# Synthetic Data–Driven AI for 2D Materials and Optoelectronic Devices

**Yong Xie**
Thiago L Vasconcelos, Eduardo R Hernández, Andres Castellanos-Gomez
*Instituto de Ciencia de Materiales de Madrid (ICMM-CSIC), 28049 Madrid, Spain*
yxie@xidian.edu.cn, eduardo.hernandez@csic.es, andres.castellanos@csic.es

## Abstract

Two-dimensional (2D) materials, such as twisted bilayer graphene and transition metal dichalcogenides (TMDs), exhibit exceptional physical properties, making them promising candidates for next-generation electronic and optoelectronic devices. Recent advances in generating large-scale synthetic data, rooted in fundamental scientific principles and unhindered by experimental limitations, have unlocked new opportunities to train and optimize artificial intelligence (AI) models. This dynamic interplay between synthetic data and AI-driven analysis accelerates the discovery of novel phenomena and deepens our understanding of the underlying physics.

We demonstrate the transformative potential of this approach through three compelling examples. First, a convolutional neural network, trained on over 10,000 synthetic images of CVD-grown bilayer TMDs—spanning hexagonal to triangular geometries—achieves high-precision prediction of rotational angles, enhancing model accuracy and robustness [1]. Second, we generated 100,000 synthetic spectra from eight distinct LED sources, creating diverse datasets that significantly improve model generalizability [2]. Third, we introduce a cost-effective spectrometer using a commercial LED array in photodiode mode, paired with a deep learning algorithm for spectral reconstruction [3]. This accessible platform lowers barriers for researchers, enabling validation and benchmarking of spectroscopic performance and advancing compact, integrated technologies. By leveraging the synergy of synthetic data and AI, this work accelerates progress in 2D materials research and lays a foundation for broader innovations.

References

[1] H. Yang, R. Hu, H. Wu, X. He, Y. Zhou*, Y. Xue, K. He, W. Hu, H. Chen, M. Gong, X. Zhang, P.-H. Tan*, E. R. Hernández*, Y. Xie*, Nano Lett., 24 (2024) 2789−2797.
[2] Y. Liu, X. He, Q. Lei, R. Hu, Y. Xue, K. He, H. Yang, Y. Wang, X. Zhang, L. Yang, Y. Zhou, Y. Xie*, European Journal of Physics, (under revision, 2025).
[3] Y. Xie*, T. L. Vasconcelos, E. R. Hernández*, A. Castellanos-Gomez*, (to be submitted, 2025).

Figures

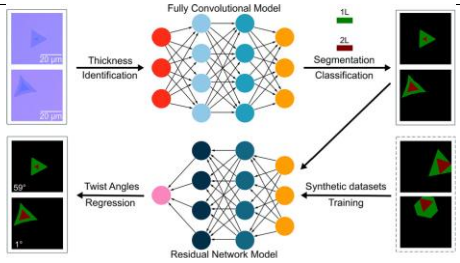


**Figure 1:** The two-step process to recognize the twisted angles of CVD-grown 2D materials using deep learning

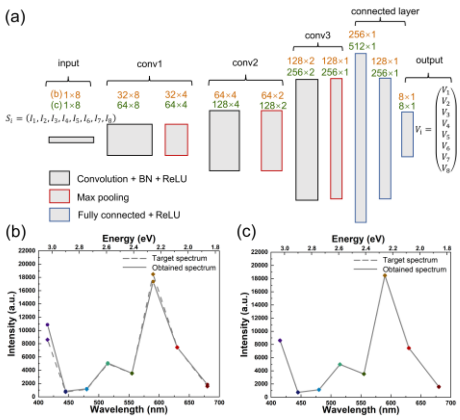


**Figure 2:** The performance of the spectrum using CNN algorithm