AI 4 Science Discovery Network+

Group: 5
Challenge: Nanopore "Defect Detection in Graphene Sheets"
AI4SD ML Summer School Report
20-24th June 2022

Project Team: Anna Bachs Herrera (Swansea University), Abdoulatif Cisse (The University of Liverpool), Emilio Alexis de la Cruz Nuñez Andrade (Swansea University), Philipp Deussen (UCL), Ivan Yankov (University of Strathclyde)

Report Date: 08/07/2022

AI4SD-SummerSchool-Series:Report-3

Group: 5
Challenge: Nanopore "Defect Detection in Graphene Sheets"
AI4SD-SummerSchool-Series:Report-3
Report Date: 08/07/2022
DOI: 10.5258/SOTON/AI3SD0246
Published by University of Southampton

Network: Artificial Intelligence and Augmented Intelligence for Automated Investigations for Scientific Discovery
This Network+ is EPSRC Funded under Grant No: EP/S000356/1

Principal Investigator: *Professor Jeremy Frey*
Co-Investigator: *Professor Mahesan Niranjan*
Network+ Coordinator: *Dr Samantha Kanza*

Contents

1 Project Details

Group Number	5
Challenge Name	Nanopore “Defect Detection in Graphene Sheets”
Project Dates	20-24th June 2022
Website	https://github.com/Ablatif6c/hackathon_group_5

2 Project Team

2.1 Project Student

Name and Title	Anna Bachs Herrera
Employer name / University Department Name	Swansea University
Work Email	2037458@swansea.ac.uk

Name and Title	Abdoulatif Cisse
Employer name / University Department Name	The University of Liverpool
Work Email	Abdoulatif.cisse@liverpool.ac.uk

Name and Title	Emilio Alexis de la Cruz Nuñez Andrade
Employer name / University Department Name	Swansea University
Work Email	2132253@swansea.ac.uk

Name and Title	Philipp Deussen
Employer name / University Department Name	UCL
Work Email	philipp.deussen.15@ucl.ac.uk

Name and Title	Ivan Yankov
Employer name / University Department Name	University of Strathclyde
Work Email	ivan.yankov.2016@uni.strath.ac.uk

2.2 Challenge Description

Graphene is a nanomaterial with excellent super conducting properties, is considered to be the strongest material available and is impermeable to gasses (even Helium -the smallest gas atom). Its properties find applications in fabrication of electronic and optoelectronic devices, gas sensors such as chemiresistors for detection of ammonium [1] and nitrogen dioxide [2], biosensors, composite materials and energy storage devices. Graphene has a regular structure - similar to a honey comb, which consists of fused six membered carbon rings into 2D array. This atomic structure is the foundation of the exceptional properties of the material. The occurrence of defects in the lattice which occur during the manufacturing process can cause significant deterioration in the property of the final material produced [3]. These defects can be detected with a number of spectroscopy-microscopic techniques, one of which is Surface Scanning Electron Microscopy (SEM).

In the challenge nanopore "Defect Detection in Graphene Sheets" a complete dataset of 180 SEM images 256x256 pixels (full-stack), from which were derived 2279 images 48x48 pixels of perfect patches (pp) and 32 images 48x48 pixels of defect containing patches (dp) were provided in numpy arrays, see Figure 1. All pixel values had been normalised to values between 0.00 and 1.00. Figure 1 Nanopore "Defect Detection in Graphene Sheets". The task was to design a classifier for SEM graphene images using the pp dataset. The two classes are as follows: the first is images without defects and the second class, images with defects, where the type of defect or its localisation are not considered. Finally, the classifier had to be evaluated using an appropriate metric on the dp dataset.

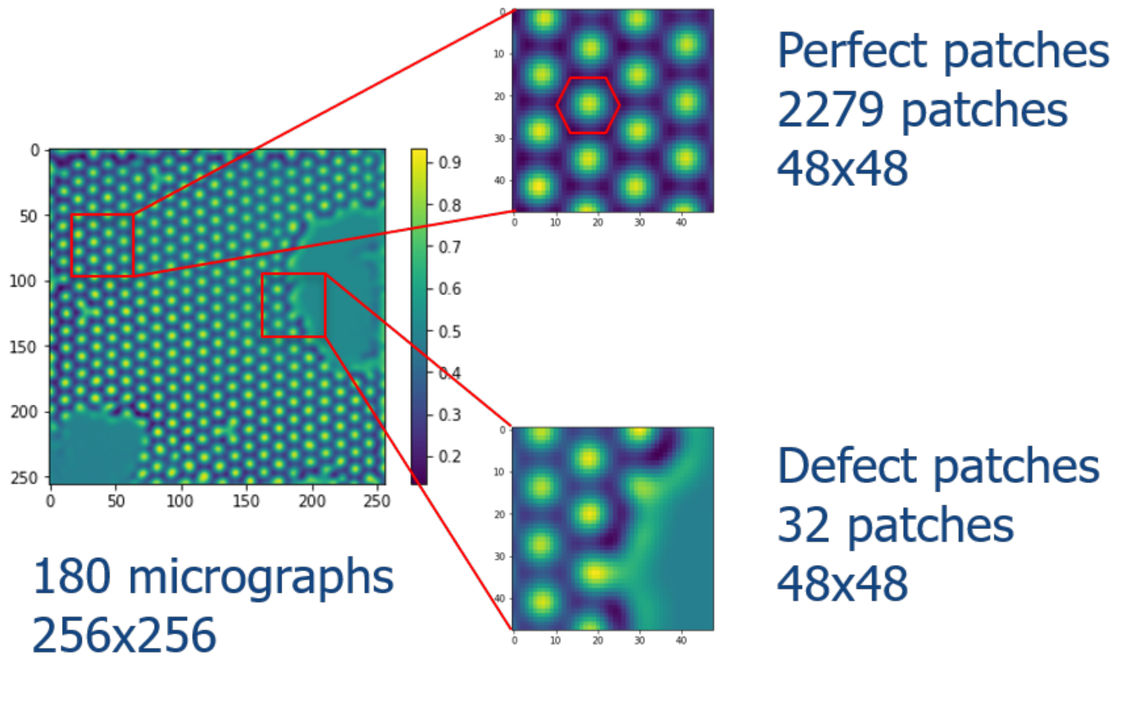

Figure 1: Nanopore "Defect Detection in Graphene Sheets".

3 Lay Summary

In the Nanopore "Defect Detection in Graphene Sheets" images from Scanning Electron Microscopy are used to inspect the quality of the hexagonal structure of graphene. To automate the process a model classifier was build using a Gaussian Process Classifier (GP) available in sklearn and a second deterministic approach comparing the eigenvalues in the latent space after

PCA to a benchmark was explored.

The GP performance was evaluated on a balanced augmented dataset of healthy and defect containing patches in a 10-fold cross validation. The performance was measured against the accuracy, neg log loss, ROC, confusion Matrix and classification report.

4 Methodology

Supervised learning approach sklearn Gaussian Process Classifier.

The dataset was faced with a class imbalance problem where pp were much more than the 32 dp. For this purpose, the dataset was augmented augmentation and 2000 samples from each augmented pp and dp were used to prepare a balanced dataset for training and evaluation of the model. Data augmentation included horizontal, vertical and both flip of images, as well as modification using blur, sin, cos, tan, exp which increased the 32 defected images to 4096 from which 2000 and another 2000 from pp were used and shuffled to form a work_dataset. The created work_dataset, constituting of both perfect and defect patches, was split into 60%:20%:20%, training, validation and testing sets.

Collective image features such as such as row sum and column sum for each individual row and column, the standard deviation (SD) for the whole image and contrast values were computed from the images. The nD arrays were flattened and the SD was added at the end of the feature vector. To select the most appropriate kernel for the following list of available kernels were tested on the validation set - RBF, DotProduct, Matern, RationalQuadratic and WhiteKernel. The selection criteria was accuracy of prediction achieved in over three 10-cross-validation loops on the validation set. Once the best kernel was selected a model was trained on the training set and evaluated against the testing set. This was done k-fold times = 10 in order to get confidence in the performance metrics. The evaluation matrices were: accuracy, NegLogLoss, ROC, Confusion Matrix and the classification report.

The workflow was executed using sklearn version (1.0.2) using The GaussianProcessClassifier from gaussian_process library; sklearn.model_selection and sklearn.metrics libraries.

The deterministic approaches included, setting a threshold value for the calculated SD over the image, contrast value or eigenvalues in the latent space after PCA analysis.

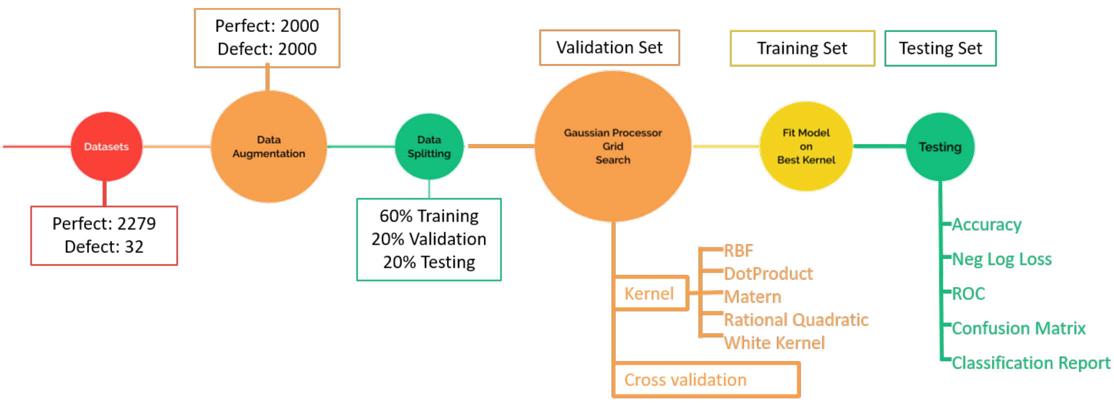

Figure 2: General Workflow for Nanopores "Defect Detection in Graphene Sheets".

5 Results

A representative example from the perfect and defect patch datasets were taken to investigate their properties. In particular, the goal was to identify a collective varible or a pattern/property that can be used to distinguish between the two classes. The sum of pixels intensities across each column for a healthy patch resulted in a nice sinusoidal patern ranging between 20 to 30, where the lowerst was across the zig-zag carbon pattern and the highest was because of 4 regular high intensity background circles which are the nanowholes in the ring. It was interesting to compare the pattern to a defect patch. It is clear to compare that defects would compress the intensity of the sinusoidal pattern where carbon attoms are missing and overall large areas with average background intensity would cause the sinosiudal pattern to levell off around the mean. This gave foundation to either compute the SD for the vertical and horizontal diraction and use it at the end of the flattened vector of an image as additional features or rather than using all 48x48 features just to apply the 48 feature in the horizontal and vertical projections.At the time of the invetigation it was selected to use the SD in each projection as it seemd to give the best differentiation between the classes.

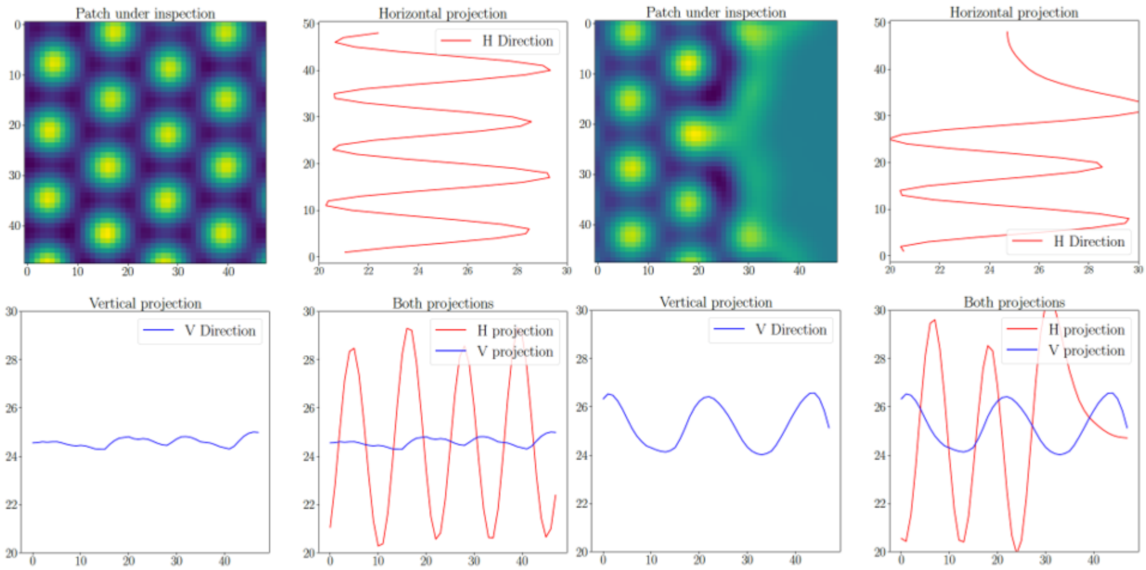

Figure 3: Image Analysis - horizontal and vertical projection of perfect and defect patch.

The augmentation process generated 2560 defect patches from the intitial 32 available. A limit of 2000 images from each class was applied when sampling from the perfect and augmented defect patches to generate a balanced shuffled dataset from which te test, validaiton and train sets were prepared. The best kernel - RBF, for the GP classifier was selected based on its highest mean accuracy in three 10 fold cross valdiaiton loops. In addition, Rational Quadratic kernel was not used in the final evaluation of the model, due to upon fitting a GP classifier to the training data, an error emphasising tha the alpha value was exceeding its recommendaed upper limit. In order to address the issue, the dataset can be pre-processed.

Best Mean Accuracy	Best Kernel Config
1.000	RBF(length_scale=1)
nan	DotProduct(sigma_0=1)
0.998	Matern(length_scale=1, nu=1.5)
1.000	RationalQuadratic(alpha=1, length_scale=1)
0.500	WhiteKernel(noise_level=1)

Table 1: Best Kernel configuration and mean accuracy

The fitted on the training data GP classifier with the RBF kernel achieved high results in the performance metrics, which gave reason to think the model was overfitting. Therefore, the model was tested on the full stack data. For this purpose, 11,520 patches from the full-size images were classified by the model. 81 48x48 patches were sampled from each of the 180 images, by taking a 48x48 patch every 26 pixels along the rows and along the columns.

'copy__train'	True,
'kernel__length_scale'	1.0,
'kernel__length_scale_bounds'	(1e-05, 100000.0),
'kernel'	RBF(length_scale=1),
'max_iter_predict'	100,
'multi_class'	'one_vs_rest',
'n_jobs'	None,
'n_restarts_optimizer'	0,
'optimizer'	'fmin_l_bfgs_b',
'random_state'	None,
'warm_start'	False

Table 2: GP Classifier RBF Kernel: Hyper parameter configuration

Metric	Mean ± SD
accuracy	0.999 ± 0.004
neg_log_loss	-0.187 ± 0.018
roc_auc	1.000 ± 0.000)

Table 3: Performance of GP CLassifier with RBF Kernel in 10-fold cross validation

400	0
0	400

Table 4: Confusion matrix of GP CLassifier with RBF Kernel

	precision	recall	f1-score	support
0.0	1.00	1.00	1.00	400
1.0	1.00	1.00	1.00	400
accuracy			1.00	800
macro avg	1.00	1.00	1.00	800
weighted avg	1.00	1.00	1.00	800

Table 5: Classification report of GP CLassifier with RBF Kernel

Each the classified images were saved in a two separate directories. By inspection over the defect containing images, it was promising to see that the model was able to identify plentiful images with defect and none with entire regular patterns. Nevertheless, when inspecting the categorised as perfect patches images, hexagonal arrays missing single atoms were the prevalent problem. That could be attributed to the fact that the dp dataset even after augmentation does not contain images with such examples. This issue can be address by labelling images containing single missing atoms and retraining the model on the enriched augmented dataset. Another approach would be to substitute the minimum intensity pixels (those representing carbon atoms) with maximum or high intensity pixels which would represent missing atoms in some of the healthy patches.

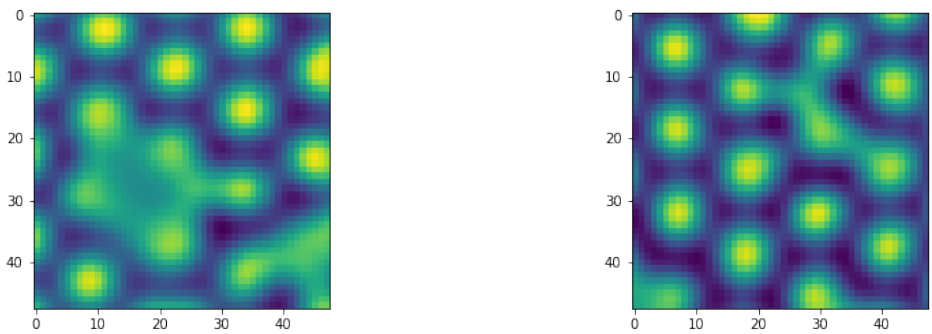

Figure 4: Single missing atoms in a graphene array

Another set of difficult to categorised images were, such the material goes outside the field of view of the SEM, thus giving a large average intensity area. This has not been defined at the start of the problem whether it should be classified as a healthy or defect patch. Such images it would be wise to be categorised into a third separate category which have undetectable regions of the material in order to remove unambiguity.

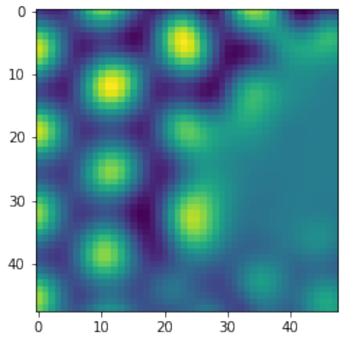

Figure 5: Graphe SEM image with average intensity region (out of focus)

6 Conclusions & Future Work

A GP classifier using RBF kernel was trained on augmented balanced dataset with two additional features which were SD in the horizontal and vertical directions. The goal of the model was to classify if an image contained a perfect/regular graphene grid or if the grid contained any defects, where the type of the defect was not important.

The performance of the model was monitored in a 10-fold cross validation quoting the accuracy, negative log loss, roc_auc and confusion matrix.

In addition, the model was applied to classify subarrays 48x48 from the 180 images using steps of 26 pixels in the horizontal and vertical direction. There were 11,520 subarrays which were used for inspection. The insight from this experiment was that the training data was poor in single missing atoms which were misclassified as regular arrays in the full stack. This could have also been the result of wrongly assigning labels in the healthy training set. Furthermore, regular patterns which fell out the field of view (being out of focus) were also categorised as regular arrays which seem plausible given the formulation of the problem.

To carry forward this project it is suggested to utilise contrast values of the images and to compress the features from 48x48 to 48x2 by utilising the sum across the horizontal and vertical direction. In addition, pre-processing the data prior to training the model could potentially improve its performance. It is strongly advised to expand the scope of the different defects in the training dataset in order to enable the model to detect single missing atoms. This can also be achieved by randomly introducing low intensity values instead of high intensity values (essentially swapping a carbon atom for transparent background, which would be approximately the equivalent of such a defect).

Lastly, threshold value approaches should further be explored as they could inform the feature design criteria.

7 Outputs, Data & Software Links

PowerPoint presentation: AI4SDMLSummerSchool_Group5 available:

- on GitHub at Ablatif6c/hackathon_group_5
- and Microsoft Teams AI4SD ML Summer School / Files / group Presentations

Jupyter Notebook Code available on GitHub at Ablatif6c/hackathon_group_5
https://github.com/Ablatif6c/hackathon_group_5

References

[1] Freddi S, Gonzalez MCR, Carro P, Sangaletti L, De Feyter S. Chemical Defect-Driven Response on Graphene-Based Chemiresistors for Sub-ppm Ammonia Detection. Angewandte Chemie International Edition. 2022 Apr;61(16). Available from: https://onlinelibrary.wiley.com/doi/10.1002/anie.202200115.

[2] Ricciardella F, Vollebregt S, Polichetti T, Miscuglio M, Alfano B, Miglietta ML, et al. Effects of graphene defects on gas sensing properties towards NO $_2$ detection. Nanoscale. 2017;9(18):6085-93. Available from: http://xlink.rsc.org/?DOI=C7NR01120B.

[3] Araujo PT, Terrones M, Dresselhaus MS. Defects and impurities in graphene-like materials. Materials Today. 2012 Mar;15(3):98-109. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1369702112700457.

List of Figures