

Image segmentation of exfoliated two-dimensional materials by generative adversarial network-based data augmentation

Xiaoyu Cheng(程晓昱), Chenxue Xie(解晨雪), Yulun Liu(刘宇伦), Ruixue Bai(白瑞雪), Nanhai Xiao(肖南海), Yanbo Ren(任琰博), Xilin Zhang(张喜林), Hui Ma(马惠), and Chongyun Jiang(蒋崇云)

Citation: Chin. Phys. B, 2024, 33 (**3**): 030703. DOI: 10.1088/1674-1056/ad23d8
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Recent advances in protein conformation sampling by combining machine learning with molecular simulation

Yiming Tang(唐一鸣), Zhongyuan Yang(杨中元), Yifei Yao(姚逸飞), Yun Zhou(周运), Yuan Tan(谈圆), Zichao Wang(王子超), Tong Pan(潘瞳), Rui Xiong(熊瑞), Junli Sun(孙俊力), and Guanghong Wei(韦广红)
Chin. Phys. B, 2024, 33 (**3**): 030701. DOI: 10.1088/1674-1056/ad1a92

Crysformer: An attention-based graph neural network for properties prediction of crystals

Tian Wang(王田), Jiahui Chen(陈家辉), Jing Teng(滕婧), Jingang Shi(史金钢),Xinhua Zeng(曾新华), and Hichem Snoussi
Chin. Phys. B, 2023, 32 (**9**): 090703. DOI: 10.1088/1674-1056/ace247

Differentiable programming and density matrix based Hartree-Fock method

Hong-Bin Ren(任宏斌), Lei Wang(王磊), and Xi Dai(戴希)
Chin. Phys. B, 2021, 30 (**6**): 060701. DOI: 10.1088/1674-1056/abeeed

Accurate Deep Potential model for the Al-Cu-Mg alloy in the full concentration space

Wanrun Jiang(姜万润), Yuzhi Zhang(张与之), Linfeng Zhang(张林峰), and Han Wang(王涵)
Chin. Phys. B, 2021, 30 (**5**): 050706. DOI: 10.1088/1674-1056/abf134

RAPID COMMUNICATION

Image segmentation of exfoliated two-dimensional materials by generative adversarial network-based data augmentation

Xiaoyu Cheng(程晓昱)[1], Chenxue Xie(解晨雪)[1], Yulun Liu(刘宇伦)[1], Ruixue Bai(白瑞雪)[1], Nanhai Xiao(肖南海)[1], Yanbo Ren(任琰博)[1], Xilin Zhang(张喜林)[1], Hui Ma(马惠)[2,†], and Chongyun Jiang(蒋崇云)[1,‡]

[1] *College of Electronic Information and Optical Engineering, Nankai University, Tianjin* 300350, *China*
[2] *School of Physical Science and Technology, Tiangong University, Tianjin* 300387, *China*

(Received 2 January 2024; revised manuscript received 26 January 2024; accepted manuscript online 30 January 2024)

Mechanically cleaved two-dimensional materials are random in size and thickness. Recognizing atomically thin flakes by human experts is inefficient and unsuitable for scalable production. Deep learning algorithms have been adopted as an alternative, nevertheless a major challenge is a lack of sufficient actual training images. Here we report the generation of synthetic two-dimensional materials images using StyleGAN3 to complement the dataset. DeepLabv3Plus network is trained with the synthetic images which reduces overfitting and improves recognition accuracy to over 90%. A semi-supervisory technique for labeling images is introduced to reduce manual efforts. The sharper edges recognized by this method facilitate material stacking with precise edge alignment, which benefits exploring novel properties of layered-material devices that crucially depend on the interlayer twist-angle. This feasible and efficient method allows for the rapid and high-quality manufacturing of atomically thin materials and devices.

Keywords: two-dimensional materials, deep learning, data augmentation, generating adversarial networks

PACS: 07.05.Pj, 68.65.–k, 84.35.+i, 87.64.M– **DOI:** 10.1088/1674-1056/ad23d8

1. Introduction

Atomically thin two-dimensional materials exhibit intriguing physical properties such as valley degree of freedom, single photon emission, strong excitonic effect, etc., which open up a roadmap for the next-generation information devices.[1–4] At present, conventional methods for preparing single- and few-layer two-dimensional materials include mechanical cleavage, liquid phase exfoliation, gas phase synthesis, etc.,[5–7] with the mechanical cleavage method being widely adopted due to its simple preparation procedure and high crystal quality. However, the obtained flakes are random in size, location, shape and thickness due to the uncontrollable interactions between the adhesive tapes and the layered crystals.[8–10] With the optical microscopy, the thickness of the single-, few-layer and bulk flakes can be distinguished upon the optical contrast, which becomes a preliminary method to fabricate two-dimensional material devices. However, determining the thickness of two-dimensional materials with the naked eye is inefficient and requires repetitive work by expert operators, limiting the development of the two-dimensional devices. To save manual effort and improve efficiency, computer vision is being investigated as a substitute for recognizing flakes of different thicknesses.

Traditional rule-based image processing methods, such as edge detection, image color contrast, and threshold segmentation can be cost-effective on condition that multiple adjustable parameters achieve their thresholds simultaneously to obtain the best contrast images, which are inefficient in recognizing thousands of images.[11–16] Meanwhile, deep learning allows computers to distinguish between thin and thick layers automatically, and insensitive to environmental changes, which is important for automation applications. Many contemporary deep learning algorithms, including as object detection, semantic segmentation and instance segmentation, can be employed to recognize two-dimensional material flakes.[17–19] However, a major challenge with deep learning-based approaches is that good performance is heavily dependent on the high quality and vast quantity of dataset used to train the model.[20,21] Meanwhile, a large number of raw images are difficult to obtain due to the low yield of flakes by mechanical cleavage.[22] Data sharing of the microscopic images may extend the training dataset, but the dataset acquired in different conditions of microscopes and cameras cannot be merged directly, making the collaborative efforts much less efficient.[23] Furthermore, the single- and few-layer targets account for a small portion of the total pixels compared with the bulk and the substrate background. This inter-category imbalance leads to a reduced accuracy in the identification of thin layers, which also calls for a large dataset.[24]

In this work, we address the issue of data scarcity by training the StyleGAN3 network to generate synthetic images and expand the dataset. Identifying different thick-

†Corresponding author. E-mail: mahuimoving@163.com
‡Corresponding author. E-mail: jiang.chongyun@nankai.edu.cn

© 2024 Chinese Physical Society and IOP Publishing Ltd

nesses of two-dimensional material is achieved by training the DeepLabv3Plus network. During the training process, semi-supervised mechanism is introduced to generate pseudo-labels, considerably reducing the cost of manual labeling. We enhance the model recognition accuracy to more than 90% using only 128 real images and synthetic images supplemented to the dataset, demonstrating that the addition of synthetic images reduces overfitting and improves recognition accuracy. Our work reduces the limitations imposed by a scarcity of training data for two-dimensional material recognition while improving recognition accuracy, which could help in further exploring the exotic properties of two-dimensional materials and speeding up the manufacture of layered materials devices with low cost.

2. Methods

2.1. Principle and process

We undertake the work of recognizing two-dimensional materials by machine learning with database augmenting in three steps (Fig. 1). With WSe_2 microscopic images as an empirical instance, the initial segment entails the annotation of raw data. Firstly, a total of 161 microscopic images are collected, manually labeled and divided into two groups with 128 (group A) and 33 (group B) images, respectively. Color space transformation and edge detection are employed as preprocessing for more accurate image classification. The 128 images in group A are then utilized in the training of the generative and segmentation networks to produce virtual images and pseudo-labels, whereas the remaining 33 images in group B are used to evaluate the accuracy of the trained model. Using 128 images to train the network is a comprehensive consideration of two factors, the collection difficulties and the high quality of generated images to improve segmentation accuracy. In the second step, StyleGAN3 is employed as the generative network, which is trained to learn the distributional features of the raw images in group A and generates virtual images. Under optimized generation conditions, the created virtual and real images are indistinguishable with the naked eye. In the third step, DeepLabv3Plus operates as the segmentation network (see the rationale and structure of the model in the supporting information section 1 and 2),[27,28] which is trained by the 128 manually labeled images in group A, and then serves to recognize the images generated in the second step and creates corresponding pseudo-labels. The synthetic images and pseudo-labels are sorted by edge detection, with the visually realistic synthetic images and the sharp edge pseudo-labels being used to expand the dataset. In this step, a semi-supervised mechanism[25,26] is used to reduce the labor cost of pixel-level label annotation. To improve the recognition accuracy, we iteratively train the segmentation model by adding 128 virtual synthetic images at a time. In the meantime, 33 images in group B are recognized using the network trained each time, and the intersection over union (IoU)[27] between the recognition results and the previously labeled results is calculated to estimate the recognition accuracy.

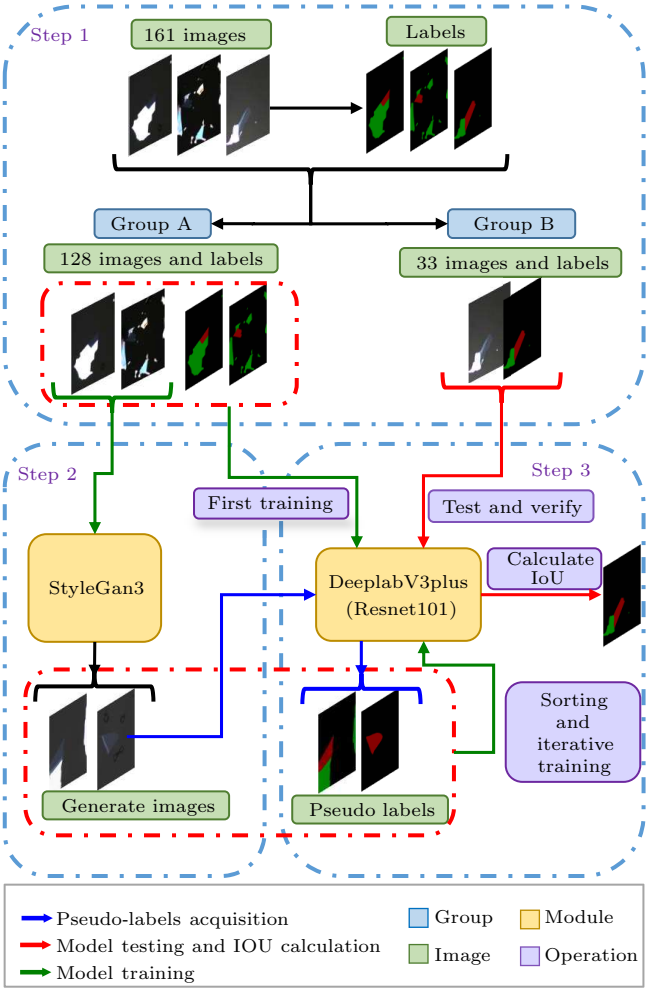


Fig. 1. Methodology and procedure. Three steps are included in the whole work, which are depicted by the enclosed blue dashed frames. Step 1: a total of 161 microscopic images of the mechanically cleaved WSe_2 flakes are collected and manually annotated, which are divided into groups A and B. Step 2: the 128 images in group A are randomly selected as input into StyleGAN3 and used to produce synthetic images. Step 3: the 128 original images together with their corresponding labels in group A are used to train a preliminary segmentation network of DeepLabv3Plus. Subsequently, the segmentation network is employed to predict the synthetic images and obtain pseudo-labels. The pseudo-labels and synthetic images are then filtered out and used to supplement the dataset for retraining the segmentation network. Finally, the 33 images in group B are recognized and the IoU is calculated to evaluate the recognition accuracy.

As the capacity of the dataset gradually increases from 128 to 640 images, the IoU increases from 88.59% to 90.38%. This increment can be observed from the segmentation results, which display sharp edges and the misclassification for contamination decreases. Our work demonstrates that training StyleGAN3 models can effectively generate visually realistic virtual images of two-dimensional materials and that using virtual images for data augmentation can improve recognition accuracy in segmentation. Furthermore, unambiguous boundaries and improved recognition accuracy will minimize

misclassification and allow for precise edge alignment during material stacking at scale, both of which are crucial for device performance.[28–31] The next sections will go through each stage in further detail.

2.2. Acquisition and annotation of datasets

In order to label the dataset initially used for training with more accurate labels, we use color space transformation and edge detection techniques. Figure 2 shows the process and results of acquisition, preprocessing and labeling of the dataset. We collect 161 microscopic images of WSe_2 thin-layer flakes in different lighting conditions to improve the generalization ability of the model. The magnification of the microscope is 50×. Figure 2(a) illustrates the process of manually annotating the 161 images assisted by traditional image processing, which improves the accuracy of the thin-layer edge labeling and reduces the manual cost of random shape labeling at pixel level. However, thin layers in the images are not easily identified by the naked eye for manually labeling the positions and edges (e.g., Fig. 2(c)). Therefore, we preprocess the images by using color space transformation and edge detection before annotation. Figures 2(b)–2(d) show the collected WSe_2 microscopic images. When the images are transformed from RGB into HSV space, the single- and few-layers become more obvious (Figs. 2(e)–2(g)). We denote 1–10 layers as thin layers or few layers, and more than 10 layers as thick layers. In order to label accurately, we grayscale the original images, perform edge detection using the Canny operator, and subsequently apply median filtering (Figs. 2(h)–2(j)).[32,33] Labeling with traditional image processing is efficient for the entire work, since it helps to recognize images more accurately.

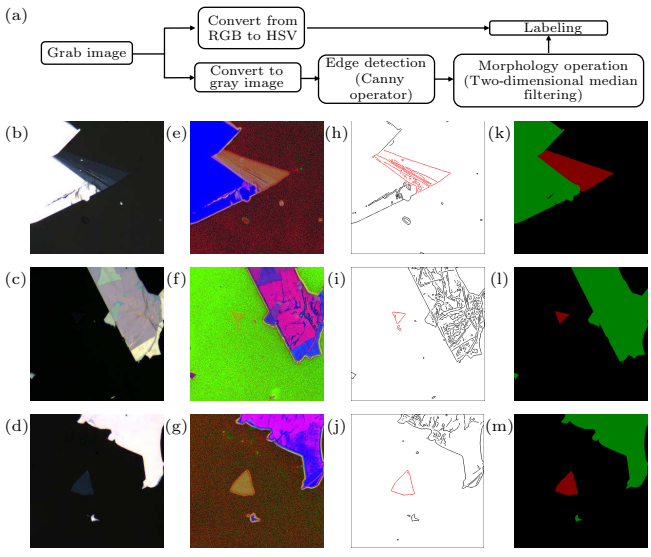


Fig. 2. Data acquisition and preprocessing. (a) Schematic diagram of the image preprocessing algorithm channel. (b)–(d) Original microscopic images of WSe_2 flakes. (e)–(g) Images after the RGB to HSV color space conversion. (h)–(j) Images after the grayscale processing, edge detection and median filtering using Canny operator. Few layers are outlined with red color manually. (k)–(m) Labels for few layers (red), bulk (green) and background (black).

3. Results and discussion

3.1. Generating synthetic images by StyleGAN3 training

Sufficient data is crucial for improving model generalization and reducing the risk of overfitting in machine learning. We choose the StyleGAN3 model to generate virtual images given that it is able to effectively control the features of the synthetic images and produce high-quality images. To evaluate the quality of synthetic images, Fréchet inception distance (FID) is adopted to quantitatively evaluate the similarity between the real and synthetic images

$$\mathrm{FID} = |\mu_A - \mu_B|^2 + \mathrm{Tr}\left(\Sigma_A + \Sigma_B - 2\left(\Sigma_A\Sigma_B\right)^{1/2}\right), \quad (1)$$

where μ_A (μ_B) represents the mean of the feature vectors extracted from the real (generated) images set using the inception network. Σ_A (Σ_B) represents the eigenvalue and eigenvector of the covariance matrix of the real (generated) images. Equation (2) indicates that the smaller the FID value, the more similar the distribution between the generated and real images.[34] We compute the FID every 40 iterations. Figure 3(a) illustrates that FID gradually decreases with the progression of training and stabilizes at 46.17 after 1720 iterations. Additionally, synthetic images at iterations 160, 400, and 840 are shown with FID scores of 140.25, 94.25, and 70.49, respectively. In these four positions, we can visually observe that the synthetic images become more similar to the real images as the FID decreases. Further examples of synthetic images that demonstrate the gradual approximation of the synthetic images to the real ones are presented in Fig. S3. This finding suggests that StyleGAN3 training can be used as an effective approach to generate high-quality images of 2D materials from limited original images. The FID score almost stops declining after 1720 iterations, implying a potential overfitting of the model. The FID is not particularly low due to the limited raw data. However, the synthetic images are sufficient to train the segmentation model. Even though not all synthetic images are realistic enough, we can obtain enough sample data by utilizing simply edge detection and a little manual sorting.

To demonstrate the advantages of StyleGAN3 in generating high-quality images, we compare it with traditional data augmentation. Figures 3(b)–3(d) show that when the input is rotated, the output of StyleGAN3 rotates as well, and the elements that do not appear in the figure are drawn when rotating. The resulting images resemble rotating the sample under a microscope. In contrast, traditional data augmentation can only reduce the size (Fig. 3(e)) or crop the image (Fig. 3(f)) in order to keep the image size when rotating it. This leads to information loss (see supporting information section 5 for details). In comparison, StyleGAN3 results can provide more information for training recognition models, and to some extent improve recognition accuracy. Subsequent experiments provide unambiguous evidence for this.

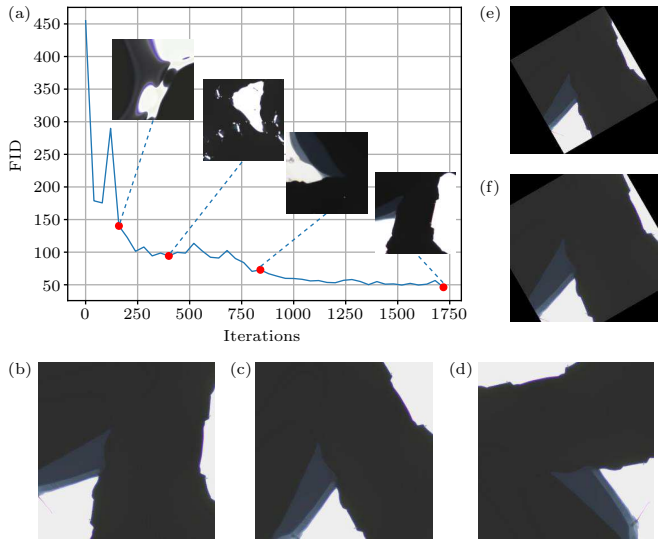


Fig. 3. Training process and results of StyleGAN3. (a) Evaluation of synthetic images on StyleGAN3. The red dots denote the FID with the iterations of 160, 400, 840, and 1720, indicating the effect of the current synthetic images. (b)–(d) Changes in the output images of StyleGAN3 when the input is rotated. (e) Traditional data augmentation of rotated graphs, using rotation followed by size reduction. (f) Traditional data augmentation of rotated graphs, rotation followed by cropping.

3.2. Recognzing images by DeepLabv3Plus training

In addition to images, the training of segmentation network also requires corresponding labels. Due to the huge amount of synthetic image data, semi-supervised approach will take an important role in reducing the labor cost in labeling production. Therefore, we first train the segmentation network to recognize the images to generate pseudo-labels. Then, the synthetic images and pseudo-labels with poor quality are filtered out by sorting, and the remaining high-quality images and pseudo-labels are added to the dataset. The network is trained with images in group A, and then its recognition accuracy is evaluated with images in group B. The recognition accuracy is estimated by the IoU, which is expressed by

$$\mathrm{IoU} = \frac{|A \cap B|}{|A \cup B|}, \tag{2}$$

where A and B represent the annotated target region and the predicted target region of the segmentation model, respectively. It is seen in Fig. 4(a) that the recognition accuracy of the background and thick layers is always higher than that of the thin layers. We extract the IoU for each category independently and use the IoU of the thin layers as the model assessment indicator rather than using the average IoU for all categories because thin layers are widely used in optoelectronic devices. The best recognition accuracy of thin layers trained by 128 images in group A reaches only 88.59% (Fig. 4(a)). The generated images and pseudo-labels are shown in Fig. 4(b). The left column shows the synthetic images generated by a model with a FID of 46.17, while the middle column marked with pseudo-labels presents the predicted segmentation of the synthetic images using the initial training of the segmentation model. Given that the synthetic images are random and the accuracy of the segmentation model is not high, we perform edge detection processing on the synthetic images and illustrate the results in the right column. It can be observed that the three columns of images coincide well, and thus synthetic images and pseudo-labels will be added to the dataset for further training. Since the input is a set of noise, we can theoretically generate an infinite amount of images, dramatically reducing the cost of real images acquisition. Together with the edge detection and sorting for labeling, which is much more efficient than the manual pixel-level labeling, this method can easily exclude images with poor generation or segmentation results.

In order to boost its performance, we repeatedly train the DeepLabv3Plus network by expanding the dataset in 128 increments. Since the Adam optimizer is employed, the training process is somewhat random. As a result, we train the dataset three times starting from scratch, using the same training parameters, and use the optimal IoU values from each time. As shown in Fig. 4(c), the optimal IoU reaches 90.38% by dataset expansion, which is around 1.79% higher than the model trained only with the real data. When the dataset capacity is increased from 256 to 640, the IoU has increased, of which recognition results are presented in Fig. 4(d). A good overlap between the segmented mask and the real images of 2D flakes can be seen from the correctly colored thin layers as red and thick layers as green. The detection process effectively removes pollutants like bubbles and tape residues, classifying them as background with minimal misjudgment (red arrows in Fig. 4(d)). As the dataset is gradually expanded by the adversarial synthetic images, the recognition of complex layered structures becomes more precise, and the segmentation of thin layer boundaries becomes clearer, while the frequency of misjudgments gradually decreases (white box in Fig. 4(d)). This finding demonstrates the effectiveness of the adversarial network for the dataset expansion and recognition accuracy improvement with limited real data. The improvement in recognition is mainly on the edges of the sample of few layers, which have a low percentage of pixels and thus are not remarkable in numerical terms. However, the recognition results in Fig. 4(d), previously misclassified, are assigned to the correct classification with this improvement. This improvement of recognition accuracy in the detail of few layers is more valuable than that in thick layers. In the automatic production of two-dimensional material devices on large scale, this reduction of misclassification will be significant in stacking high quality devices and reducing production waste.

It is demonstrated by Fig. 4(c) that our method consistently improves recognition accuracy during the initial stages of data augmentation. The training results under the same con-

ditions using the traditional method of data augmentation are given in Fig. S7. In contrast, the traditional method initially shows some improvement but lacks stability and eventually drops below the previous level as the dataset expands. We attribute this observation to the fact that traditional method loses image information during the augmentation process, while our approach generates new images by mimicking the existing distribution, providing more information and effectively enhancing the model performance. It should be noted that even with image generation-based augmentation, segmentation accuracy may decrease after 640 images. Synthetic images can reflect most of the information in real images, but not all of it. Therefore, as the proportion of synthetic images becomes too high, there is a decrease in the recognition accuracy of real images. Consequently, it is not possible to keep improving accuracy using this method indefinitely. Nevertheless, this approach allows for a stable increase in recognition accuracy during the early stages without incurring physical costs associated with material preparation and collection. It is highly practical for researchers aiming to develop machine learning algorithms tailored to their experimental environments and to improve device fabrication efficiency. We demonstrate that the trained network weights can be used directly for the recognition of other two-dimensional materials, and the recognition results can be viewed in the supplementary materials (Fig. S9). We suggest using a few pending data to fine-tune the model before applying the pre-trained weights, since this might result in higher recognition results.

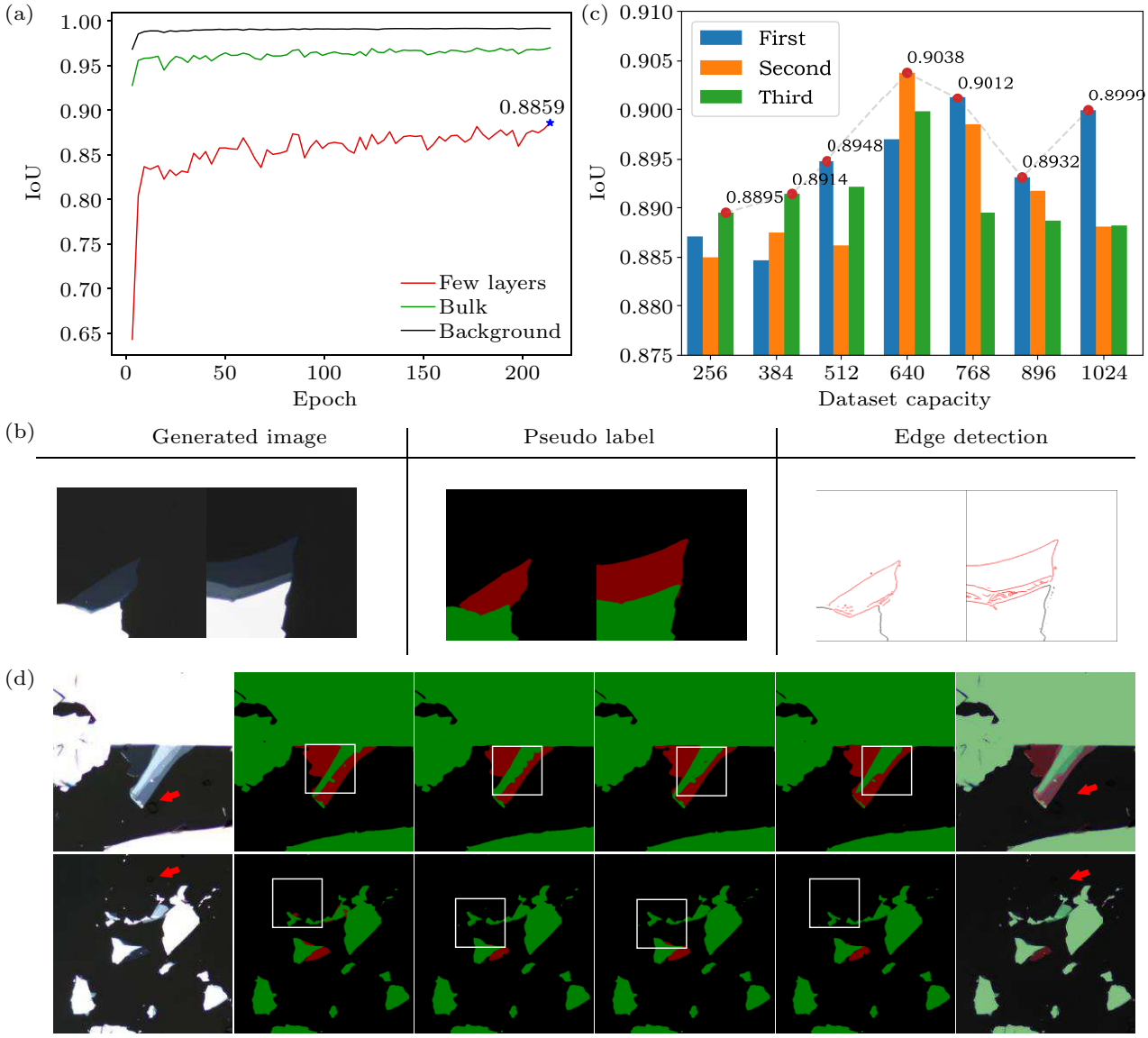


Fig. 4. DeepLavV3Plus training results. (a) IoU changes with training epoch in the first training session for different thickness of two-dimensional materials and backgroud. (b) Generated images, pseudo-labels, and edge detection results. (c) IoU depends on the expanded dataset. The dashed line shows the optimal changing trend after three rounds of training (blue, orange and green histograms) with the same data volume. (d) From left to right: 1st column, images to be recognized in group B. 2nd–4th columns, recognition results with different accuracies for GAN augmented dataset sizes of 256, 384, 512, and 640, respectively. 5th column, masking of 640 training model recognition results on the original images. Bubbles and contaminants indicated by the red arrows are classified as background in the predicted outcomes. White boxes present improvements in recognition details.

4. Conclusion

In conclusion, we demonstrate the feasibility of StyleGAN3 in generating synthetic images of two-dimensional materials and expanding dataset. We confirm that employing synthetic images for data augmentation aids in the recognition of two-dimensional materials and improves the recognition accuracy of segmentation network DeepLabv3Plus. The proposed data augmentation approach that we demonstrated is applicable to a wide range of two-dimensional materials. Our feasible and reliable method, prompted by the demand of scalable production of atomically thin materials, could helpfully explore the intriguing properties of layered materials and enable the rapid manufacturing of layered information devices.

Acknowledgments

Project supported by the National Key Research and Development Program of China (Grant No. 2022YFB2803900), the National Natural Science Foundation of China (Grant Nos. 61974075 and 61704121), the Natural Science Foundation of Tianjin Municipality (Grant Nos. 22JCZDJC00460 and 19JCQNJC00700), Tianjin Municipal Education Commission (Grant No. 2019KJ028), and Fundamental Research Funds for the Central Universities (Grant No. 22JCZDJC00460). C.Y.J. acknowledges the Key Laboratory of Photoelectronic Thin Film Devices and Technology of Tianjin and the Engineering Research Center of Thin Film Optoelectronics Technology, Ministry of Education of China.

References

[1] Tan Q, Rasmita A, Zhang Z, Cai H, Cai X, Dai X, Watanabe K, Taniguchi T, MacDonald A H and Gao W 2023 *Nat. Mater.* **22** 605
[2] Sauer M O, Taghizadeh A, Petralanda U, Ovesen M, Thygesen K S, Olsen T, Cornean H and Pedersen T G 2023 *NPJ Comput. Mater.* **9** 35
[3] Bertoldo F, Ali S, Manti S and Thygesen K S 2022 *NPJ Comput. Mater.* **8** 56
[4] Zhou X, Zhang R W, Zhang Z, Feng W, Mokrousov Y and Yao Y 2021 *NPJ Comput. Mater.* **7** 160
[5] Li M Y, Pu J, Huang J K, Miyauchi Y, Matsuda K, Takenobu T and Li L J 2018 *Adv. Funct. Mater.* **28** 1706860
[6] Dong J, Zhang L and Ding F 2019 *Adv. Mater.* **31** e1801583
[7] Zhao B, Shen D, Zhang Z, Lu P, Hossain M, Li J, Li B and Duan X 2021 *Adv. Funct. Mater.* **31** 2105132
[8] Zhang X W, D W He, He J Q, Zhao S Q, Hao S C, Wang Y S and Yi L X 2017 *Chin. Phys. B* **26** 97202
[9] Qi X, Gao M, Ding C, Zhang W, Qu R, Guo Y, Gao H and Zhang Z 2021 *Phys. Status Solidi (RRL)* **15** 2100052
[10] Hu B, Zhang T, Wang K, Wang L, Zhang Y, Gao S, Ye X, Zhou Q, Jiang S, Li X, Shi F and Chen C 2023 *Small* **19** e2207538
[11] Hu X, Qiu C and Liu D 2020 *Nano Res.* **14** 840
[12] Masubuchi S, Morimoto M, Morikawa S, Onodera M, Asakawa Y, Watanabe K, Taniguchi T and Machida T 2018 *Nat. Commun.* **9** 1413
[13] Masubuchi S and Machida T 2019 *NPJ 2D Mater. Appl.* **3** 4
[14] Dong X, Yetisen A K, Tian H, Gueler I, Stier A V, Li Z, Koehler M H, Dong J, Jakobi M, Finley J J and Koch A W 2020 *ACS Photonics* **7** 1216
[15] Li H, Wu J, Huang X, Lu G, Yang J, Lu X, Zhang Q and Zhang H 2013 *ACS Nano* **7** 10344
[16] Nolen C M, Denina G, Teweldebrhan D, Bhanu B and Balandin A A 2011 *ACS Nano* **5** 914
[17] Han B, Lin Y, Yang Y, Mao N, Li W, Wang H, Yasuda K, Wang X, Fatemi V, Zhou L, Wang J I, Ma Q, Cao Y, Rodan-Legrain D, Bie Y Q, Navarro-Moratalla E, Klein D, MacNeill D, Wu S, Kitadai H, Ling X, Jarillo-Herrero P, Kong J, Yin J and Palacios T 2020 *Adv. Mater.* **32** e2000953
[18] Saito Y, Shin K, Terayama K, Desai S, Onga M, Nakagawa Y, Itahashi Y M, Iwasa Y, Yamada M and Tsuda K 2019 *NPJ Comput. Mater.* **5** 124
[19] Masubuchi S, Watanabe E, Seo Y, Okazaki S, Sasagawa T, Watanabe K, Taniguchi T and Machida T 2020 *NPJ 2D Mater. Appl.* **4** 3
[20] Talaei Khoei T, Ould Slimane H and Kaabouch N 2023 *Neural. Comput. Appl.* **35** 23103
[21] Whang S E, Roh Y, Song H and Lee J G 2023 *VLDB J* **32** 791
[22] Yang R, Mei L, Zhang Q, Fan Y, Shin H S, Voiry D and Zeng Z 2022 *Nat. Protoc.* **17** 358
[23] Miao Q, Lin H, Wang X and Hassan M M 2021 *Comput. Netw.* **197** 108327
[24] Weiss K R and Khoshgoftaar T M 2016 *Proceedings of the 15th IEEE International Conference on Machine Learning and Applications*, December 18–20, 2016, Anaheim, CA, USA, pp. 207–213
[25] Wang K Z, Zhang D Y, Li Y, Zhang R M and Lin L 2017 *IEEE Trans. Circuits Syst. Video Technol.* **27** 2591
[26] Karras T, Aittala M, Laine S, Harkonen E, Hellsten J, Lehtinen J and Aila T 2021 *Proceedings of the Conference on Neural Information Processing Systems*, December 06–14, 2021, Electr Network, pp. 852–863
[27] Rezatofighi H, Tsoi N, Gwak J, Sadeghian A, Reid I, Savarese S and Soc I C 2019 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, June 16–20, 2019, Long Beach, CA, pp. 658–666
[28] Jin C, Tao Z, Li T, Xu Y, Tang Y, Zhu J, Liu S, Watanabe K, Taniguchi T, Hone J C, Fu L, Shan J and Mak K F 2021 *Nat. Mater.* **20** 940
[29] Yang Y, Li J, Yin J, Xu S, Mullan C, Taniguchi T, Watanabe K, Geim A K, Novoselov K S and Mishchenko A 2020 *Sci. Adv.* **6** eabd3655
[30] Liu Y, Dini K, Tan Q, Liew T, Novoselov K S and Gao W 2020 *Sci. Adv.* **6** eaba1830
[31] Liu F, Wu W J, Bai Y S, Chae S H, Li Q Y, Wang J, Hone J and Zhu X Y 2020 *Science* **367** 903
[32] Tang R, Chen W, Wu Y, Xiong H and Yan B 2023 *Sensors* **23** 3834
[33] Nhat-Duc H, Quoc-Lam N and Van-Duc T 2018 *Autom. Constr.* **94** 203
[34] Borji A 2022 *Comput. Vision Image Understanding* **215** 103329