Dear R.F. Egerton and Reviewers:

Thank you for your letter and for the reviewers’ comments concerning our manuscript entitled ‘Fully-automatic defects classification and restoration for STM imageshttps://www.evise.com/evise/adf/images/t.gif’(ID:JMIC\_2019\_232). Those comments are all valuable and very helpful for revising and improving our paper, as well as the important guiding significance to us researches. We have studied comments carefully and have made correction which we hope meet with approval. Revised portion are marked in red in the paper. The main corrections in the paper and the point-by point reply to reviewers’ comments are attached below.

Yours sincerely,

Xin Wang

Reply to Reviewer 1

We thank the reviewer for his/her constructive criticisms that have helped us to improve our manuscript. The point-by-point response to the comments is given below. (Reviewer’s comments in blue)

· Is there a real demand for such a type of procedure? Contemporary STM instruments working in the proper environment and operated by suitably trained and skilled personal do practically not any more produce such type of artefacts. Especially the global noise level is usually very small. STM images are published nowadays mostly as raw data (may be, sometimes gently polished by a low-pass filtering). Appearing of “local noise”, on the other hand, has often real physical reasons (e.g. a specific feature on the surface as a cluster, a molecule or something else might be for some reason less tightly bound to the substrate than others or possessing a less stable internal configuration or interacting in a different way with the probing tip). Such real features in the raw data are sometimes very important findings and should never be “automatically” removed.

Reply:

It is really true as reviewer suggested that contemporary STM instruments is indeed a very precise instrument in design, and nanoscale accuracy is usually achieved when it is working in low temperature or noise reduction environment. However, we are not able to expect that the microscope always works perfectly as designed. In practice, many reasons may lead to defects in STM images:

. the inevitable external interference may penetrate the control system (the periodic noise due to CCD refrigeration of Raman spectrum acquisition in TERS applications).

. STM instruments is open loop when they work in constant height mode, thus, external weak vibrations (artificial or environmental) may cause the temporary jitter or the scanning probe which generate long stripe defects.

. When STM instruments work in constant current mode, the strong irregularity of the samples’ morphology (a sudden bulge or sag) may lead the STM close loop system adjust the z axis displacement to maintain the tunneling current rapidly. While because of the hysteresis of piezoelectric ceramics, short stripe defects may appear before and after the raised or hollow morphologies.

Defects caused by the above reasons usually have distinctions from the defects caused by real physical reasons.

. Defects caused by the above reasons have relationship with the systems’ adjustment time for the external interference and the hysteresis time of the piezoelectric ceramics. Thus, the size of this kind of defects will not be change along with the adjustment of the scan range and scan accuracy as the fig.1 show. However, the noise caused by a specific feature on the surface as a cluster, a molecule or something else may be scale up with the reduction of the scanning area.

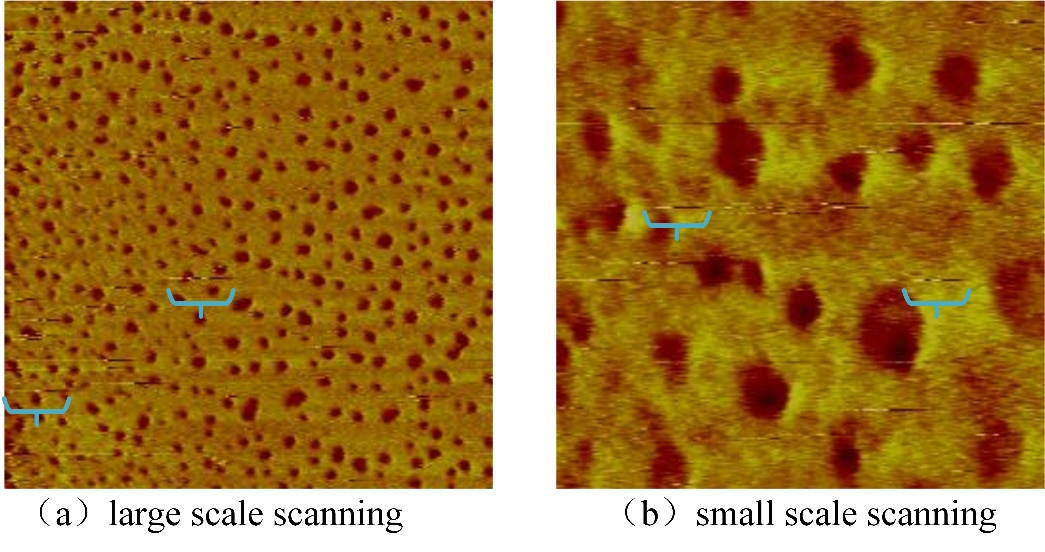


Fig. 1 different scanning scales for the same samples (defects have generally consistent size)

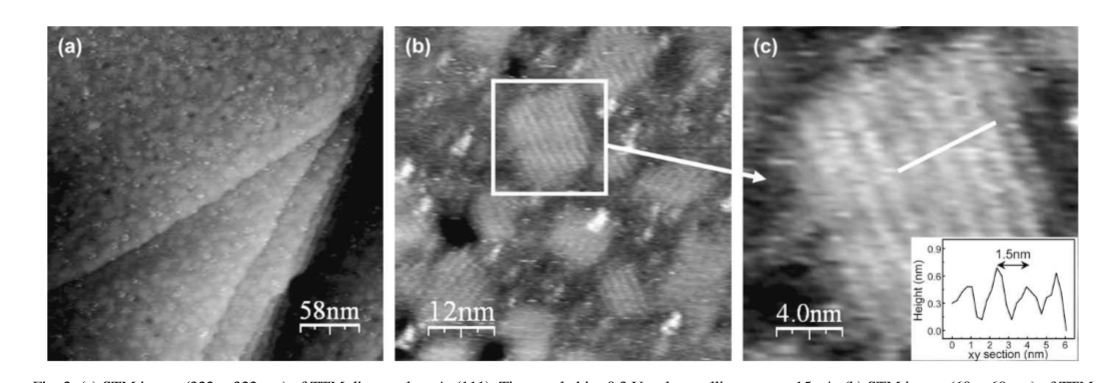
 . The mentioned long stripe defects and short stripe defects in the article have the same characteristics that they all extend along the scanning direction and the margin of defects and morphology is very sharp. These are all indicate that the defects are produced by the mistakes in probes during the scanning procedure. While, the noise produced by specific cluster or a molecule usually don’t have such sharp margin and they usually have specific morphologies related with their physical characteristics instead of all stretching along the scanning direction. For example, (Mugnaini et al., 2009) detect one single spin cattier molecule deposited on surface by using Electron Spin Noise Scanning Tunneling Microscopy(ESN-STM),as fig. 2.

Fig. 2 TTM (tris (2,4,6-thrichlorophenyl) methyl radical) dispersed on Au(111)

Mugnaini, V., Fabrizioli, M., Ratera, I., Mannini, M., Caneschi, A., Gatteschi, D., Manassen, Y., Veciana, J., 2009. Towards the detection of single polychlorotriphenylmethyl radical derivatives by means of Electron Spin Noise STM. Solid State Sciences 11, 956-960.

Action:

1. Add a state mentioning simply the distinction with the two local defects and the noises generated by real physical reasons at section 1 paragraph 2:

‘Because of that these two noises have relationship with the ‘mistakes’ during scanning procedure, they have their own characteristics distinguished from real physical morphologies which make the automatic recognition and restoration possible.’

· The authors should at least mention the very high importance and necessity of a separate storage of the real raw data as the original data. These are always the only basic data to be considered and discussed. Especially critical seems to be the effect which is inherent in the use of neuronal networks that there is no way to retrace what actually did happen to the original data after the particular processing. “Automation” of scientific work in such a way is in contradiction to good scientific practice and might endanger the user to be blamed for scientific misconduct.It might be, that possible applications of the method could be found in routine imaging e.g. in industrial processes as quality control but I cannot imagine and would even not tolerate the application in basic research. In any case, the authors should give a very clear statement on these items.

Reply:

As reviewer suggested that the storage of the original data is indeed not only important but also necessary. In effect, the original intention of the restoration is not to cover up the original data but hints of the probable morphologies for the researchers when the scanning images have defects.

The procedure of get a STM morphologies image is relatively complex and time consuming. Some samples have complex preparation process and need re-prepare before each experiment. The preparation of scanning probes through electrochemical and the scanning procedure are very time consuming. And also, some special phenomenon doesn’t always have the second chance to observe. The defects caused by these scanning mistakes usually cover up the true morphologies and physic phenomenon of the observed samples which are supposed to be found. Thus, the orientation of restoration of the original scanning images is not to change the original data but to give research useful hints to the real morphology with less time cost. The quick restorations make the researchers can judge if the samples studied contain the phenomenon they interested in or judge the correctness of their prepared samples quickly. The restored images can guide the researchers decide whether it is necessary to do more detailed research on the samples (such as scanning more times to get the correct morphologies in the quitter and purer environment).

In practice, taking the State Key Laboratory of Physical Chemistry of Solid Surfaces, Xiamen University which we cooperate with as example, the researchers usually need many times of sample preparation and need dozen times of scanning on each sample to ensure their interested scanning area. Among these experiments, many repeats are on account of that the real morphologies are covered up by the noises caused by the improper operations and the surrounding noises. Thus, the restoration of STM images mentioned in our article are very useful as a kind of research aid of indicating the real morphologies which can improve the research efficiency.

As to the non-interpretability of deep learning the reviewer mentioned. Indeed, it is the shortcoming of the artificial intelligence(AI). And also, it is a hot research direction nowadays. Some types of restoration which fully rely on neural networks may indeed make the restored image unexplainable and even have serious distortions. However, the deep convolutional neural network our article used is designed to help people determine which kinds of defects exists in a STM image. The input of the network is an original image, and after the ‘parsing’ of the input image, the output of the network is labels which indicate what defects the image have. Thus, at the classification stage, the deep convolutional neural network (DCNN) doesn’t make non-interpretable change to the original image. After the DCNN get the defects label, our article used particular algorithm to position and restore each kinds of defects. These algorithms are combined by traditional image processing which restore the possible morphologies of the image before the damage. These restorations are all interpretable.

About the storage of the original image, it is indeed very important and necessary for the basic research. As mentioned above, our restoration is aimed at giving hints to the researchers instead of covering up the real phenomenon. Thus at the image post-process software, the original image will always be shown. Researchers can choose whether use the automatic restoration, and the restoring result will be shown in an individual form. And also, only the original image can be exported.

Action:

1. Add a paragraph in the section 1 to mention the real intention of the introduced introduction which avoid readers from misunderstanding.

‘Because of the difficulties of samples and probes preparation and the long period of scanning, getting a STM image is time consuming. And also, researchers usually need scanning many times to get an entire perfect image without any defects covering up the real morphologies. The original intention of this article is to introduce a restoration method for giving researchers quick hints of the real morphologies when the STM images have defects.’

1. Add ‘the processed results are shown in an individual form as hints for the researchers (the original data is always stored).’to the section 4 paragraph 1 to declare the automatic process is not to cover up original data but to give hints.

· STM has not been introduced by J.Tersoff (who is a theoretician) but by G.Binnig and H.Rohrer (both of them were awarded by the Nobel prize in physics 1986 for the invention of the STM). Suitable references are also missing in the following (this refers to [1-7] which are all not the groundbreaking papers as which they are announced here).

Reply:

We are very sorry for our incorrect writing in the mentioning of the introducer of STM. We have corrected this point in the revised manuscript. As to the references [2-7], we originally intended to list some latest applications of STM in different fields, so we didn’t refer some references in the earlier times. Now we maintain some latest reference and add some earlier previous references.

Action:

1. Correct the statement of the introducer of STM in section 1 paragraph 1:

‘Scanning tunneling microscope(STM),which has the ability of direct atomic-precise surface structure determination of objects, was introduced by G.Binnig and H.Rohrer in 1983(Binnig and Rohrer, 1983). Both of were awarded by the Nobel Prize in physics 1986 for the invention(Binnig and Rohrer, 1987).’

1. Alter the references [2-7] in the previous manuscript:

Add (Binnig et al., 1983, Voelker et al., 1988, Rosei et al., 2003)

Delete (Tempe et al., 2006, Motin et al., 2018)

· The description of the defect classification networks (2.2.1. … 2.2.3.) is not very helpful (even obscure) for an outsider. The referee understands the inherent problem with this for any type of neuronal network. Nevertheless, the authors should try to explain better why these different versions are just well adapted to the particular defects.Reply:

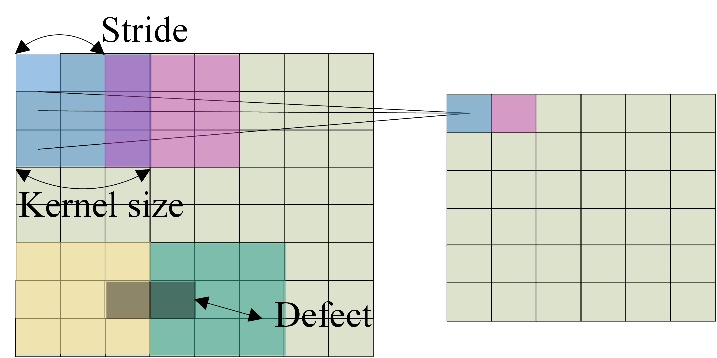
We are very grateful for the reminder that the expression in 2.2.1 ... 2.2.3 is a little obscure for an outsider. Thus, in the revised manuscript, we alter some expression and add some figures to try our best to explain. However，the deep learning is actually a practical discipline to some extents and the set of the hyper-parameters are adjusted according to the training result. Sometimes we spent a lot of time in adjusting the hyper-parameters to make the training result better. Thus, it’s hard to explain why the particular parameter is better. But some recognized experience (which have been added in the revised manuscript) can be used to determine the direction of the adjustment of parameters.

Action:

1. Add some explanation of the reason of select the particular kernel size for the periodic noise classification in the section 2.2.1.

‘Thus, it’s necessary to select appropriate size of the convolution kernel. Convolution kernel is also known as receptive field. The pixels in the kernel will be weighted sum as a new pixel in the next layer, as the Fig. 4show. When the defects’ size is small, if the kernel size is set too large, the characteristics of defects may be neglected during training. And also, when the stride is set large (even equals to the kernel size), if the defects are in middle of the kernel margin, the defects’ characteristics may loss.’

1. Add a figure in section 2.2.3 to show the meanings of the stride and the kernel size and explain why the stride can’t be set large.



1. Add an explanation about what is the overfitting of the network in section 2.2.2.

‘(the training set’s accuracy is 100% but the testing set’s accuracy is very low)’

1. Add an explanation about why the kernel size should be expanded in the section 2.2.3.

‘On account of that the short stripe is sparse, if the stride is small, the sum of each kernel’s pixels may become similar and the sparse defects may loss its’ features during the training.’

· The rule (1) looks strange. It is not satisfying to refer just to “prior experience”, especially if this kind of hierarchy seems to influence on the output.Reply:

We are very sorry about that we didn’t explain clearly about why the hierarchy is the best in the previous manuscript. In fact, we have done many experiments in each restoration algorithm of each particular defects and find that one algorithm may influence the other one. First, periodic noise restoration may make the margin of the long stripes become a little blur which may cause difficulties of their positioning. Second, short sparse stripes restoration uses a kind of low rank recovery algorithm which can eliminate the outliers of the image (not only the short sparse stripes but also the sparse components of other defects). Thus, the original character of the periodic noise and long stripes may be changed which make their restoration more difficult. Above all, we set that the periodic noise restoration should after the long stripe restoration and the short sparse stripe restoration should be in the last.

Action:

1. Add the explanation of the hierarchy in section 2.3 paragraph 1.

· What is meant by “positive and negative samples” in 3.1.?

Reply:

The positive label represents that the samples have the specified kind of defects. Similarly, the negative label represents that the samples don’t have the specified kind of defects. The classification networks of each kind of defects study large numbers of the positive and negative samples and then when a new image inputs the network, it may output the predicted label indicating whether the image have the defects.

Action:

1. Add simple explains after the ‘positive’ and ’negative’ in section 3.1

· Several symbols have to be explained (e.g. “Chl” in (A4,)).Reply:

The CPI represents P-chloro-isonitrile benzene molecule, Autip1 represents using the Au tip 1 to scanning.

Action:

1. Add explains to the notes of fig.11

· It is unusual to have “formula” which contain whole words as well as symbols ((1), (3), (4)).Reply:

we have replaced the whole words in the formula (1) (3), the formula (4) which is to describe using which algorithm to restore the long stripe defects may not be simplified.

Action:

1. Change the eq.1 to ‘’, and add notes to explain ‘where LS represents long stripe, PN represents periodic noise, SS represents short sparse stripe.’
2. Change the ‘threshold’ to ‘thr’ in eq.3.

· The references and the spelling of author names have to be rechecked carefully (e.g. [11], [12]) as well as misspellings have to be removed (e.g. “resored image” in Fig.6). The term “et al.” should not be used in the reference list – each author and coauthor deserves to be announced by his name.

Reply:

We are very sorry for our negligence of the wrong formats and the misspellings. We have corrected these problems in the revised manuscript.

Action:

1. Correct all references in the manuscript.
2. Redraw the fig.6 to correct the misspelling.

· The English has to be improved considerably (e.g. plural forms are missing almost throughout the whole paper).

Reply:

We are very sorry for our negligence of the syntax errors in our previous manuscript. We have checked all the manuscript carefully and correct the errors we have found.

Action:

In the abstract section:

1. Revise the ‘judgement’ to ‘judgements’ in the line 3.
2. Revise the ‘predicted’ to ‘predicts’ in the line 7.
3. Revise the ‘network’ to ‘networks’, the ‘label’ to ‘labels’.
4. Revise ‘difficulty’ to ‘difficult’, ’complexity’ to ‘complex’.
5. Revise ‘noise’ to ‘noises’, ’image’ to ‘images’.

In the section 1:

1. Revise the ‘theory’ to ‘theories’.in paragraph 6 line 2.

In the section 2:

1. Revise the ‘defect’ to ‘defects’ in section 2.1 paragraph 1 line 4.
2. Revise the ‘difficulty’ to ‘difficulties’ in section 2.2 paragraph 2 line 1.
3. Revise the ‘pseudo texture’ to ‘pseudo textures’ in the section 2.3.1 fig.4.
4. Revise the ‘algorithm’ to ‘algorithms’ in the section2.3.1 paragraph 4 line 1.
5. Revise the ‘difference’ to ‘differences’ in the section 2.3.1 paragraph 5 line2.
6. Revise the ‘value’ to ‘values’ in the explanation of the eq.2.
7. Revise the ‘kernel work’ to ‘core work’ in section 2.3.1 paragraph 5.

In the section 3

1. Delete ‘three’ in the section 3.1 paragraph 1 line 7.
2. Revise ‘samples’ to ‘samples’’ in the section 3.1 the title of table 1.
3. Revise ‘data’ to ‘data’s’ in the section 3.1 paragraph 3 line 1.
4. Revise ‘network’ to ‘networks’, ‘the’ to ‘one’ in the section 3.1 paragraph 3 line 2.
5. Add ‘correspondingly’ after ‘training’ in the section 3.1 paragraph 3 line 3.
6. Revise the ‘network’ to ‘networks’, ‘present’ to ‘presents’ in the section 3.1 paragraph 3 line 7.
7. Revise the ‘curve’ to ‘curves’ in the section 3.1 paragraph 3 line 8.
8. Add ‘values’ after ‘AUC (area under ROC curve)’.
9. Revise ‘network’ to ‘networks’, add ‘is’ after ‘precision’ and ‘recall’ in section 3.1 table 3.
10. Revise ‘image’ to ‘images’, ’represent’ to ‘represents’, add ‘network’ after ‘classification’ in section 3.1 fig 10.
11. Revise ‘represent’ to ‘represents’ in section 3.1 table 4.

Reply to Reviewer 2

We thank the reviewer for his/her constructive criticisms that have helped us to improve our manuscript. The point-by-point response to the comments is given below. (Reviewer’s comments in blue)

The manuscript presents some interesting ideas on the validation of STM images. The English is not adequate for presentation in Micron and must be carefully revised before publication.

Reply:

We are very sorry for our negligence of the syntax errors in our previous manuscript. We have checked all the manuscript carefully and correct the errors we have found.

Action:

In the abstract section:

1. Revise the ‘judgement’ to ‘judgements’ in the line 3.
2. Revise the ‘predicted’ to ‘predicts’ in the line 7.
3. Revise the ‘network’ to ‘networks’, the ‘label’ to ‘labels’.
4. Revise ‘difficulty’ to ‘difficult’, ’complexity’ to ‘complex’.
5. Revise ‘noise’ to ‘noises’, ’image’ to ‘images’.

In the section 1:

1. Revise the ‘theory’ to ‘theories’.in paragraph 6 line 2.

In the section 2:

1. Revise the ‘defect’ to ‘defects’ in section 2.1 paragraph 1 line 4.
2. Revise the ‘difficulty’ to ‘difficulties’ in section 2.2 paragraph 2 line 1.
3. Revise the ‘pseudo texture’ to ‘pseudo textures’ in the section 2.3.1 fig.4.
4. Revise the ‘algorithm’ to ‘algorithms’ in the section2.3.1 paragraph 4 line 1.
5. Revise the ‘difference’ to ‘differences’ in the section 2.3.1 paragraph 5 line2.
6. Revise the ‘value’ to ‘values’ in the explanation of the eq.2.
7. Revise the ‘kernel work’ to ‘core work’ in section 2.3.1 paragraph 5.

In the section 3

1. Delete ‘three’ in the section 3.1 paragraph 1 line 7.
2. Revise ‘samples’ to ‘samples’’ in the section 3.1 the title of table 1.
3. Revise ‘data’ to ‘data’s’ in the section 3.1 paragraph 3 line 1.
4. Revise ‘network’ to ‘networks’, ‘the’ to ‘one’ in the section 3.1 paragraph 3 line 2.
5. Add ‘correspondingly’ after ‘training’ in the section 3.1 paragraph 3 line 3.
6. Revise the ‘network’ to ‘networks’, ‘present’ to ‘presents’ in the section 3.1 paragraph 3 line 7.
7. Revise the ‘curve’ to ‘curves’ in the section 3.1 paragraph 3 line 8.
8. Add ‘values’ after ‘AUC (area under ROC curve)’.
9. Revise ‘network’ to ‘networks’, add ‘is’ after ‘precision’ and ‘recall’ in section 3.1 table 3.
10. Revise ‘image’ to ‘images’, ’represent’ to ‘represents’, add ‘network’ after ‘classification’ in section 3.1 fig 10.
11. Revise ‘represent’ to ‘represents’ in section 3.1 table 4.