UNIVERSITY OF BIRMINGHAM

DEPARTMENT OF PHYSICS AND ASTRONOMY

Using Machine Learning to Investigate Rotations in HAADF-STEM Images of Large Nano-particle Structures

Joe Laurie

Supervisor: Wolfgang Theis

Monday 11th December, 2023

Abstract

This preliminary report outlines the plan for a project focusing on the analysis of nanoparticles (NPs) through High Angle Annular Dark Field Scanning Transmission Electron Microscopy (HAADF-STEM). The key objective is to integrate Convolutional Neural Networks (CNNs) in NP orientation analysis, utilizing stereo imaging for unsupervised learning on an artificial dataset. This is done in Python using Tensorflow/Keras[14]. Two proposed CNN-based methods are tasked with identifying orientations of regular polygons and polyhedra without explicit angle information. Initial progress involves angle identification of polygons on a single axis of rotation using the first of the two methods. Future work includes refining networks and addressing nuanced rotational symmetry issues, allowing progression to projected images of polyhedra rotated about multiple axes. Success would further the utility of HAADF-STEM imaging for synthesis and investigation of NP structure, which can be often be found at random orientations in a solution.

1 Introduction

In contrast to traditional microscopes, STEM employs a focused electron beam instead of light to scan specimens. By measuring transmitted electrons, STEM achieves superior resolution at the atomic level, surpassing the limitations of optical microscopy[3] [17] [10]. This enables scientists to unveil intricate details of materials and biological structures, marking a significant advancement in nanoscale imaging and analysis across various scientific fields.
This project aims to use machine learning to aid the analysis of NPs imaged by HAADF-STEM. The main tools used are CNNs, which are a class of deep learning model designed for visual processing tasks[16]. They use convolutional layers to learn hierarchical representations from input data, making them highly effective in image recognition and object detection[2]. By building such networks, this project will be able to determine the orientations of large NP structures from their 2D projection. In order to be useful, the NNs (neural networks) used must be able to outperform lab personnel in assessing complicated structures. This will be made possible through the use of stereo imaging. By taking pairs of images at a fixed angular offset, the NNs will be better able to map the features of the projected shapes without knowledge of the shape or angle in training [13].
There are two proposed methods of analysis to be used and compared. The first is to pass both stereoimages through the same CNN, hence determining the angles of each with the use of the known angular separation. The second will use a CNN on the first image to determine it's angle, and generate an image of the stereo-partner using a second NN (comparing this to the original second image). In STEM imaging, the time required to produce large datasets is too great for efficient use of neural networks because each image must be individually set up. A solution can be found in artificial generation of such datasets for the training of image recognition algorithms. This will have to be done within the project. Ultimately this will result in an NN capable of determining angles of previously unseen shapes to a high degree of accuracy.

2 Aims

The project will have to be successful at creating a dataset of images that are reasonably similar to STEM images[8], such that the methods used could be easily adapted to use with real scientific data. These will be used in stereopairs to create a neural network for identifying the orientation of a given shape, without being supplied the ground truth angles for training. The most useful input data would be 2D projections of 3D shapes; as is produced by HAADF-STEM[8][7]. However, this is not feasible as an introductory task so the first step will be use of purely 2D images (this can be thought of as the simplest cross section of prisms in 3D). With this data, the

neural networks should be able to achieve accuracy of 95% or greater when attempting to determine the orientation angle of the shapes along the axis perpendicular to the stereo pair separation angle (which is also the plane of the polygon).
Once this is done, the project will move on to the use of projections. This will turn a "top-down" image of a 2D shape into an intensity graph along an axis through it. By using these projections, the trained networks will encounter similar problems to the use of 3D structures. If there is not significant difficulty encountered, progression to 3D shapes is desired as an end goal to the project. This allows for two further angles to be used simultaneously with the first, although they might not be introduced in the same stage. This will present significant difficulty as the angular symmetries are not confined to the rotational axes[9]. Composite rotation angles are able to map back to the original set of unique orientations in the angle phase space, which will be much more difficult to account for in the neural network design. In two dimensions this can be done easily by labelling all ranges beyond the first symmetry with $0°$ ($90°$ for a square) as identical to the first range ($0° - 90°$).

3 Approach

The training data is constructed from a collection of "randomly" generated start conditions. Shapes are created with a start position, size and orientation that makes the results non identical. This is done to prevent the network attempting to learn the dataset, overfitting[11]. These are turned into a low-pixel, greyscale image with intensity-scaled, noise. This is true to STEM output [7] as the NPs cannot be expected to always be found in the same position relative to the image "frame". These images are then passed through a pre-prossessing function as if they were experimental data. This will centre the image and remove the noise, in order to prepare it for the learning tasks. The networks will be trained without the ground truth of the image angle[13]. This is an important feature to the reliability as the networks cannot be expected to learn from any dataset that is already labelled with angles because they are being trained to eliminate the need for human labour.
Both structures will receive the same input, a pair of stereoimages. The first method will involve two identical CNNs and a FCN (fully connected network) which feed into a third, simpler network. The CNNs will be tasked with producing the angle of the image which they are given without access to the ground truth. The FCN will take a zero-dimensional input and be tasked with providing the output of the internal angle of the polygon. They then feed these three angles to the final network which consists of a single layer asking for both: the difference between the two orientations and the difference plus the internal angle. This will produce pairs of results, from which the minimum absolute value will be extracted. The expected pairs are [10,100] and [-80,10] for a square. The ground truth of the dataset will then be checked against

the output of the overall structure. If the training is successful, the networks will be able to determine, accurately, the angle offset between the images. This will enable the CNNs to be tuned to correctly identify the orientation. Using the trained CNN, the angle of any single image can be determined. This plan will inevitably lead to the networks providing the orientation angle and adding some arbitrary offset, as the network has no concept of the "zero position". However, this will be overcome by asking it for the angle of an image with our desired zero position, and removing this constant from future data.

The second method will use a structure similar to a U-net[12][6]. The contracting path will be of a similar structure to the CNN above and the expanding path will be an IGN (image generation network) designed for the project. One of the two stereoimages will be passed into the CNN, which will produce a single output (the angle). This number will then be fed into IGN; along with the angular separation. Using these two inputs, the IGN will create the second image, with its success being measured by how closely it is able to reproduce the unseen partner image. Once again, the CNN will be separable and can be used to identify the angle of a new image that is passed into the trained network. The loss of this model will be measured by how many pixels of the generated image are correctly coloured.

The network structures will need to be altered to accommodate the different input formats of the project progressions. While a CNN will always be used where they have been described, they will require layers of different types and in alternate combinations. For example, the initial structures will make heavy use of Conv2D Keras layers[4], as these take a two dimensional array as their input. That however, would be totally unsuitable for the projections of the 2D images, which will require Conv1D layers instead[4]. As a consequence, the networks will have only be useful in successive progressions as a guideline to the structures required, making each step a new challenge.

4 Progress

Most progress so far has been in the form of general research and understanding of neural networks in anticipation of the challenges to come. A simple network has been made for method 1 using images of squares. This has resulted in verifiable data generation and preprocessing functions being written that will be used throughout the project. Numpy meshgrid arrays are made and then populated with 1s and 0s to describe the image of a shape.

One of the main challenges, namely the periodicity of the angle definition, has already been encountered. The CNN struggles to define images in the $0-10°$ and $80-90°$ ranges. This was expected but a suitable method of resolving the issue has not yet been found. The expected solution is to create a new loss function[15]. The current

method uses the inbuilt Keras function 'mse'[5]. This computes the mean square error between the calculated and desired value/ground truth. This encounters an error when the values do not have 10 degree separation, eg $85°$ & $95°$ will be found as having $5 - 85 = -80°$ separation due to the rotational symmetry of a square. If a new loss function can be made that will not confuse this area of the results then the network will be improved.
The CNN structure was built as a series of three repeated structures consisting of a Conv2D layer, a MaxPooling2D layer and a Dropout layer[4][11]. The convolutional layers had 64, 32 & 16 filters respectively, all had 3*3 kernal size, the maxpooling layers had a 2*2 pool size and the dropout layers had 0.1 activation (values borrowed and edited from[1]). Finally, it was flattened and fully connected to a single output layer. The secondary network was simply a subtraction of the two outputs.

5 Plan & Conclusion

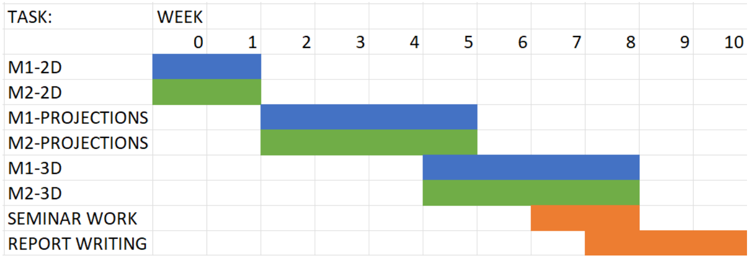

Figure 5.1: Gantt Chart showing project plan, If one of the two methods (M1&M2) proves less suitable to the task, it will be dropped.

This preliminary report establishes the groundwork for a multi-stage project for applying and comparing two unsupervised machine learning methods to the field of NP research. The integration of CNNs for the analysis of NPs through HAADF-STEM holds potential for insights into orientation determination of imaged NP structures, such as proteins. Artificial data generation has been provided already, along with an unrefined method for single rotational axis measurements. It is hoped this research will aid in NP synthesis for use in research and industry by aiding experimental measurements. Anticipated future work includes refining network structures and addressing rotational symmetry nuances, ensuring a robust foundation for the final report.

6 References

All websites checked as of 01:18 on Monday 11th December, 2023

References

[1] Jason Brownlee. https://machinelearningmastery.com/convolutional-layers-for-deep-learning-neural-networks/.

[2] Rahul Chauhan, Kamal Ghanshala, and R. Joshi. Convolutional neural network (cnn) for image detection and recognition, 12 2018.

[3] HAADF-STEM high-angle annular dark-field scanning transmission electron microscopy. www.jeol.com/words/emterms/20121023.031059.php#gsc.tab=0.

[4] Keras. https://keras.io/api/layers/.

[5] Keras. https://www.tensorflow.org/api_docs/python/tf/keras/losses/meansquarederror.

[6] Yafen Li, Wen Li, Jing Xiong, Jun Xia, and Yaoqin Xie. Comparison of supervised and unsupervised deep learning methods for medical image synthesis between computed tomography and magnetic resonance images. *BioMed Research International*, 2020:1–9, 11 2020.

[7] Katherine Macarthur, Lewys Jones, Sergio Lozano-Perez, Dogan Ozkaya, and P Nellist. Structural quantification of nanoparticles by haadf stem. *Journal of Physics: Conference Series*, 522:012061, 06 2014.

[8] Shinya Maenosono and Derrick Mott. Structural analysis of nanoparticles using scanning transmission electron microscopy, 07 2013.

[9] Aurélien Martinet, Cyril Soler, Nicolas Holzschuch, and François Sillion. Accurate detection of symmetries in 3d shapes. *ACM Trans. Graph*, 25, 04 2006.

[10] Max T. Otten. High-angle annular dark-field imaging on a tem/stem system. *Journal of Electron Microscopy Technique*, 02 1991.

[11] Alexander J. Pattison. Applying automation and machine learning to scanning transmission electron microscopy, 2020.

[12] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation, 10 2015.

[13] Saman Siadati. What is unsupervised learning, 08 2018.

[14] Tensorflow. Keras: The high-level api for tensorflow, https://www.tensorflow.org/guide/keras.

[15] Juan Terven, Diana-Margarita Cordova-Esparza, Alfonzo Ramirez-Pedraza, and Edgar Chavez-Urbiola. Loss functions and metrics in deep learning. a review, 07 2023.

[16] Anjeel Upreti. Convolutional neural network (cnn). a comprehensive overview, 08 2022.

[17] various Wikipedia. https://en.wikipedia.org/wiki/annular_dark-field_imaging.