1. **Introduction**

Scanning tunneling microscope(STM)，which has the ability of direct atomic-precise surface structure determination of objects, was introduced by J. Tersoff in 1984[1]. As a mighty imaging tool, STM has been widely applied in surface science[2, 3], material science[4], chemistry[5], biology[6] and nanotechnology[7],etc. Whereas on account of the STM system needs very high stability to generate high quality image, just slight outside vibration, motor periodic vibration or control system hysteresis may cause defects in topographies. Hence, image post-processing of STM morphology is a very significant issue for STM application.

Many image process methods have been introduced in STM image restoration and [denoising](javascript:;), such as Wavelet Denoising[8], Sparse Coding[9], Wiener filter[10], TV and criminisi algorithms[11] and so on. While all these algorithms were rely on manual classification and manual positioning of defects. Moreover, the complexity and time consumption can be added along with the defects amount [increasement](javascript:;) and position randomness augment. Therefore, in this study, a novel method for STM image post-processing is provided. It combines with defects classification network which using CNN and unique image processing algorithm. Using this method, STM images can be fast restored automatically by defects self-positioning and self-processing and the entire process minimize the manual intervention. Three typical types of STM images’ common defects to be processed are long stripe, short sparse stripe and periodic noise, as presented in Fig.1.



Long stripe and short sparse stripe are both local noise. Long stripe is usually caused by the pinpoint tinny shake when the whole precise mechanical system having outside disturbance. On account of the pinpoint scanning direction is horizontal, almost all long stripe defects present as horizontal linear stripes occupying the whole row or at least quarter. Short stripe is usually generated by system hysteresis before and after the pinpoint scanning to a raised surface topography which abundantly and sparsely appears in an image. The randomness and quantity of these two defects make identification and positioning a very time-consuming and inefficient work（label waste？）.

Periodic noise usually appears globally in an image because of periodic mechanical vibration causing by motor or refrigerant fan rotation. Although several researchers have introduced some methods of restoring strict periodic noise in the time domain[12-14] , methods in the frequency domain are more efficient and generally used for generic periodic noise[15-17]. While on account of the diversity of the periodic noise, positioning the frequency content in the frequency map becomes a difficulty and complexity task.



Materials and methods

(构思)

**材料**：处理的图像材料来源，大小，样本生成时间等等，主要针对的瑕疵对象

**图像处理策略**：

0、总写一下流程---流程图

1. 分类方法，cnn原理，网络构架，训练，三种瑕疵网络的差异改动
2. 长条纹特征分析与处理
3. 随机短条纹特征分析与处理
4. 周期性噪声特征分析预处理

（行文）

2.1 processing materials

The processed STM images were generated by (仪器名称) from 2007 to 2015 which total number is around ten thousand. The image size is 256pixels\*256pixels which has RGB three channels. It is analyzed and found that long stripe, short sparse stripe, periodic noise is the top three main defects of scanning images in spite of the various morphologies these images present.

2.2 image process strategy

As observed in fig1, each defect has its unique feature which has nothing to do with the morphology presented. The valid restoration methods become diverse because of the specificity of each defect which makes distinguishing defect as separate necessary. Furthermore, on account of the independence between defect features and the content presented, classifying defects by computer without manual works are possible. Therefore, the central idea of the developed algorithm consists of two parts:(1) classifying defects and marking all defect types that one image to be processed may have. (2) self-positioning and self-processing of each marked defect by different means according to appointed priority. The flowchart of the developed algorithm is shown in fig.2, and the various detailed steps involved in the algorithm are discussed in subsequent chapters.



2.3 defects classification and marking

A STM image may contain more than one type defects which makes defects recognizing a multi-label learning problem. The key challenge of multi-label learning problems is the large size of output space. As the types of label increases, the size of label sets grows exponentially[18]. Thus, a solving method called cross-training which transform the defects classification problem into three independent binary classification problems is to be used[19]. Because of the irrelevance between the different defect labels on each image, the coexistence of other labels can be ignored when training one label[20]. Then, how to set up and train the unique classification network for the three defects become the core problem.

On account of the difficulty in features extraction of traditional supervised machine-based learning methods, such as k-NN, SVM, these methods need some extraction techniques (e.g., SIFT, HOG, GIST)[21]. Hence, the convolutional neural network (CNN) which can extract features automatically and does not need hand designed features is be chosen[22]. The base structure of the proposed CNN network contains 5 convolution layers and 2 fully connected layers as fig.3 presented. The detailed parameters are different according to the diverse features the three defects have as fig.4 presented.



|  |  |  |  |
| --- | --- | --- | --- |
| **Layer name** | **Periodic noise** | **Long stripe** | **Short stripe** |
| **input** | 3@256\*256 | 3@256\*256 | 3@256\*256 |
| **Conv1** | 64@256\*256+(RELU) | 64@256\*256+(RELU) | 16@256\*256+(RELU) |
| **Conv2** | 64@256\*256+(RELU) | 64@256\*256+(RELU) | 32@256\*256+(RELU) |
| **maxPool1** | 64@128\*128+(RELU) | 64@128\*128+(RELU) | 64@128\*128+(RELU) |
| **Conv3** | 128@128\*128+(RELU) | 128@128\*128+(RELU) | 128@128\*128+(RELU) |
| **maxPool2** | 128@64\*64+(RELU) | 128@64\*64+(RELU) | 128@64\*64+(RELU) |
| **Conv4** | 128@64\*64+(RELU) | 128@64\*64+(RELU) | 128@64\*64+(RELU) |
| **maxPool3** | 128@32\*32+(RELU) | 128@32\*32+(RELU) | 128@32\*32+(RELU) |
| **Conv5** | 64@32\*32+(RELU) | 64@32\*32+(RELU) | 64@32\*32+(RELU) |
| **maxPoo4** | 64@16\*16+(RELU) | 64@16\*16+(RELU) | 64@16\*16+(RELU) |
| **Full1** | 1@1\*1024+(RELU) | 1@1\*256+(RELU) | 1@1\*128+(RELU) |
| **Full2** | 1@1\*32+(RELU) | 1@1\*32+(RELU) | 1@1\*32+(RELU) |
| **output** | 1@1\*1+(Sigmoid) | 1@1\*1+(Sigmoid) | 1@1\*1+(Sigmoid) |
| **Parameters** | *17,145,025* | *4,583,617* | *4,443,617* |

2.3.1 periodic noise classification network

CNN is a deep neural network which structure contains input layer, hidden layers, output layer. The hidden layers mainly include convolution layers, pooling layers and full connected layers. The size of pending STM image is 256\*256\*3. So the input layer shape is (256,256,3). The hidden layer of the network consists of 5 convolution layers and each layer flowed with a pooling layer. Although the periodic noise is global noise, the features of it are weak relative to the diverse background morphologies. So the kernel size of each five convolution layers are set to 3\*3 to ensure the effect that features of periodic noise can be learned by the network. All stride of each five convolution layers are set to 1 and filters number are 64,64,128,128,64 respectively. All pooling layers are use max-pooling method and 2\*2 kernel size. The nodes number of followed 2 fully connected layers are 1024 and 32. In the hidden layer section, each output will be nonlinearized by the Relu activation function. The output layer is a 1 node full-connected layer and activation function is sigmoid. Thus, the output 1 represent pending image has periodic noise and 0 represent it does not have. The specific parameters and output shape of each layer are presented in the second column of fig.4.

2.3.2 long stripe classification network

Long stripe is local noise, the evidence of its features are weaker than periodic noise. If using the same parameters of periodic noise classification network, the overfitting phenomenon, that the accuracy of training set rocket to nearly 100% while the accuracy of validation set rises slowly or even declines, may appear in training. Thus, in order to avoid the network being trained to remember all the samples, the size of the first full connected layer are reduced to 256 as shown in the third column of fig.4.

2.3.3 short sparse stripe classification network

Short sparse stripe is also local noise, and the sparsity and randomness are higher than long stripe. Furthermore, on account of the tinny size of each stripe, the feature of short sparse stripe is the weakest among three defects. In order to make the weeny feature can be retained during convolution, the stride of first two convolution layers are increased and the filters number change to 16 and 32 as demonstrated in fig.4.

2.4 defects restoration methods

What defects a pending image contained has been marked by the classification network in section 2.3. Thus, the following work is restoring the image in what way and in what priorities. Because of the restoring methods of three defects have different slight side effects which may affect other defects restoration of pending images, the order of each process is important to ensure the remediation effect. The priority level of three defects is defined as eq.1:

(1)

The reason of the rank will be mentioned in the following sections along with details of each restoration algorithm.

2.4.1 long stripe restoration

Long stripe defect is a kind of image damage which main causes are pinpoint tinny shaking. Some studies have proposed some effective algorithms of image damages restoration for different types of damages.

Total Variation(TV) algorithm was proposed by L.Rudin & S.Osher in 1992,and was amended and used for image inpainting by Chan in 2002. The principle of TV is that restore the pending region in pixels through the information of the boundary pixels from the outside of marked region. The mathematical formula of TV model is eq.2:

(2)

Among the expression, E is represented the marked pending region and D is represented the other region and u is represented the pixel value in the pending region. If u0 is defined as the original pixel and u is defined as the restored pixel, the expression of u is eq.3:

(3)

Because of TV using the outside boundary information, this restoring method are effective only when the damage size is relatively small. Otherwise the restoration result will become blur in center section of marked pending region as fig.5 shown.

Criminisi algorithm is a global searching algorithm used widely in texture repairing. The principle of criminisi is that chose one pixel (denoted by p) which has the highest confidence weight (eq.4) among the boundary of pending area, and chose the area of specified size (9\*9 in this article) around p as a template. Then find the best-matched (SSD standard, eq.5) area in unbroken region to replace the template area, and meanwhile update the confidence weight of p. By parity of reasoning, repeat the preceding procedure and the pending area will be restored after the end of the iteration.

But because of the globality of the searching and matching, when the pending area is smaller than the template, the restoring result may have some pseudo texture, as fig 5.



To solve the shortcoming of the two algorithm and make the best of the two method, [11]proposed a method which combing the two in restoring STM image damages.

While this method need manual positioning and judging of the defects. Hence, a self-positioning, self-judging and self-restoring algorithm is proposed.

Through the analysis of the long stripe defects, a phenomenon is found that the texture difference between damaged rows and adjacent lines are much larger than that between undamaged rows. Thus, we proposed a novel index parameter called ‘row gradient’ (denoted by RG) which present the sharpness of transition between two rows. The formula of RG is eq6:

(6)

Where x is the row of the image; n is the width of the image; p is the image matrix. As the fig.6 shown, the RG value of the damaged rows are prominently higher than others, and RG value the rows above and under the damaged line are about half of damaged rows’.



Then, the kernel work is finding the suitable threshold value of RG to divide the damaged rows and undamaged rows. In one pending image, the majority of the rows are undamaged, thus the average of the RG is nearly represent the datum RG value of undamaged rows. Also, the standard deviation of the RG can represent the extent of the RG augment of the damaged rows. Thus, the threshold value is defined as eq.7:

（7）

After remove the rows that adjoin the damaged row, the pending rows can be marked as fig.6.

Then, choose the restoring method automatically through whether the damaged rows are intensive, the judge formula is eq.8:

(8)

To sum up, after the algorithm mentioned above, the pending image marked having long stripe defects can be restored automatically without any manual intervention.

2.4.2 Short sparse stripe restoration

Short stripe defects can be viewed as global sparse outliers, although they are separately looked as local damages. Beside the sparse defects, the undamaged area of the pending image is identified as low-rank matrix. On this assumption, the restoring process is become a work that recover the low-rank section of the pending image, namely separating M= L0+ S0 into the terms L0 and S0[23]. Although the problem is N-P hard, a super-duper solution to those low-rank and sparse decomposition problem was introduced by Emmanuel Cand´es et al in 2010 called Robust Principal Component Analysis (RPCA). In this method, the low-rank L0 and sparse S0 which generated data can be exactly recovers by solving a simple convex program[24]:

(9)

where ,. This principle has been widely used in image processing and video processing[25], and inspire applied work in image video analysis[24]. On account of that the short sparse stripe defects can be viewed as the S matrix, restoring this kind of defects can be solved by RPCA as fig.7 present.



2.4.3 periodic noise restoration

Periodic noise is usually hard processed directly in time domain although it has obvious periodic feature. Because of the ability of Fourier transform that converting complex convolution operation in spatial domain into simple product operation in frequency domain, frequency domain filtering has been widely used in periodic de-noising of images. While，the shape and location of the filter determine the effect of de-noising , thus positioning the frequency domain that removed is a very important task. The Fourier transform of two-dimensional image is eq.10:

(10)

The frequency of periodic noise is two light vertical lines centrosymmetric distributing around the center of the spectrogram as the fig.8 presented. Also, analytically the distance between the light line and the center column is range from 1pixels to 6 pixels. The light lines’ position can be ensured by an index parameter called ‘column gradient’ (denoted by CG) which is defined like the way mentioned in 2.4.1, as the eq.11.

(11)

So, the column which has the highest CG value is the light line which contain the spectrogram of the periodic noise, as shown in fig.8.



After ensuring the light lines’ position, the next task is to estimate whether the light lines are distributing as ‘left high right low’ mode or ‘left low right high’ mode. Through analyzing the pixels value in the left light line in fig.9, when the distribution mode is ‘left high right low, the average of the 1~128 pixels is obvious higher than that of the 129-256 pixels, vice versa.

Then, after ensuring the distribution of the light line, choosing corresponding filter can restore the periodic noise of the pending image.

Results and discussion

1. 分类部分结果（data set preparation（数据获取，数据增强）、单个二分类结果（混淆矩阵，评价指标，roc）、总体分类示例）。
2. 修复结果（长条纹修复结果、短条纹修复结果、周期性噪声修复结果，均需要物质名称）。

The processed STM images were generated from 2007 to 2015 by the STM module of NanoScope®E produced by Digital Instruments department which total number is around ten thousand. The image size is 256pixels\*256pixels which has RGB three channels. It is analyzed and found that long stripe, short sparse stripe, periodic noise is the top three main defects of scanning images in spite of the various morphologies these images present.

3.1 Classification result

Through manual recognition and tagging, we collect three original data sets as table 1 shown. The positive samples and negative samples are roughly equally distributed. Distinguishingly, in the long stripe defect data set, we collect 126 positive samples and 126 negative samples which has totally same morphology except for long stripes in order to help the network study the characteristics of long stripes. Then, after shuffling the data set, respectively divide the three dataset into three set for training, validating and testing.

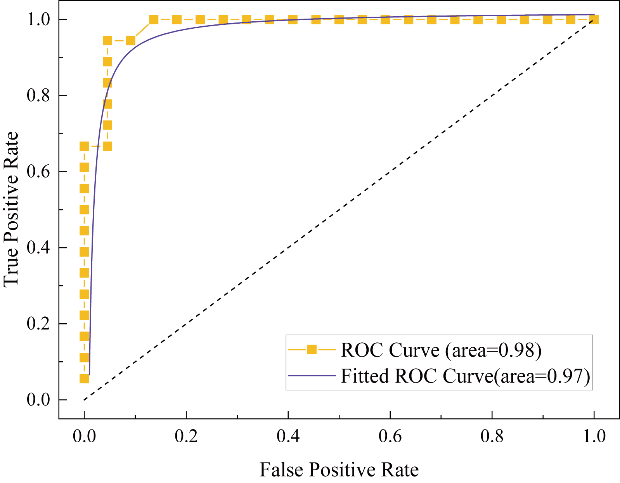
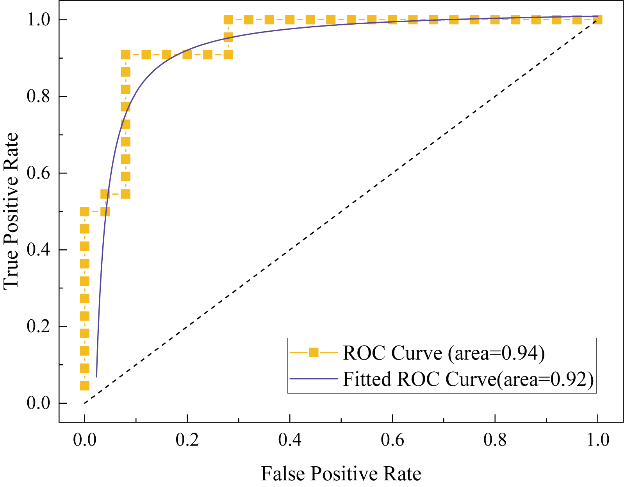
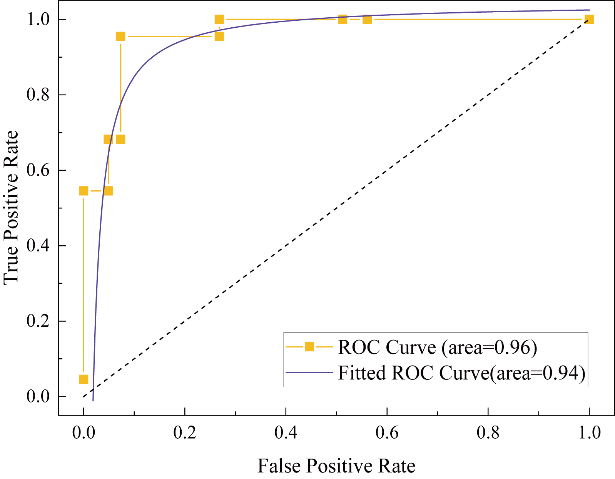
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Data Set Size** | **Original** | | **Divided** | | |
| YES | NO | Train | Valid | Test |
| **Periodic Noise** | 142 | 139 | 281 | 23 | 40 |
| **Long Stripe** | 126**\***+76 | 126**\***+60 | 300 | 44 | 44 |
| **Short Stripe** | 233 | 204 | 300 | 69 | 65 |

Then, in order to get good classification performance, we make some minor alterations such as flips, zooming and rotations to the existing training data sets as table 2 shown. Because of the periodic noise is global noise and the characteristics of it are in sensitive to direction, we randomly rotate the data in the range of and zoom the data in the range of 0.2% and flip the data both horizontally and vertically. While, to the short stripe and long stripe, in account of the locality and the sensibility of direction, we only use the horizontal flip and only zoom the short stripe data in the range of 0.1%. Through these minor changes, the training data amount of each set are increased to around 5000 which is more proportional to the parameters of the net.

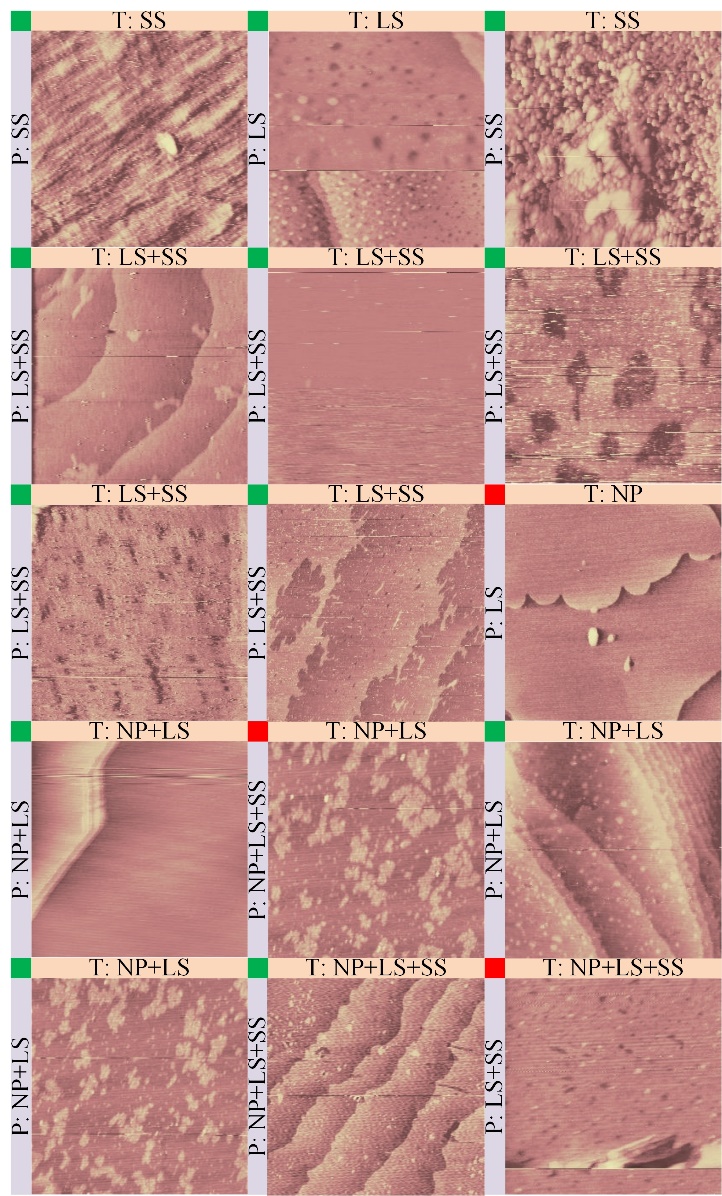
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Data Augment** | **Train Set** | **Augment Method** | | | **Augmented**  **Set** |
| Rotate | Zoom | Flip |
| **Periodic Noise** | 281 |  | 0.2% | H/V | 5708 |
| **Long Stripe** | 300 | ╳ | ╳ | H | 5216 |
| **Short Stripe** | 300 | ╳ | 0.1% | H | 5216 |

After the training data preparation, we use the augmented training data sets to train the correspondent binary network. And the accuracy of the validating data set indicates the degree of the training. We test each binary network in correspondent testing data set, and the resulting confusion matrix of each binary network is shown in table 3. The accuracy of the three binary network is respectively 92.500%, 91.489% and 92.063%. We also use the precision value, recall value and f-measure(a=1) to evaluate the binary network. Fig.10 present the three binary networks’ ROC (receiver operating characteristic) curve and the AUC (area under ROC curve). Putting all the evaluation indexes together, the three binary networks works well separately in each classifying work.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Periodic Noise** | **Predicted** | | | | **Evaluation Criterion** | | | |
|  | YES | NO | TOTAL | ACC | Precision | Recall | F-1 |
| **Actual** | YES | 20 | 2 | 22 | 92.500% | 95.238% | 90.909% | 93.023% |
| NO | 1 | 17 | 18 |
| TOTAL | 21 | 19 | 40 |
| **Long**  **Stripe** | **Predicted** | | | | **Evaluation Criterion** | | | |
|  | YES | NO | TOTAL | ACC | Precision | Recall | F-1 |
| **Actual** | YES | 23 | 2 | 25 | 91.489% | 92.000% | 92.000% | 92.000% |
| NO | 2 | 20 | 22 |
| TOTAL | 25 | 22 | 47 |
| **Short Stripe** | **Predicted** | | | | **Evaluation Criterion** | | | |
|  | YES | NO | TOTAL | ACC | Precision | Recall | F-1 |
| **Actual** | YES | 37 | 4 | 41 | 92.063% | 97.374% | 90.244% | 93.674% |
| NO | 1 | 21 | 22 |
| TOTAL | 38 | 25 | 63 |



In order to verify the universality of the classification net, we choose 15 image samples which are newly generated by the STM to validate the accuracy of the classification. As the fig.11 shown, the true label which were tagged manually is attached at the top of each image, and the predict label is attached on the left of the image. The color of the rectangle at the top left corner represent the correctness of the classification (▉green represent correct, ▉red represent incorrect). Among the 15 test samples, 12 images are classified correctly, 3 images have little deviation on the prediction, which are the 9th,11th and 15th respectively. Through analyzing, the periodic noise of the 9th image is very sharp and the single stripe of the noise is similar as a long stripe defect which may cause the mistake of the prediction. The short light line in the long stripe of the 11th image may cause that it was predicted not only having long stripe defects but also having short stripe defects. The periodic noise of the 15th image is slightly blurred which may lead to the mistake.



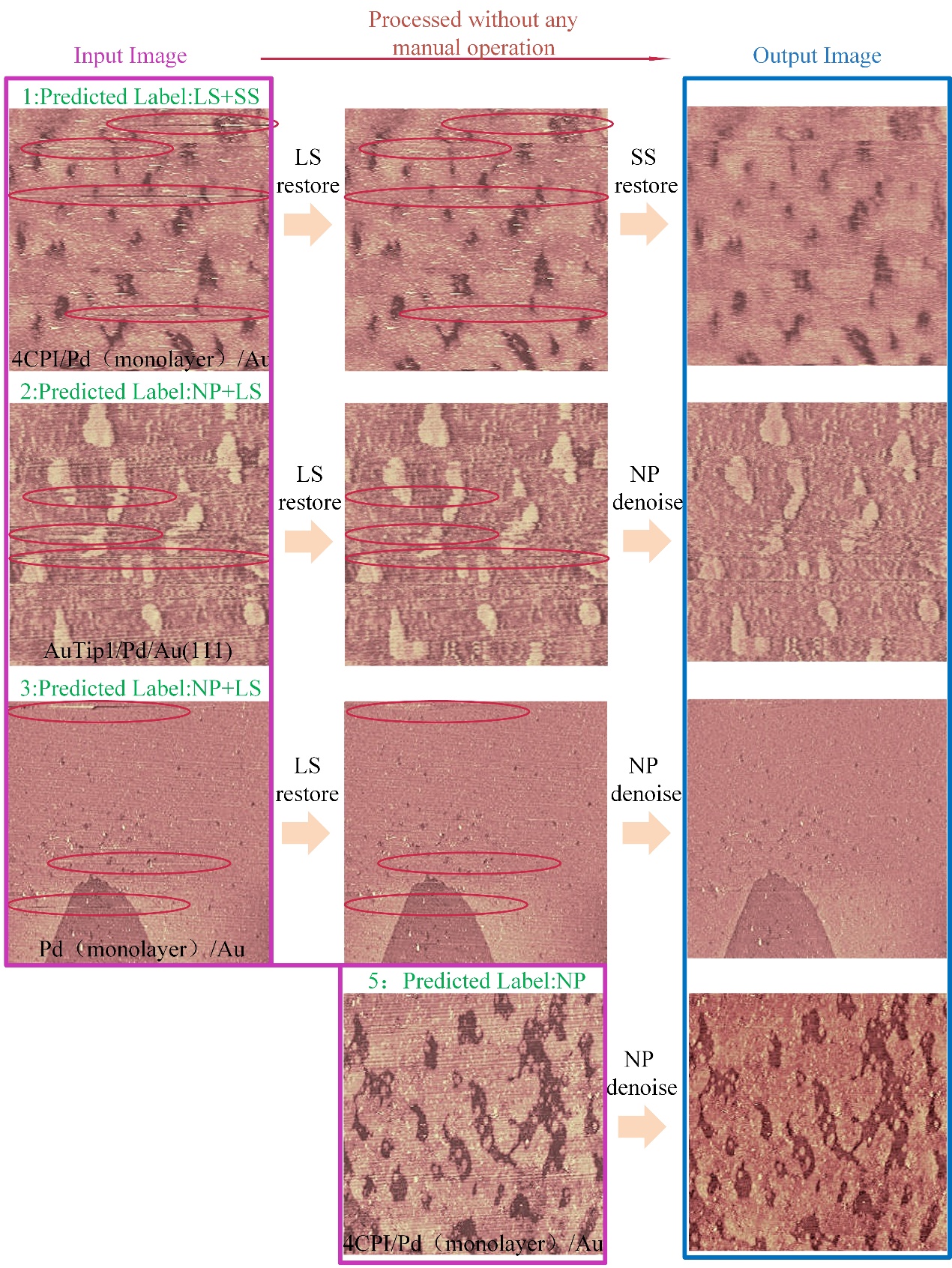
The overall results indicated that the deep convolutional neural network provided above have a quite well performance in the defect classification and prediction work.

3.2 Restoration result

Table 4 indicates the effectiveness of the provided restoration methods mentioned before. Three provided algorithm respectively process 50 images which have typical corresponding defects, and only a few of images are not restored good remaining some defects. The integral success rate of restoration is around 95%().

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Image Restoration** | **Restoration Result** | | | **Success**  **Rate** |
| total | good | bad |
| **Periodic Noise** | 50 | 48 | 2 | 96% |
| **Long Stripe** | 50 | 46 | 4 | 92% |
| **Short Stripe** | 50 | 49 | 1 | 98% |

fig.11 vividly presents the entire restoring process which without any manual intervention. The no.1 topography (CPI/Pd（monolayer）/Au) is correctly predicted by the network as LS and SS. It is restored automatically by the long stripe restoration algorithm (self-locating TV&CRISMINI) and short stripe restoration algorithm (RPCA) successively. No.2 (Autip1/Pd/Au(111))and No.3 (Pd(monolayer)/Au)are both predicted as NP and LS. After the process, the long stripes are restored well and the periodic noise are repressed obviously not entirely (remain little slight grains). No.4 (4CPI/Pd(monolayer)/Au are predicted as NP, and the periodic noise is removed well.



4 conclusion

We have proposed a novel full-automatic method for STM post-processing combing CNN classification network and three specific restoring algorithms. The predicted network use 3 deep CNN binary network which parallel distinguish whether pending images has corresponding defects. The long stripe restoration algorithm uses a self-positioning and self-deciding mechanism combined with TV and Criminisi to restore the long stripe automatically. The short stripe restoration algorithm uses a low-rank restoration algorithm, RPCA, which has relatively high robustness to sparse random outliers. The periodic noise denosing algorithm uses a self-positioning mechanism to locate the noise spectrum and uses a specific band-pass filtering to depress the noise. The processed results demonstrate the effectiveness of the method adequately. The proposed novel method has been written in the software and can process one image in seconds automatically without any manual intervention.

Further research should be focus on expanding the kinds of defects that can be handled and on how to change the process order adaptively to get the best process result. In the current method, the order of process is determined by prior experience and can’t be changed in different conditions.

1. Tersoff, J. and D.R. Hamann, *Theory of the scanning tunneling microscope.* Physical Review B, 1985. **31**(2): p. 805-813.

2. Motin, A.M., et al., *Surface science approach to Pt/carbon model catalysts: XPS, STM and microreactor studies.* Applied Surface Science, 2018. **440**: p. 680-687.

3. Nguyen, H.A., et al., *STM Imaging of Localized Surface Plasmons on Individual Gold Nanoislands.* J Phys Chem Lett, 2018. **9**(8): p. 1970-1976.

4. Kumar, A., et al., *STM study of exfoliated few layer black phosphorus annealed in ultrahigh vacuum.* 2D Materials, 2018. **6**(1).

5. Zhao, A., et al., *STM tip-assisted single molecule chemistry.* Phys Chem Chem Phys, 2013. **15**(30): p. 12428-41.

6. Tempe, D., et al., *Multisite protein kinase A and glycogen synthase kinase 3beta phosphorylation leads to Gli3 ubiquitination by SCFbetaTrCP.* Mol Cell Biol, 2006. **26**(11): p. 4316-26.

7. Tapaszto, L., et al., *Tailoring the atomic structure of graphene nanoribbons by scanning tunnelling microscope lithography.* Nat Nanotechnol, 2008. **3**(7): p. 397-401.

8. Schimmack, M. and P.J.J.o.t.F.I. Mercorelli, *An On-line Orthogonal Wavelet Denoising Algorithm for High-Resolution Surface Scans.* 2017: p. S0016003217302776.

9. Oliveira, J.P., et al., *Restoring STM images via Sparse Coding: noise and artifact removal.* 2016.

10. Pancorbo, M., et al., *A WIENER FILTER WITH CIRCULAR-APERTURE-LIKE POINT SPREAD FUNCTION TO RESTORE SCANNING TUNNELING MICROSCOPY (STM) IMAGES.* Pattern Recognition Letters, 1990. **11**(8): p. 553-556.

11. Xian-guang, F., et al., *Image processing for three defects of topography images by SPM.* Chemometrics and Intelligent Laboratory Systems, 2019. **185**: p. 12-17.

12. *<选择迭加法消除周期性噪声\_穆信良.pdf>.*

13. *<基于柔性形态学滤波优化的周期性噪声消除算法\_卫星.pdf>.*

14. *<基于权重自适应形态学的周期性噪声去除方法\_戴丹 (1).pdf>.*

15. *<2-D FFT for Periodic Noise Removal on Strain Image.pdf>.*

16. *<Periodic Noise Removal using Local Thresholding.pdf>.*

17. Moallem, P., M. Masoumzadeh, and M. Habibi, *A novel adaptive Gaussian restoration filter for reducing periodic noises in digital image.* Signal, Image and Video Processing, 2013. **9**(5): p. 1179-1191.

18. Zhang, M.-L. and Z.-H. Zhou, *A Review on Multi-Label Learning Algorithms.* IEEE Transactions on Knowledge and Data Engineering, 2014. **26**(8): p. 1819-1837.

19. Boutell, M.R., et al., *Learning multi-label scene classification.* Pattern Recognition, 2004. **37**(9): p. 1757-1771.

20. Zhang, M.-L. and Z.-H. Zhou, *ML-KNN: A lazy learning approach to multi-label learning.* Pattern Recognition, 2007. **40**(7): p. 2038-2048.

21. Zhang, M., W. Li, and Q. Du, *Diverse Region-Based CNN for Hyperspectral Image Classification.* IEEE Trans Image Process, 2018. **27**(6): p. 2623-2634.

22. Liu, X.L., et al., *On fusing the latent deep CNN feature for image classification.* World Wide Web-Internet and Web Information Systems, 2019. **22**(2): p. 423-436.

23. Candès, E.J., et al., *Robust principal component analysis?* Journal of the ACM, 2011. **58**(3): p. 1-37.

24. Candes, E.J., et al., *Robust Principal Component Analysis?* 2011. **58**(3): p. 1-37.

25. Bouwmans, T., et al., *On the Applications of Robust PCA in Image and Video Processing.* 2018. **106**(8): p. 1427-1457.

2.4 defects