

Convolutional Neural Network Applied for Nanoparticle Classification using Coherent Scaterometry Data

D. Kolenov^{1*}, D. Davidse¹, J. Le Cam², and S.F. Pereira¹

¹Optics Research Group, Imaging Physics Department, Faculty of Applied Sciences, Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands

²Institut d'Optique Graduate School, 2 Avenue Augustin Fresnel, 91120 Palaiseau, France

Abstract. The analysis of 2D scattering maps generated in scatterometry experiments for detection and classification of nanoparticle on surfaces is a cumbersome and slow process. Recently, deep learning techniques have been adopted to avoid manual feature extraction and classification in many research and application areas, including optics. In the present work, we collected experimental dataset of nanoparticles deposited on wafers for four different classes of polystyrene particles (with diameters of 40, 50, 60, 80 nm) plus background (no particles) class. We trained a convolutional neural network, including its architecture optimization, and achieved 95% accurate results. We compared the performance of this network to a existing method based on line-by-line search and thresholding, demonstrating up to a twofold enhanced performance in particle classification. The network is extended by a supervisor layer that can reject up to 80% of the fooling images at the cost of only rejecting 10% of original data.

1 Introduction

In this paper, we propose a method to classify 2D scattered maps of containing isolated nanoparticles on surfaces using CNN. These maps have been generated in a Coherent Fourier Scatterometry (CFS) setup [1–4]. In order to generate the scattered maps to train the data, polystyrene latex (PSL) nanospheres of calibrated sizes were spin-coated on a silicon wafer. The diameters of the particles ranges from 40 to 80 nm. Furthermore, for the classification, we study the areas of the wafer where the nanoparticles were absent, contributing to the “background” class. We also target at a novelty detection, by looking at ways for the network to separate the “unknown” class from the input data, i.e., classes that have been unseen in the training. In order to do that, we rely on a simple approach of baseline and also a more sophisticated approach of introducing the OpenMax layer [5].

We realized an experiment by adding noise to the scattered maps (degrading thus the SNR), fooling images as well as some unfamiliar (reversal) images to the network. As one of the main goals for the CNN is to accurately discriminating between classes, we studied the samples that contain multiple classes of PSL particles with diameters of 40 and 50 nm, 50 and 60 nm, 60 and 80 nm. The results show that our model can successfully discriminate between the 5 classes with accuracies up to 95%. By providing the samples that were unseen

*e-mail: d.kolenov@tudelft.nl

during the training, we highlight the importance of the novelty detection in order to capture the confusing inputs. The results show that the proposed method has superior capabilities compared to the classification with the traditional search algorithm.

2 Closed set classification

In order to see which classes the network struggles to distinguish and to what degree, we built the confusion matrix (Figure 1 A). The horizontal axis represents the particle classes predicted by our model, and the vertical axis represents the true input image labels. It is clearly visible that most misclassification involves the background class. The best model

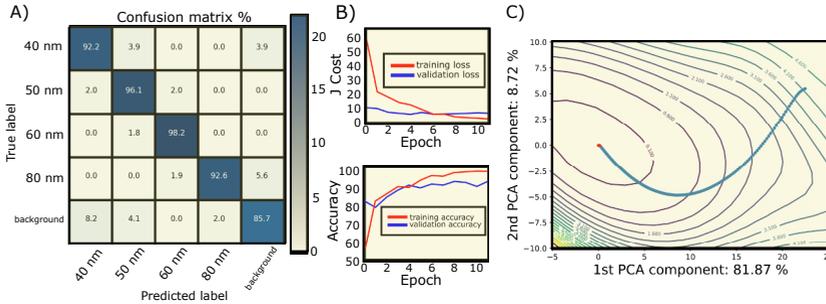


Figure 1. A) The accuracy for the test set in the confusion matrix. B) The training and validation loss - top and accuracy - bottom, is evaluated across different number of epochs based on the optimal architecture of the CNN model. C) 2D visualization of the loss surface of the CNN model with the projected learning trajectories using normalized PCA direction (batch size of 20, Adam optimizer and 15 epochs of training).

has to be selected based on the Accuracy metric calculated on the test data. For the closed set of five classes, we found that after approximately 12 epochs (12 times through all the training examples), the loss no longer decreased significantly. The top performing network ($Accuracy = 95\%$) was stored to be used in the further tasks (see Figure 1 B). Finally, we built the landscape, where we demonstrate the convergence to minima as our learning procedure follows the loss in a gradual manner. The projected learning trajectory is estimated using normalized principal component analysis (PCA) directions. The squared nature of the loss function leads to mostly convex loss landscape (Figure 1 C).

References

- [1] S. Roy and A. C. Assafrao and S. F. Pereira and H. P. Urbach, " *Opt. Express*, pp. 13250-13262, Jun 2014.
- [2] S. Roy and S. F. Pereira and H. P. Urbach and X. Wei and Omar El Gawhary, " *Phys. Rev. A*, 96, pp. 013814, Jul 2017.
- [3] Omar El Gawhary and Stefan J.H. Petra, "Method and apparatus for determining structure parameters of microstructures," *US Patent*, 2015.
- [4] D. Kolenov and R. C. Horsten and S. F. Pereira, "Heterodyne detection system for nanoparticle detection using coherent Fourier scatterometry," *Optical Measurement Systems for Industrial Inspection XI*, 11056, pp. 336-342, 2019.
- [5] Abhijit Bendale and Terrance E. Boult, "Towards Open Set Deep Networks," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1563-1572, 2015.