PAPER

Check for updates

High-throughput thickness analysis of 2D materials enabled by intelligent image segmentation

Jun Chen Ng,[a,b] Farina Muhamad,[b] Pauline Shan Qing Yeoh,[b] Ziyi Han,[a] Zanlin Qiu,[a,c] Khin Wee Lai [ID] *[b] and Xiaoxu Zhao [ID] *[a,c]

Thickness measurement of two-dimensional (2D) materials is essential due to their thickness-dependent physical and optical properties. However, current thickness characterization techniques, *e.g.*, Atomic Force Microscopy (AFM), suffer from limitations such as slow scanning, tip–sample artifacts, and low throughput. To address this, an Artificial Intelligence-based pipeline was proposed for estimating the thickness of 2D materials from Optical Microscopy (OM) images, offering a significantly faster and more efficient alternative. OM captures colour contrast due to thin-film interference, explained by Fresnel's law. These colour cues, along with morphological features (area and perimeter), were extracted from the regions of interest (ROIs) segmented using Otsu's thresholding. Several regression models, including Random Forest Regressor (RFR) and a shallow Multi-Layer Perceptron (MLP), were trained on augmented paired OM-AFM data. Both models performed well on representative 2D materials, *e.g.*, In_2Se_3, under threshold-based segmentation, but only the MLP maintained strong accuracy with automated ROI detection using Cellpose, achieving excellent predictive performance ($R^2 = 0.947$, MSE = 34.580 nm^2, MAE = 4.696 nm, RMSE = 5.881 nm). Statistical analysis validated the model's generalizability across segmentation methods. Shapley Additive Explanations (SHAP) identified red and green intensities as key predictors, aligning with thin-film interference theory. Overall, this AI-based model provides a non-destructive, efficient alternative to AFM, allowing precise and continuous thickness estimation from small datasets with high robustness and generalizability.

Introduction

2D materials have attracted growing interest due to their unique electronic,[1,2] optical,[3,4] and mechanical properties,[5,6] with promising applications in electronics,[7,8] energy storage,[9,10] sensing,[11,12] and quantum information.[13,14] Characterized by their atomically thin, sheet-like structures, often just a few nanometres thick, these materials allow free electron movement within the plane while confining motion in the third dimension due to quantum effects, resulting in properties that significantly vary with the number of layers.[15,16]

Focusing on 2D materials, *e.g.*, In_2Se_3, they typically exhibit triangular or hexagonal shapes and exist in multiple phases. These phases can be differentiated based on their band gaps and are commonly characterized by Raman[17,18] and X-ray diffr-

action (XRD) analyses.[19,20] For more accurate phase identification at the atomic level, advanced techniques such as high-resolution cross-sectional transmission electron microscopy (TEM) and scanning transmission electron microscopy (STEM) are employed, enabling atom-resolved imaging of the crystalline structures.[21–23]

Accurate thickness determination is essential, *e.g.*, In_2Se_3, because its structural, electronic, and optical properties vary significantly with thickness, directly influencing device performances.[24] Thickness plays a critical role in governing phase stability by modulating interlayer coupling, strain distribution, and symmetry breaking.[25–29] These trends highlight how thickness is not merely a geometric parameter but a key driver of phase selection and functional performance in In_2Se_3-based applications. Therefore, precise thickness quantification is essential for understanding and optimizing the material's properties for targeted applications.

Consequently, several methods have been introduced to determine the thickness of 2D materials. Among the techniques, AFM is a popular technique for analyzing material surface properties, including thickness and topography, by providing high-resolution 3D surface profiles.[30–32] Initially,

[a] School of Materials Science and Engineering, Peking University, Beijing 100871, China. E-mail: khinwee.lai@um.edu.my, xiaoxuzhao@pku.edu.cn
[b] Department of Biomedical Engineering, Faculty of Engineering, Universiti Malaya, Kuala Lumpur 50603, Malaysia
[c] AI for Science Institute, Beijing, China

AFM was limited to insulating materials, but advancements have evolved it into a versatile technique for characterizing surfaces at the atomic and nanometre scales.[33,34] Its ability to deliver absolute height measurements makes it a preferred method for quantifying the thickness of 2D materials. However, AFM suffers from inherent limitations that restrict its practicality for high-throughput and large-area characterization. These include a slow scanning speed, limited vertical range, tip–sample interaction artifacts, and issues such as surface deformation, laser misalignment, tip contamination, and ambiguous contact points.[35,36] Furthermore, the relatively low repeatability and acquisition speed hinder the generation of large, consistent datasets necessary for robust Machine Learning (ML) model development.[37]

Given these limitations, there is a growing need for automated and scalable alternatives, particularly those driven by Artificial Intelligence (AI), to accelerate the characterization of 2D materials. Even though AI-enhanced AFM workflows have already been explored to accelerate data analysis,[38,39] reliance on AFM hardware still imposes throughput constraints. To address this, optical microscopy (OM) has been explored as a faster, non-destructive alternative. While OM does not provide direct thickness measurement, the thin-film interference effect produces colour contrast that correlates with flake thickness. While early approaches have leveraged optical contrast by creating empirical colour charts of fitting optical spectra using Fresnel-based models,[40,41] these conventional methods are often labour-intensive and lack scalability. In response, AI has emerged as a powerful tool to automate and accelerate thickness estimation directly from OM images, enabling scalable and data-driven workflows suitable for high-throughput screening.[42–46]

Building on these insights, this study advances beyond conventional analytical and classification-based methods by implementing an AI-driven approach for continuous thickness estimation. Rather than classifying OM images into discrete layer numbers, this work aims to predict the actual height values, offering more detailed and quantitative information. Focusing specifically on In_2Se_3, a representative 2D material, the proposed method utilized surface morphological features, namely colour intensity and geometric measurements (area and perimeter), extracted from OM images. Several regression models were evaluated, and the MLP model was selected to learn both linear and non-linear correlations between these features and thickness due to its superior performance, offering an automated, data-driven alternative to manual or physics-based estimation techniques.

Results and discussion

In_2Se_3 crystals, exhibiting typical polygonal or hexagonal morphologies with colour variations induced by thin-film interference, were imaged using OM. The ROI was identified and manually marked (red rectangles), and its corresponding height values were extracted from aligned AFM scans (Fig. 1a).
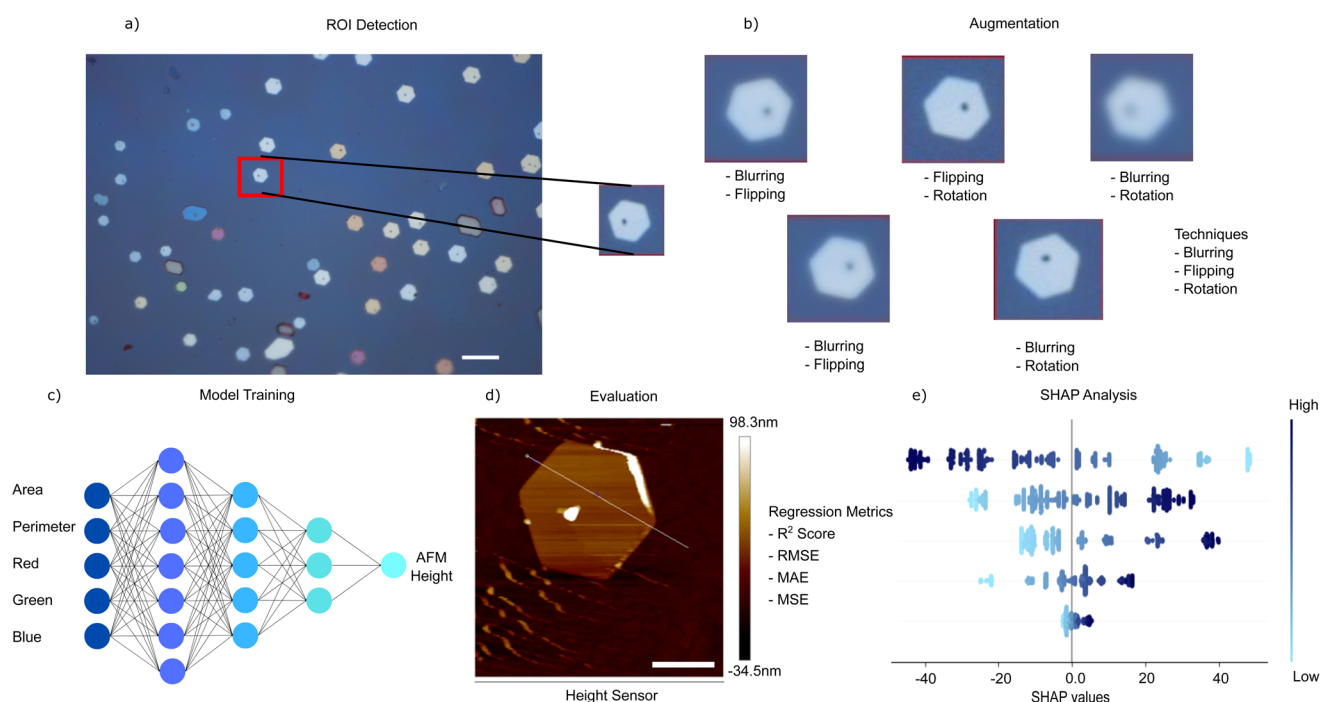
Fig. 1 Overview of the proposed AI pipeline for height prediction from OM images. (a) OM images of In_2Se_3 with a manually selected ROI (red box). (b) Augmented ROI examples used to increase the training robustness. (c) Model training using morphological and RGB features. (d) Evaluation against AFM-measured ground truth using standard regression metrics. (e) SHAP analysis showing feature contributions to predicted AFM height. Scale bars: a, 10 μm; d, 2 μm.

To improve the model robustness and generalizability, the dataset was expanded using data augmentation techniques, such as flipping, rotation, and blurring (Fig. 1b). Morphological features (area and perimeter) and RGB intensities were extracted and used to train multiple regression models, including Extreme Gradient Boosting Regressor (XGBR), RFR, Support Vector Regressor (SVR), Linear Regressor (LR), Ridge Regressor (RR), and a shallow neural network, MLP. Feature scaling was applied using StandardScaler for morphological features and MinMaxScaler for RGB values to ensure consistent input ranges during training (Fig. 1c). The dataset was split into 60% training and 40% testing sets, and a random seed of 42 was set during model training. The model performance was assessed through R^2, MAE, MSE, and RMSE metrics (Fig. 1d). To interpret the model's decision-making process, SHAP analysis was performed, revealing the relative contribution of each feature to the AFM height prediction and supporting model interpretability in the context of thin-film interference (Fig. 1e). SHAP analysis applies Shapley values, derived from game theory, to deliver both local and global insights into feature importance. Beyond capturing global feature interactions, it also ensures consistent and reliable feature attribution throughout the analysis.[47]

Among the trained regression models, RFR demonstrated the best performance based on standard regression metrics (SI Table S1). The results indicate that LR and RR performed poorly, even during the training phase, with R^2 scores below 0.5, suggesting that these models could explain less than half of the variance in the data. This poor performance may be attributed to the simplicity of these models, which likely limits their ability to capture complex, non-linear relationships among the features. However, only RFR maintained consistent results between the training and testing phases, suggesting no signs of overfitting. Both XGBR and SVR, on the other hand, exhibited a significant performance drop in the testing phase, indicating overfitting.

After applying Grid Search Cross Validation (CV) and Randomized Search CV, respectively, to optimize model performance,[48] the training performance of XGBR, RFR, and SVR improved compared to their default settings (SI Tables S2 and S3). Despite the improvement in training metrics, XGBR and SVR still exhibit clear signs of overfitting. The testing performance remained significantly lower than their training performance, indicating that the models failed to generalize well to unseen data. As such, both models were deemed unsuitable for further implementation for height prediction. For the RFR model, while training performance slightly improved after hyperparameter tuning, the performance on the testing dataset decreased, with the R^2 score dropping from 0.997 (default) to 0.976 (Grid Search CV tuned) and 0.949 (Randomized Search tuned). This suggests that the default RFR model generalized better than the tuned version despite having slightly lower training accuracy. The smaller gap between training and testing performances in the default RFR model supports its superior generalization ability.

Overfitting is a common problem in ML, especially when working with limited datasets. It occurs when the model tends to memorize the training data rather than learning the underlying patterns or relationships.[49] Consequently, the model performs well on training data but fails to generalize to unseen data, leading to poor performance during testing.[50] For the MLP, the relatively small dataset used in this study necessitated careful tuning to prevent overfitting (SI Table S4). The MLP was selected because, unlike conventional machine learning algorithms, it incorporates interconnected layers that can capture complex relationships among features. At the same time, the network was kept relatively shallow, making it more suitable for this application by reducing the risk of over-interpreting the relationships between input features. The model's best performance was achieved when the number of neurons in each layer decreased from 512 to 64, the learning rate was 0.001, the weight decay was 0.001, and the cross-validation fold was set to 5 (model 6). It showed an average R^2 of 0.973 across the cross-validation folds, with an MSE of 45.996 nm^2, an MAE of 5.743 nm, and an RMSE of 6.782 nm. These values indicate strong model performance in predicting the AFM height. The model's generalization ability was further validated using the test dataset, achieving an R^2 of 0.947, an MSE of 34.580 nm^2, an MAE of 4.696 nm, and an RMSE of 5.881 nm. Since these values remained consistent with the training results, it confirmed that no overfitting occurred, demonstrating the model's robustness in AFM height prediction. Although some models, such as models 7 and 13, also showed no signs of overfitting, their predictive performance was inferior to that of the selected model, particularly in terms of MAE, MSE, and RMSE. As a result, model 6 was identified as the most reliable and was selected for further exploration.

Model performance comparison: RFR *vs.* MLP

To evaluate the predictive performance of the trained models, predicted AFM height values were plotted against the corresponding ground truth, as shown in Fig. 2. In Fig. 2a, the RFR demonstrates strong performance, with all 14 validation samples closely aligning with the ideal diagonal ($y = x$). This indicates high prediction accuracy and low variance across the threshold-based validation set. In contrast, Fig. 2c displays the performance of the MLP model trained using cross-validation, resulting in a denser and more dispersed distribution of data points. Although the MLP shows a greater spread compared to the RFR, a substantial number of predictions still fall near the perfect fit line, suggesting reasonable accuracy and generalization.

According to SHAP analysis for model interpretability,[51] as illustrated in Fig. 2b, the RFR model ranked area as the most influential feature, followed by perimeter and RGB intensity features. However, the relatively low mean SHAP values (<1.0) suggest that RFR does not heavily rely on any single feature, suggesting a more distributed learning pattern and possibly robust learning behaviour. The dispersion of SHAP values in the summary plot (Fig. 2b) further implies non-linear and inconsistent feature interactions. In contrast, the MLP model
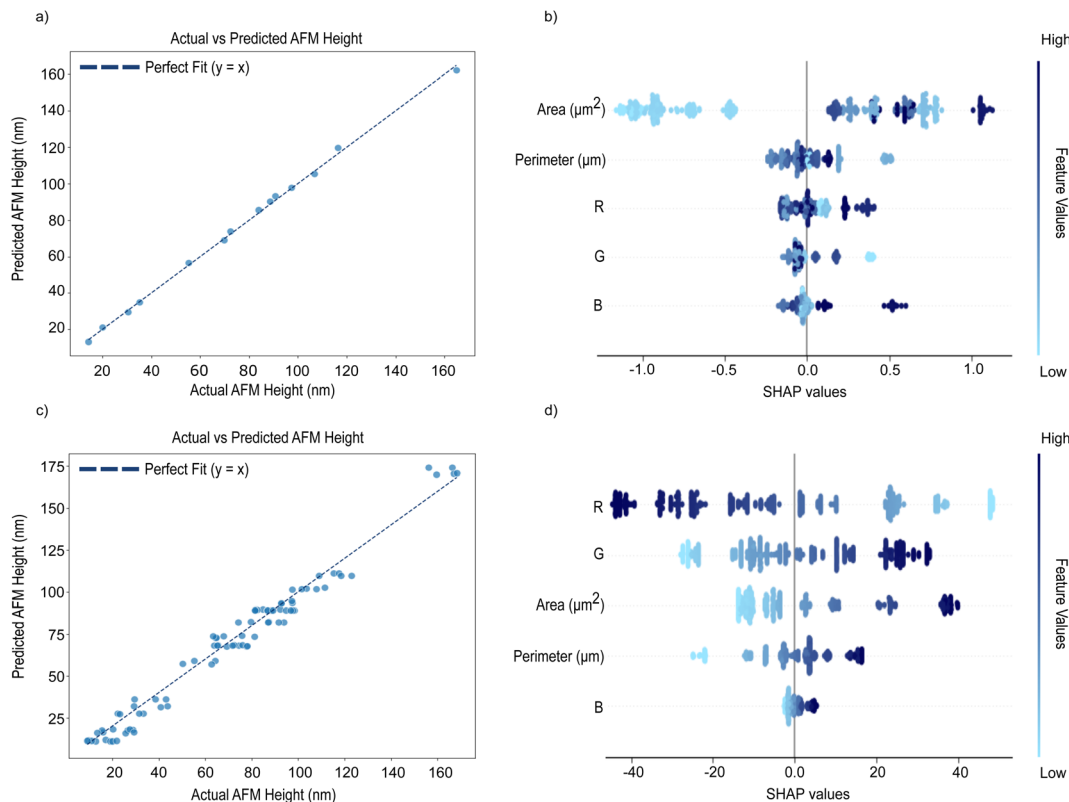
Fig. 2 Performance analysis and comparison between RFR and MLP models. (a) Predicted *vs.* actual AFM-measured heights using RFR on the validation set. (b) SHAP summary plot for RFR, showing feature impact on individual predictions. (c) Predicted *vs.* actual AFM-measured heights using MLP under cross-validation. (d) SHAP summary plot for MLP.

exhibits clearer and more concentrated feature dependencies (Fig. 2d). Red intensity emerged as the most influential factor, followed by green intensity, area, blue intensity, and perimeter. The summary plot (Fig. 2d) highlights red and green intensities as the most influential features in AFM height prediction. While their exact directional effects vary, both channels show stronger contributions compared to other features, supporting their importance in the model's decision-making. The dominance of red intensity indicates a strong feature-to-output correlation, consistent with thin-film interference effects in OM imaging. In short, the MLP exhibits more pronounced and interpretable feature contributions than RFR, aligning well with the underlying physical principles. This reinforces the suitability of MLP for AFM height prediction based on OM features, especially when interpretability and physical relevance are critical.

Evaluation of automated height prediction

While the thresholding method required region-specific tuning, limiting scalability, the Cellpose model[52] was introduced for segmentation as an automated contour detection method for RFR and MLP models to evaluate generalizability. According to the statistical tests, for the RFR model on the test dataset ($n = 4$), the Shapiro–Wilk test confirmed the normal distribution of the prediction differences ($p > 0.05$), allowing a

paired *t*-test, which indicated no statistically significant difference between threshold-based and automated approaches ($p > 0.05$). However, when applied to the full dataset ($n = 20$), comprising the original 14 training samples, 4 held-out test samples, and an additional 2 completely unseen random test images, the Shapiro–Wilk test revealed non-normality ($p < 0.05$), prompting the use of the Wilcoxon signed-rank test, which again showed no statistically significant difference ($p > 0.05$). Paired *t*-test analysis across the full dataset further supported this conclusion, with small Hedges' *g* values between the original and automated methods, indicating only minor differences attributable to the segmentation method. Despite this statistical consistency, a drop in performance ($R^2 = 0.584$, MSE = 583.838 nm^2, RMSE = 24.16 nm, and MAE = 13.266 nm) showed the RFR model was sensitive to segmentation method variations, as minor shifting affected the model performance, aligning with the SHAP analysis, which demonstrated low overall feature dependence.

For the MLP model, the Shapiro–Wilk test indicated non-normality of prediction differences between segmentation methods ($p < 0.05$), leading to Wilcoxon signed-rank testing, which showed no significant difference ($p > 0.05$). Across both the test and full datasets, predictions from both segmentation methods were statistically consistent with ground truth ($p > 0.05$) as confirmed by paired *t*-tests. Effect size analysis sup-

ported this finding, with only a small Hedges' *g* value, indicating that the automated method produces measurements highly comparable to the original baseline. The MLP's strong regression performance (R^2 = 0.978, MSE = 30.112 nm^2, RMSE = 5.487 nm, and MAE = 3.820 nm) confirms high predictive accuracy and generalization to automated segmentation, without signs of overfitting, corroborating the SHAP analysis, which revealed the MLP model's strong reliance on specific features, particularly red and green intensities, followed by area, allowing the MLP model to maintain predictive stability, even when segmentation-induced variations.

Collectively, these results highlight the practical advantage of the automated segmentation workflow, which achieves prediction performance statistically indistinguishable from the threshold-based approach while offering greater scalability and consistency. Due to the limited sample size, evaluations were conducted on the full dataset to prevent a fully independent generalization assessment. Nevertheless, consistency across statistical tests, effect-size analysis, and SHAP-based interpretability provides strong evidence that the automated method can reliably replace threshold segmentation in this context. Using the MLP as the baseline, the automated workflow generates predictions in approximately 1 second, allowing for high-throughput applications. Additional evaluation on the held-out validation data confirmed that the model generalizes effectively to unseen data, supporting its use for robust and scalable thickness prediction. The segmented outcome of the augmented images (see SI Fig. 1) shows slight variations in the segmentation masks, which can lead to minor differences in predicted values. However, because the masks are largely consistent, these differences do not result in statistically significant deviations from the ground truth.

Thin film interference theory

Focusing on the MLP model, which demonstrated strong generalization and robustness, the SHAP summary plot (Fig. 3b) offers insight into the model's decision-making process. Red intensity emerges as the most influential feature for AFM height prediction, followed by green intensity, area, blue intensity, and perimeter. This ranking aligns with the general trend observed in Fig. 3a, where particles exhibiting higher AFM heights tend to display lower red intensity and higher green intensity. The SHAP value distribution further supports that higher feature values (shown in dark blue) positively influence predictions, whereas lower values (in blue) have a negative impact. An interesting case can be observed in Fig. 3a, a particle with an AFM height of 30.429 nm shows relatively high RGB intensities. Rather than contradicting the overall SHAP trends, as shown in Fig. 3b, this case exemplifies the model's capability to learn multi-feature interactions, where the influence of one feature is contextually adjusted based on the values of others. This interaction-aware behavior highlights the model's flexibility and supports its predictive consistency. The SHAP distribution also revealed that area and perimeter contribute positively to height prediction, suggesting that larger or more geometrically spread particles are often associ-
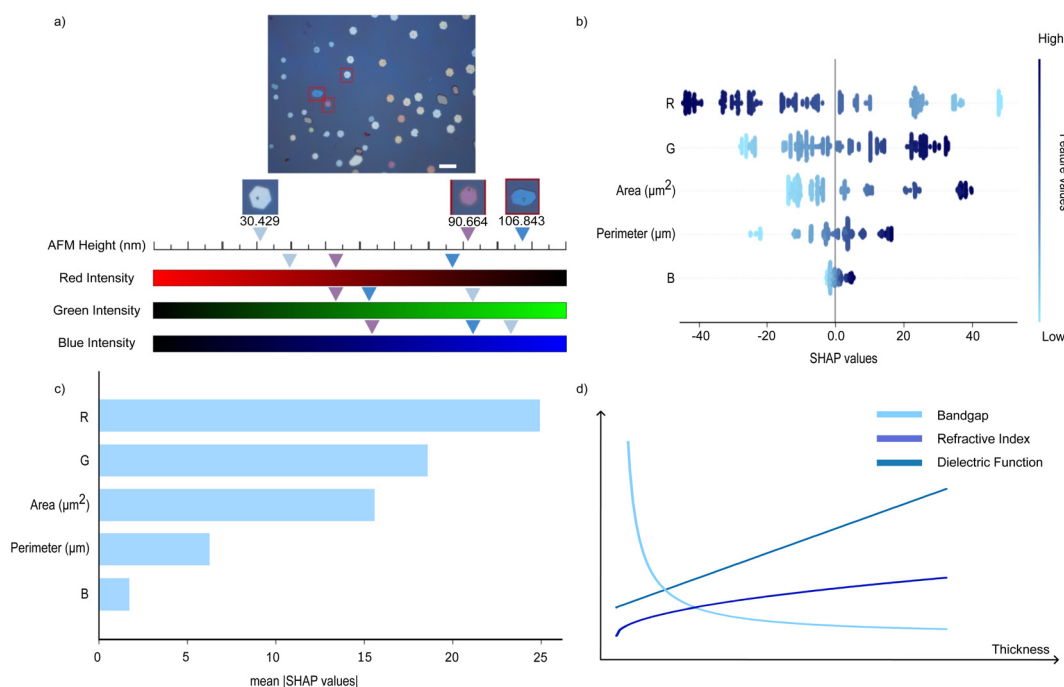
Fig. 3 Interpretability analysis of the MLP model aligned with thin-film interference principles. (a) OM image examples and corresponding AFM heights illustrating the relationship between particle features and predicted thickness. (b) SHAP summary plot showing the impact of each feature on the MLP model's output. (c) Mean absolute SHAP values indicating the overall feature importance. (d) Schematic showing theoretical trends of optical parameters (bandgap, refractive index, and dielectric function) with increasing thickness based on thin-film interference. Scale bar: 10 μm.

ated with increased thickness. Fig. 3c further quantifies each feature's impact, confirming the dominant role of red intensity. Overall, SHAP analysis affirms the relevance of both morphological and colour-based cues, providing a transparent and interpretable pathway for translating OM-derived features into accurate AFM height predictions.

As In_2Se_3 exhibits ferroelectricity, its spontaneous polarisation alters the internal electronic distribution, directly impacting the material's complex dielectric function, $\varepsilon = \varepsilon_1 + i\varepsilon_2$, where ε_1 and ε_2 are the real and imaginary parts, respectively. Given that the refractive index, n, is related to the dielectric function by

$$n = \sqrt{\frac{|\varepsilon| + \varepsilon_1}{2}}$$

for real n in simple cases, or more generally, $n + ik = \sqrt{\varepsilon}$. Any modification in the dielectric function due to polarization can affect the refractive index. It has also been proven that the dielectric constant increases monotonically with the number of layers and saturates at the bulk value around eight layers, indicating a thickness-dependent dielectric response.[24]

This thickness-dependent optical behaviour, combined with the material's anisotropy, is crucial for understanding its visual appearance. In_2Se_3 is an anisotropic material, meaning its optical properties vary with direction.[53] One manifestation is birefringence, where an incident light ray splits into two rays with different velocities and polarizations. Fresnel's law describes how reflection and transmission at an interface depend on both the angle and polarization, with the overall complex reflection coefficient for a thin film system given by:

$$r = \frac{r_{12} + r_{23}e^{2i\beta}}{1 + r_{12}r_{23}e^{2i\beta}}$$

Here, r_{12} and r_{23} are the complex amplitude reflection coefficients at the top and bottom interfaces, respectively (which are polarization- and angle-dependent), and $e^{2i\beta}$ accounts for the phase accumulation across the film, with

$$\beta = \frac{2\pi nd \cos \theta_r}{\lambda}.$$

While the ferroelectric polarization and anisotropy-driven effects modulate the reflected light intensity, the primary factor contributing to the significant colour contrast observed in the OM images of In_2Se_3 is the dominant thin-film interference arising from its thickness variation. This interference originates from light reflections at the top and bottom surfaces of the thin film, leading to constructive or destructive interference depending on the optical path difference, governed by: $2nd \cos \theta = m\lambda$ where n is the refractive index, d is the film thickness, θ is the refraction angle of the light within the film, m is the order of interference, and λ is the wavelength. As a result, regions of differing thicknesses appear as different colours under OM.[40,41] This makes OM a fast, non-destructive, and widely accessible way for estimating the thickness of 2D materials, aligning with the core objective of this study. This contrast-based method has been previously validated for

thickness estimation. For instance, it was successfully applied to determine the number of layers in graphene using its refractive index.[54]

Consistent with findings, uniform colour in OM images reflects uniform thickness, as explained by thin-film interference theory; regions of equal thickness produce the same interference conditions and, thus, reflect similar colour.[27] It is further demonstrated that thinner regions appear orange (lighter), while intermediate regions appear blue (darker), forming a colour gradient consistent with varying optical path lengths. Thicker regions often appear white due to broadband constructive interference and additional scattering effects.[55] Complementing these optical insights, a positive correlation between PL intensity and In_2Se_3 thickness was reported, suggesting that both colour contrasts in OM and PL response serve as indirect indicators of thickness, shaped by the nanoscale geometry and optical interactions.[56]

Generally, AFM-observed higher thickness corresponds to a higher refractive index and a lower bandgap, factors that influence the observed colour *via* thin-film interference. The refractive index and bandgap are inversely related, as described by the Moss relation, $n^4 \times E_g$ = constant. A higher refractive index implies more densely packed electronic states and, consequently, a reduced bandgap due to quantum confinement effects. The thickness-dependent bandgap of α-phase In_2Se_3 was experimentally confirmed by using electron energy loss spectroscopy (EELS) to demonstrate that the bandgap increases from 1.44 eV at 48 nm thickness to 1.64 eV at 8 nm.[57] Their results, supported by Density Functional Theory (DFT) calculations, are consistent with the quantum confinement model. These interlinked changes, thickness, refractive index, and bandgap directly affect the OM colour contrast, making it a valuable proxy for assessing both the thickness and electronic structure. Fig. 3d illustrates how these properties vary with thickness. These observations are reinforced by SHAP analysis, which identified red and green colour intensities as the most influential features in predicting the AFM-measured height. This supports the conclusion that OM contrast, particularly colour, is significantly informative of surface topography.

Building on the Fresnel interference theory, the observed optical contrast in In_2Se_3 is closely tied to its thickness-dependent dielectric behaviour. As a ferroelectric material, In_2Se_3 exhibits spontaneous polarization, which alters its internal electronic distribution and modulates its complex dielectric function. This directly affects the refractive index and consequently the optical contrast observed in OM images. Such contrast is not merely a visual artifact but reflects real changes in the electronic structure, enabling thickness estimation through colour-based analysis.

Moreover, the dielectric constant of In_2Se_3 has been shown to increase monotonically with the number of layers, saturating at the bulk value beyond eight layers. This indicates a strong correlation between thickness and intrinsic properties such as phase stability, interlayer coupling, and ferroelectric switching behaviour. The proposed AI-based approach, by

providing continuous height values rather than discrete layer classification, enables a more nuanced analysis of these thickness-dependent properties. This not only aids in rapid and non-destructive thickness estimation but also facilitates phase identification and material characterization, which are crucial for optimizing the use of In_2Se_3 in memory devices, optoelectronics, and other semiconductor applications.

To enhance the generalizability of the trained AI model across OM images captured under varying lightning conditions, histogram matching was applied as a preprocessing step to unseen images. This technique adjusts the RGB intensity distribution of the new images to match that of the training dataset. Such correction is particularly critical in OM-based thickness estimation, where contrast variations arising from thin-film interference serve as key predictive features. Differences in illumination or imaging parameters can alter these colour cues, potentially leading to inaccurate predictions. By standardizing the colour distribution, histogram matching improves the consistency of feature extraction and supports more reliable thickness estimation across diverse imaging scenarios. SI Fig. 2 and 3 illustrate the application of this correction under different brightness and lighting conditions, respectively, demonstrating the enhanced robustness and generalization capability of the trained model.

Conclusion

In conclusion, this study proposed an automated, high-throughput pipeline for estimating the thickness of 2D materials, using In_2Se_3 as the case study. Despite being trained on a small dataset, the MLP achieved strong predictive accuracy, achieving R^2 = 0.973, MSE = 45.996 nm^2, MAE = 5.743 nm, and RMSE = 6.782 nm, and R^2 = 0.978, MSE = 30.112 nm^2, MAE = 3.820 nm, and RMSE = 5.487 nm, for thresholding and automated methods, respectively, for segmentation. The remarkable robustness and generalizability, even across different segmentation methods, were supported by statistical validation. This highlights the effectiveness of the proposed approach in low-data scenarios. SHAP analysis further revealed that red and green intensities were the most influential features in height prediction. This finding aligns with thin-film interference theory, where colour contrast in OM reflects variations in 2D material thickness. The ROI is determined through a simple rectangular crop, with the model automatically extracting the relevant area for evaluation, ensuring both ease of use and high throughput. Overall, the proposed pipeline offers a reliable, fast, and non-destructive alternative to conventional AFM-based methods, making it highly suitable for efficient 2D material characterization, even with limited training data.

Future work could extend this approach to a broader range of 2D materials and incorporate alternative loss functions or physics-informed regularization strategies, such as physics-informed neural networks, to further improve both understanding and predictive accuracy. Since the current model was trained on a limited dataset and relies solely on AFM and OM

imaging, future studies could benefit from expanding the dataset and expanding the model across different material systems or varying imaging conditions to better assess its generalizability. Such expansion would help evaluate the model's robustness against variations in surface morphology, imaging artifacts, and sample preparation procedures. Building on the strong foundation established in this work, the proposed pipeline has significant potential for integration into high-throughput experimental workflows, rapid quality control during material synthesis, and scalable industrial applications where non-destructive and automated characterization, particularly for thickness estimation, is crucial.

Experimental section

Sample preparation

For the chemical vapour deposition (CVD) synthesis, 50 mg of indium chloride ($InCl_3$, 99.999%, Alfa Aesar) and 150 mg of selenium (Se, 99+%, Alfa Aesar) were thoroughly mixed in a quartz boat. The mica was utilized as the growth substrate. Prior to heating, the CVD system was purged with argon (Ar) gas to remove residual air and moisture. Subsequently, a gas mixture of hydrogen (H_2) and argon (Ar) was introduced into the reaction chamber. The synthesis followed a typical oxidation–reduction reaction mechanism:

$$2InCl_3 + 3Se + 6H_2 = In_2Se_3 + 6HCl \uparrow.$$
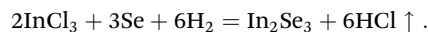
After the growth process, the system was naturally cooled to room temperature. The resulting samples were attached to a SiO_2/Si substrate (the thickness of SiO_2 = 285 nm) for optical characterization. A Nikon optical microscope (ECLIPSE LV100D) was used for morphological observation. The thickness of In_2Se_3 was measured using AFM with a Bruker Dimension Icon system.

Model training

Fourteen OM-AFM image pairs were used for model training, with 30 augmentations per image, resulting in 420 samples. Contours were extracted using Otsu's thresholding, and features, area, perimeter, and RGB intensities were computed. Pixel measurements were converted to nanometres using scale bar calibration. To account for measurement variability inherent to AFM, a ±10 nm tolerance was applied. This tolerance was determined based on the measurement uncertainty across the material surface, as quantified using NanoScope Analysis software (SI Fig. 4).

Features were standardized or MinMax scaled based on type. Five conventional regressors (XGBR, RFR, SVR, LR, RR) were tuned through Grid Search CV. The MLP included ReLU activations, dropout, and batch normalization, and was optimized using the Adam optimizer with MSE loss.

Evaluation

Performance was assessed using standard regression metrics (R^2, RMSE, MSE, and MAE). SHAP was applied to interpret the

trained model's predictions, with the results analysed in the context of thin-film interference theory to validate the physical relevance of the identified feature contributions.

Author contributions

K. W. L. and X. Z. supervised and led the research project. M. F., P. S. Q. Y., and Z. Q. provided insights throughout the project and assisted in reviewing the manuscript. Z. H. carried out the AFM and OM imaging and contributed to manuscript editing. J. C. N. conducted the model training and contributed to manuscript writing. All authors discussed the results and contributed to the final version of the manuscript.

Conflicts of interest

There are no conflicts to declare.

Data availability

Data for this article are available upon request.

Supplementary information: discussion on the hyperparameter tuning during model training, the application of histogram matching as a correction step, and its effect on the model performance. See DOI: https://doi.org/10.1039/d5nr03320a.

Acknowledgements

References

1 N. Briggs, *et al.*, A roadmap for electronic grade 2D materials, *2D Mater.*, 2019, **6**, DOI: 10.1088/2053-1583/aaf836.

2 T. Dutta, *et al.*, Electronic properties of 2D materials and their junctions, *Nano Mater. Sci.*, 2024, **6**, 1–23, DOI: 10.1016/j.nanoms.2023.05.003.

3 N. Baig, Two-dimensional nanomaterials: A critical review of recent progress, properties, applications, and future directions, *Composites, Part A*, 2023, **165**, DOI: 10.1016/j.compositesa.2022.107362.

4 J. W. You, S. R. Bongu, Q. Bao and N. C. Panoiu, Nonlinear optical properties and applications of 2D materials: theoretical and experimental aspects, *Nanophotonics*, 2019, **8**, 63–97, DOI: 10.1515/nanoph-2018-0106.

5 D. Akinwande, *et al.*, A review on mechanics and mechanical properties of 2D materials-Graphene and beyond, *Extreme Mech. Lett.*, 2017, **13**, 42–77, DOI: 10.1016/j.eml.2017.01.008.

6 C. Androulidakis, K. H. Zhang, M. Robertson and S. Tawfick, Tailoring the mechanical properties of 2D materials and heterostructures, *2D Mater.*, 2018, **5**, DOI: 10.1088/2053-1583/aac764.

7 K. S. Novoselov, A. Mishchenko, A. Carvalho and A. H. C. Neto, 2D materials and van der Waals heterostructures, *Science*, 2016, **353**, DOI: 10.1126/science.aac9439.

8 M. C. Lemme, L. J. Li, T. Palacios and F. Schwierz, Two-dimensional materials for electronic applications, *MRS Bull.*, 2014, **39**, 711–718, DOI: 10.1557/mrs.2014.138.

9 M. Chhowalla, *et al.*, The chemistry of two-dimensional layered transition metal dichalcogenide nanosheets, *Nat. Chem.*, 2013, **5**, 263–275, DOI: 10.1038/nchem.1589.

10 X. Y. Zhang, L. L. Hou, A. Ciesielski and P. Samorì, 2D Materials Beyond Graphene for High-Performance Energy Storage Applications, *Adv. Energy Mater.*, 2016, **6**, DOI: 10.1002/aenm.201600671.

11 Y. F. Shi, N. T. Duong and K. W. Ang, Emerging 2D materials hardware for in-sensor computing, *Nanoscale Horiz.*, 2025, **10**, 205–229, DOI: 10.1039/d4nh00405a.

12 J. H. Kim, J. H. Jeong, N. Kim, R. Joshi and G. H. Lee, Mechanical properties of two-dimensional materials and their applications, *J. Phys. D:Appl. Phys.*, 2019, **52**, DOI: 10.1088/1361-6463/aaf465.

13 X. L. Liu and M. C. Hersam, 2D materials for quantum information science, *Nat. Rev. Mater.*, 2019, **4**, 669–684, DOI: 10.1038/s41578-019-0136-x.

14 M. Turunen, *et al.*, Quantum photonics with layered 2D materials, *Nat. Rev. Phys.*, 2022, **4**, 219–236, DOI: 10.1038/s42254-021-00408-0.

15 X. D. Duan, C. Wang, A. L. Pan, R. Q. Yu and X. F. Duan, Two-dimensional transition metal dichalcogenides as atomically thin semiconductors: opportunities and challenges, *Chem. Soc. Rev.*, 2015, **44**, 8859–8876, DOI: 10.1039/c5cs00507h.

16 S. Haastrup, *et al.*, The Computational 2D Materials Database: high-throughput modeling and discovery of atomically thin crystals, *2D Mater.*, 2018, **5**, DOI: 10.1088/2053-1583/aacfc1.

17 X. Cong, X. L. Liu, M. L. Lin and P. H. Tan, Application of Raman spectroscopy to probe fundamental properties of two-dimensional materials, *npj 2D Mater. Appl.*, 2020, **4**, DOI: 10.1038/s41699-020-0140-4.

18 S. S. Zhang, *et al.*, Spotting the differences in two-dimensional materials - the Raman scattering perspective, *Chem. Soc. Rev.*, 2018, **47**, 3217–3240, DOI: 10.1039/c7cs00874k.

19 C. F. Holder and R. E. Schaak, Tutorial on Powder X-ray Diffraction for Characterizing Nanoscale Materials, *ACS Nano*, 2019, **13**, 7359–7365, DOI: 10.1021/acsnano.9b05157.

20 Y. B. Jiang, *et al.*, Simulating Powder X-ray Diffraction Patterns of Two-Dimensional Materials, *Inorg. Chem.*, 2018, **57**, 15123–15132, DOI: 10.1021/acs.inorgchem.8b02315.

21 S. Lopatin, *et al.*, Aberration-corrected STEM imaging of 2D materials: Artifacts and practical applications of threefold astigmatism, *Sci. Adv.*, 2020, **6**, DOI: 10.1126/sciadv.abb8431.

22 R. G. Mendes, *et al.*, Electron-Driven In Situ Transmission Electron Microscopy of 2D Transition Metal Dichalcogenides and Their 2D Heterostructures, *ACS Nano*, 2019, **13**, 978–995, DOI: 10.1021/acsnano.8b08079.

23 L. F. Fei, *et al.*, Direct TEM observations of growth mechanisms of two-dimensional MoS2 flakes, *Nat. Commun.*, 2016, **7**, DOI: 10.1038/ncomms12206.

24 D. Wu, *et al.*, Thickness-Dependent Dielectric Constant of Few-Layer In2Se3 Nanoflakes, *Nano Lett.*, 2015, **15**, 8136–8140, DOI: 10.1021/acs.nanolett.5b03575.

25 X. Tao and Y. Gu, Crystalline-Crystalline Phase Transformation in Two-Dimensional In2Se3 Thin Layers, *Nano Lett.*, 2013, **13**, 3501–3505, DOI: 10.1021/nl400888p.

26 Y. T. Huang, *et al.*, Two-dimensional In2Se3: A rising advanced material for ferroelectric data storage, *InfoMat*, 2022, **4**, DOI: 10.1002/inf2.12341.

27 K. P. Si, *et al.*, Quasi-equilibrium growth of inch-scale single-crystal monolayer α-In2Se3 on fluor-phlogopite, *Nat. Commun.*, 2024, **15**, DOI: 10.1038/s41467-024-51322-9.

28 C. H. Ho, Amorphous effect on the advancing of wide-range absorption and structural-phase transition in γ-In2Se3 polycrystalline layers, *Sci. Rep.*, 2014, **4**, DOI: 10.1038/srep04764.

29 N. Balakrishnan, *et al.*, Quantum confinement and photo-responsivity of β-In2Se3 nanosheets grown by physical vapour transport, *2D Mater.*, 2016, **3**, DOI: 10.1088/2053-1583/3/2/025030.

30 S. Wu, *et al.*, Progress on mechanical and tribological characterization of 2D materials by AFM force spectroscopy, *Friction*, 2024, **12**, 2627–2656, DOI: 10.1007/s40544-024-0864-9.

31 R. Xu, *et al.*, Advanced atomic force microscopies and their applications in two-dimensional materials: a review, *Mater. Futures*, 2022, **1**, DOI: 10.1088/2752-5724/ac8aba.

32 Y. H. Li, *et al.*, Mapping the elastic properties of two-dimensional MoS2 via bimodal atomic force microscopy and finite element simulation, *npj Comput. Mater.*, 2018, **4**, DOI: 10.1038/s41524-018-0105-8.

33 P. C. Uzoma, *et al.*, AFM: An important enabling technology for 2D materials and devices, *Nanotechnol. Rev.*, 2025, **14**, DOI: 10.1515/ntrev-2025-0154.

34 B. de la Torre, *et al.*, Atomic-Scale Variations of the Mechanical Response of 2D Materials Detected by Noncontact Atomic Force Microscopy, *Phys. Rev. Lett.*, 2016, **116**, DOI: 10.1103/PhysRevLett.116.245502.

35 F. Gołek, P. Mazur, Z. Ryszka and S. Zuber, AFM image artifacts, *Appl. Surf. Sci.*, 2014, **304**, 11–19, DOI: 10.1016/j.apsusc.2014.01.149.

36 N. Gavara, Combined strategies for optimal detection of the contact point in AFM force-indentation curves obtained on thin samples and adherent cells, *Sci. Rep.*, 2016, **6**, DOI: 10.1038/srep21267.

37 T. Ando, T. Uchihashi and N. Kodera, High-speed AFM and applications to biomolecular systems, *Annu. Rev. Biophys.*, 2013, **42**, 393–414, DOI: 10.1146/annurev-biophys-083012-130324.

38 M. Petrov, *et al.*, Mechanical spectroscopy of materials using atomic force microscopy (AFM-MS), *Mater. Today*, 2024, **80**, 218–225, DOI: 10.1016/j.mattod.2024.08.021.

39 N. Oinonen, *et al.*, Electrostatic Discovery Atomic Force Microscopy, *ACS Nano*, 2022, **16**, 89–97, DOI: 10.1021/acsnano.1c06840.

40 S. Puebla, A. Mariscal-Jiménez, R. S. Galán, C. Munuera and A. Castellanos-Gomez, Optical-Based Thickness Measurement of MoO3 Nanosheets, *Nanomaterials*, 2020, **10**, 1272, DOI: 10.3390/nano10071272.

41 P. Gant, *et al.*, Optical contrast and refractive index of natural van der Waals heterostructure nanosheets of franckeite, *Beilstein J. Nanotechnol.*, 2017, **8**, 2357–2362, DOI: 10.3762/bjnano.8.235.

42 Y. H. Li, *et al.*, Rapid identification of two-dimensional materials via machine learning assisted optic microscopy, *J. Materiomics*, 2019, **5**, 413–421, DOI: 10.1016/j.jmat.2019.03.003.

43 J. T. Yang and H. M. Yao, Automated identification and characterization of two-dimensional materials via machine learning-based processing of optical microscope images, *Extreme Mech. Lett.*, 2020, **39**, DOI: 10.1016/j.eml.2020.100771.

44 S. Masubuchi and T. Machida, Classifying optical microscope images of exfoliated graphene flakes by data-driven machine learning, *npj 2D Mater. Appl.*, 2019, **3**, DOI: 10.1038/s41699-018-0084-0.

45 S. Mahjoubi, F. Ye, Y. Bao, W. A. Meng and X. Zhang, Identification and classification of exfoliated graphene flakes from microscopy images using a hierarchical deep convolutional neural network, *Eng. Appl. Artif. Intell.*, 2023, **119**, DOI: 10.1016/j.engappai.2022.105743.

46 L. Zhu, *et al.*, Artificial Neuron Networks Enabled Identification and Characterizations of 2D Materials and van der Waals Heterostructures, *ACS Nano*, 2022, **16**, 2721–2729, DOI: 10.1021/acsnano.1c09644.

47 J. Su, *et al.*, Intelligent synthesis of magnetic nanographenes via chemist-intuited atomic robotic probe, *Nat. Synth.*, 2024, **3**, 466–476, DOI: 10.1038/s44160-024-00488-7.

48 D. Krstajic, L. J. Buturovic, D. E. Leahy and S. Thomas, Cross-validation pitfalls when selecting and assessing regression and classification models, *J. Cheminf.*, 2014, **6**, DOI: 10.1186/1758-2946-6-10.

49 Y. Yang, X. Xia, D. Lo and J. Grundy, A survey on deep learning for software engineering, *ACM Comput. Surv.*, 2022, **54**, 1–73, DOI: 10.1145/3505243.

50 V. Aceña, I. M. de Diego, R. R. Fernández and J. M. Moguerza, Minimally overfitted learners: A general framework for ensemble learning, *Knowl.-Based Syst.*, 2022, **254**, DOI: 10.1016/j.knosys.2022.109669.

51 L. Antwarg, R. M. Miller, B. Shapira and L. Rokach, Explaining anomalies detected by autoencoders using Shapley Additive Explanations, *Expert Syst. Appl.*, 2021, **186**, DOI: 10.1016/j.eswa.2021.115736.

52 C. Stringer, T. Wang, M. Michaelos and M. Pachitariu, Cellpose: a generalist algorithm for cellular segmentation, *Nat. Methods*, 2021, **18**, 100–106, DOI: 10.1038/s41592-020-01018-x.

53 S. Y. Wang, *et al.*, Strong Anisotropic Two-Dimensional In2Se3 for Light Intensity and Polarization Dual-Mode High-Performance Detection, *ACS Appl. Mater. Interfaces*, 2023, **15**, 3357–3364, DOI: 10.1021/acsami.2c19660.

54 Z. Ni, *et al.*, Graphene thickness determination using reflection and contrast spectroscopy, *Nano Lett.*, 2007, **7**, 2758–2763, DOI: 10.1021/nl071254m.

55 J. M. Czerniawski and J. L. Stickney, Electrodeposition of In2Se3 Using Potential Pulse Atomic Layer Deposition, *J. Phys. Chem. C*, 2016, **120**, 16162–16167, DOI: 10.1021/acs.jpcc.6b00320.

56 Z. J. Wang, M. Wang, H. Y. Nan, J. Bai and C. L. Wang, Effect of thickness on optical properties of InSe/In2Se3 heterojunction, *AIP Adv.*, 2024, **14**, DOI: 10.1063/5.0222672.

57 F. J. Lyu, *et al.*, Thickness-dependent band gap of α-In2Se3: from electron energy loss spectroscopy to density functional theory calculations, *Nanotechnology*, 2020, **31**, DOI: 10.1088/1361-6528/ab8998.