
- [3]. J.-A. Huang, M. Z. Mousavi, Y. Zhao, A. Hubarevich, F. Omeis, G. Giovannini, M. Schütte, D. Garoli, F. De Angelis, Nat Commun 10, 5321 (2019).

Figures

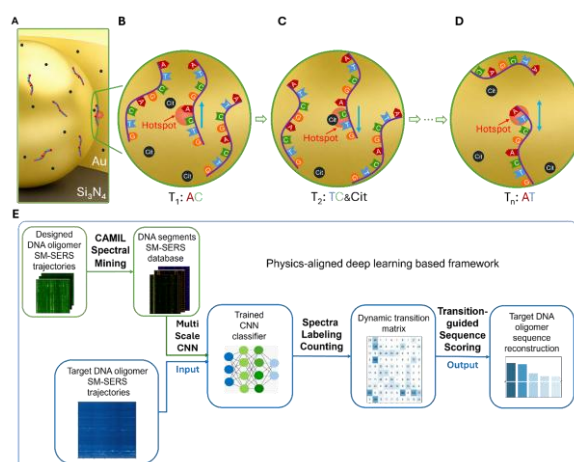


Figure 1: Plasmonic nanocavity-based SM-SERS heterogeneity and the physics-aligned deep learning framework for dynamic DNA oligomer decoding.