Applying Deep Learning Approaches to Estimate the Number of Layers in Nanomaterials from Optical Images

25th Symposium on Photonics and Optics (SPO-24)

November 4-8, 2024
Kyiv, Ukraine

I.S. Babichuk[1,2], R. Zhou[1], H. Liu[1], I.V. Babichuk[3] and J. Yang[1]

[1] School of Mechanical and Automation Engineering, Wuyi University, 529020, Jiangmen, P.R. China.

[2] V. Lashkaryov Institute of Semiconductor Physics, NAS of Ukraine, 03680, Kyiv, Ukraine.

[3] National Center "Minor academy of sciences of Ukraine", 04119, Kyiv, Ukraine.

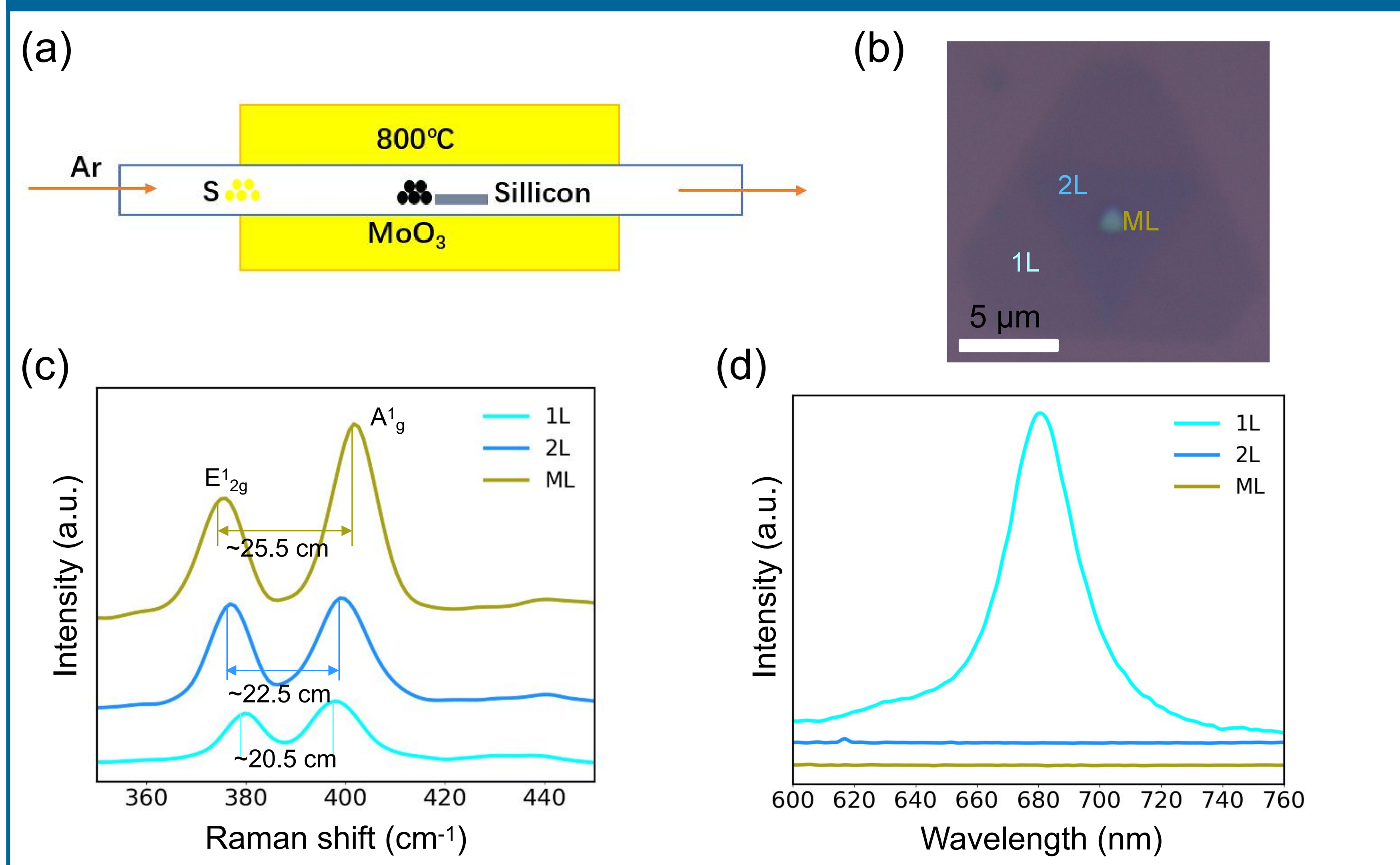
INTRODUCTION

Atomic-level engineering enables the creation of hybrid structures that enhance light-matter interactions and advance next-generation optical devices. Two-dimensional (2D) materials are important in nanophotonics due to their unique properties.

In this study, the authors introduce a hybrid vision transformer model, termed 2D-HVT, specifically designed for identifying and analyzing 2D materials. This model incorporates the FastViT encoder alongside a composite decoder that merges Lite Reduced Atrous Spatial Pyramid Pooling (LRASPP) and Knet, optimizing inference speed without compromising accuracy in assessing the thickness of molybdenum disulfide (MoS_2) layers, which may range from monolayer to multilayer configurations.

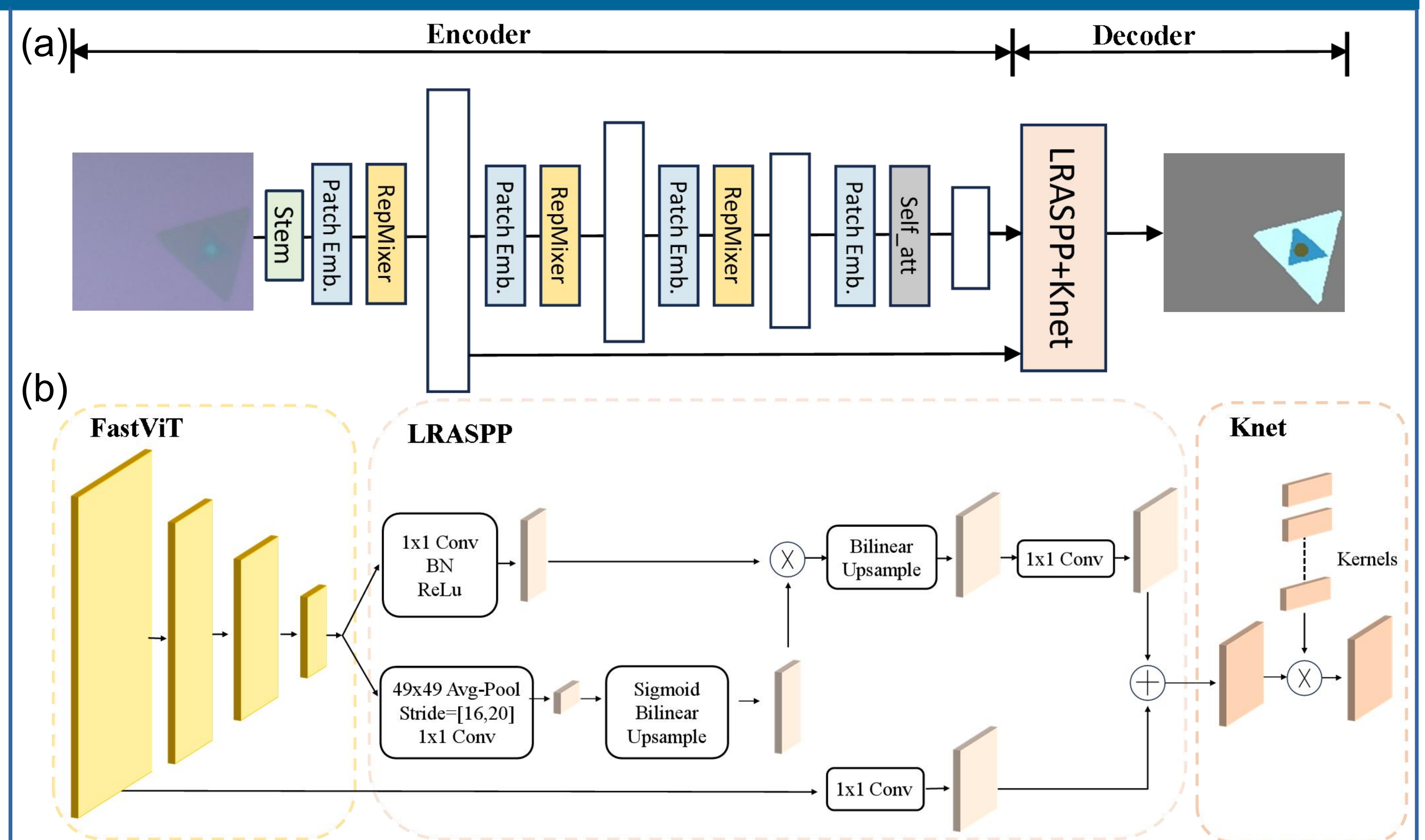
To facilitate practical application, we developed a graphical user interface (GUI) that enables real-time data acquisition from a microscopic camera. This interface effectively applies the trained model to produce predictive outputs based on the acquired images. The process begins with the microscope scanning a substrate chip coated with a MoS_2 film synthesized through CVD (Chemical Vapor Deposition), followed by image capture and processing. The resulting optical images serve as input for our model, which generates predictive outcomes that are further analyzed using histogram techniques to establish a layer distribution profile.

PREPARATION



CVD growth of MoS_2 (a). Microscope image of MoS_2 thin film (b). Raman spectra (c) for monolayer (1L), bilayer (2L) and multilayer (ML) positions in (b). PL spectra (d) for the same positions as (c).

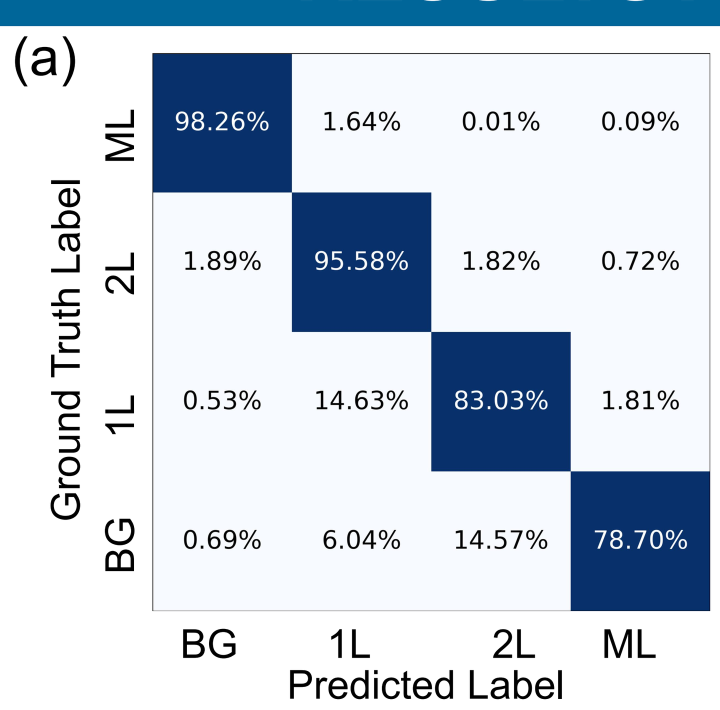
MODEL

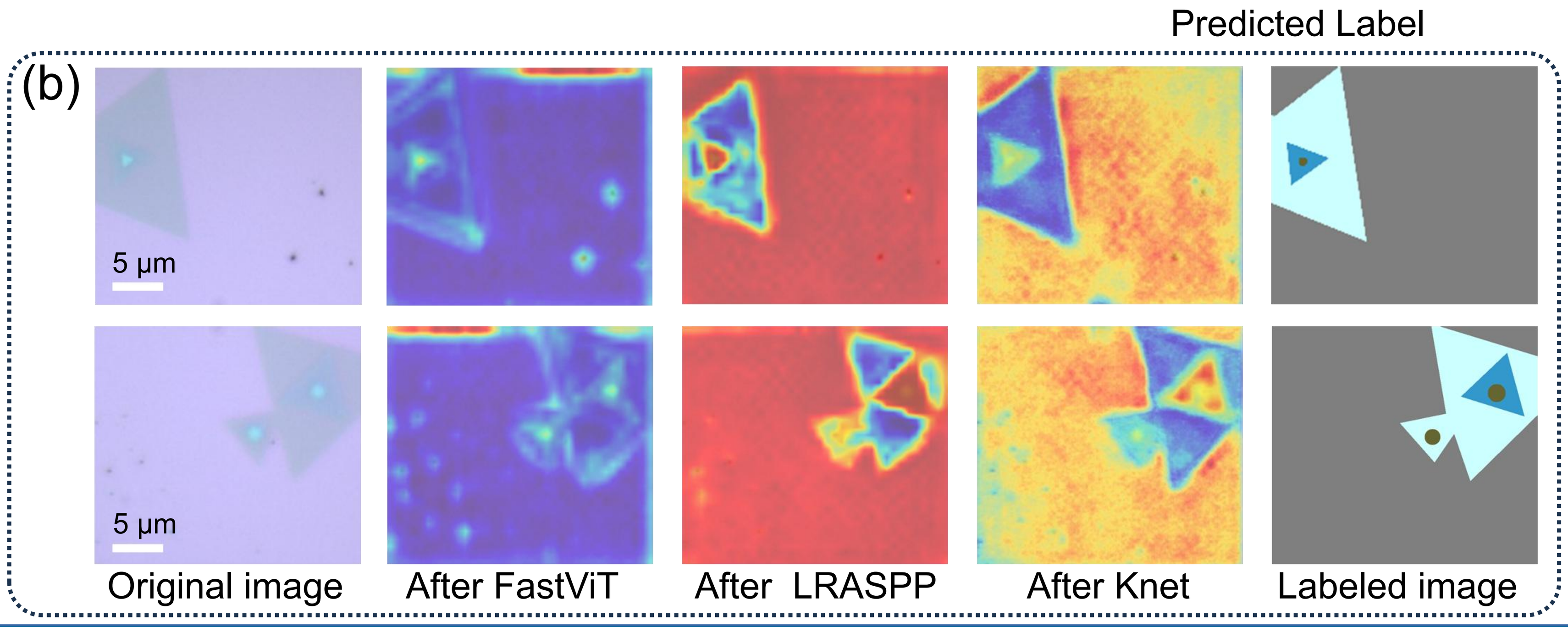


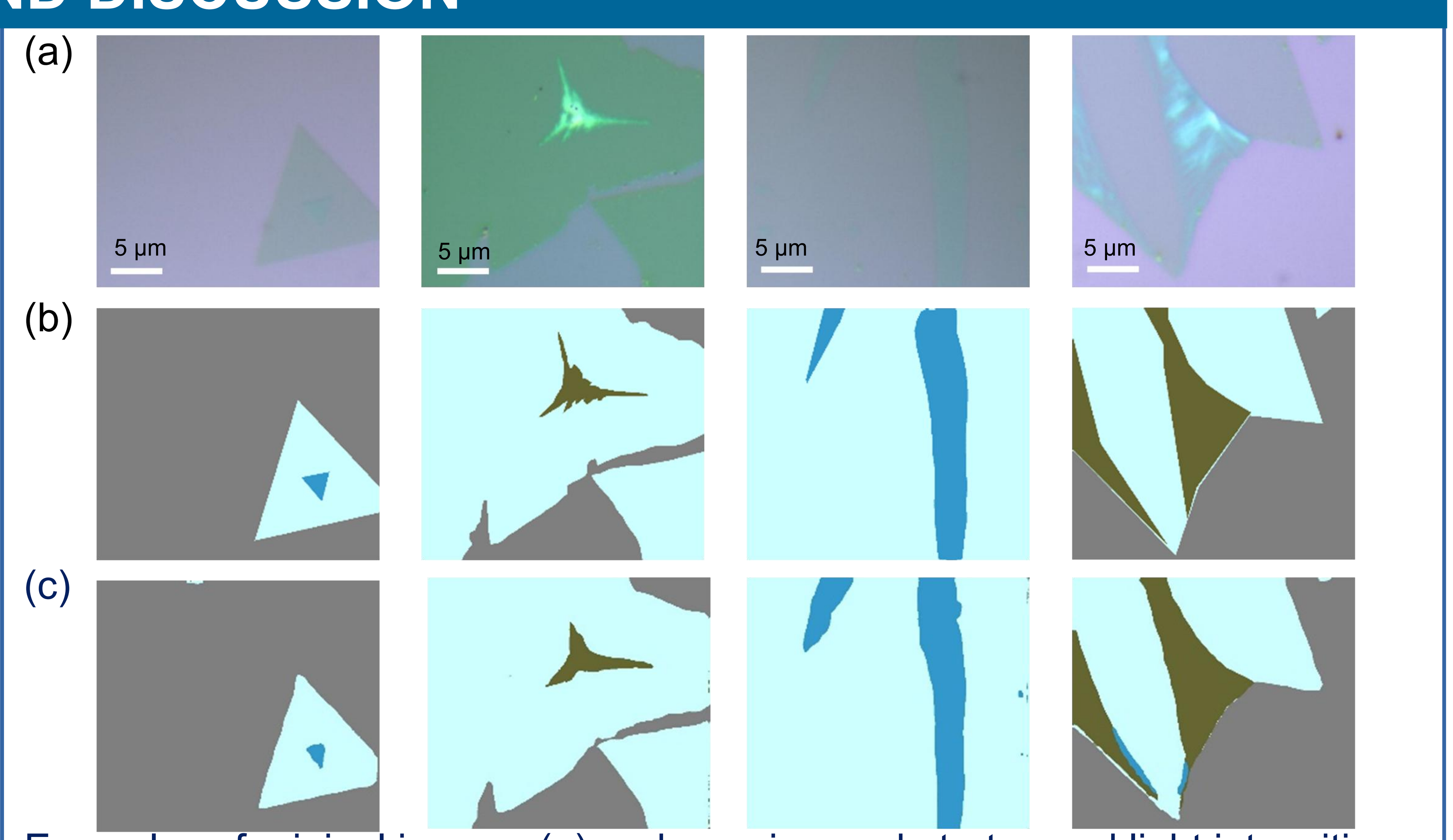
Overview of 2D-HVT (a). It consists of two main modules: a fast vision transformer encoder (FastViT) and a decoder with LRASPP and Ket (b).

RESULTS AND DISCUSSION

A confusion matrix showing classification results obtained by applying the trained model to the testing set (a). Feature map visualization output results on different layers of the trained 2D-HVT model (b). It illustrates the model's capability to progressively refine input data into an increasingly accurate and detailed representation.







Examples of original images (a) under various substrates and light intensities. Labeled images (b) and generated images (c) using the 2D-HVT model. In the labeled and generated images, cyan, blue, and brown indicate monolayer, bilayer, and multilayer regions, respectively. Black represents other areas.

CONCLUSIONS

1. The 2D-HVT model effectively utilizes advanced modules to process high-resolution images, ensuring high performance and computational efficiency.
2. The model demonstrates an exceptional ability to differentiate between various material layer formations, achieving high precision in segmentation tasks.
3. Future work aims to broaden the dataset with more annotated examples of 2D materials beyond MoS_2.

E-mail:
Zhou Ruiliang
3256405049@qq.com
Ivan S. Babichuk
ivan@szu.edu.cn

Our code is available at https://github.com/zhouruiliangxian/2D_materials