

Nanoparticle classification in wide-field interferometric microscopy by supervised learning from model

OGUZHAN AVCI,¹ CELALETTIN YURDAKUL,¹ AND M. SELIM ÜNLÜ^{1,2,*}

¹Department of Electrical and Computer Engineering, Boston University, Boston, Massachusetts 02215, USA

²Department of Biomedical Engineering, Boston University, Boston, Massachusetts 02215, USA

*Corresponding author: selim@bu.edu

Received 10 March 2017; revised 18 April 2017; accepted 20 April 2017; posted 20 April 2017 (Doc. ID 290361); published 11 May 2017

Interference-enhanced wide-field nanoparticle imaging is a highly sensitive technique that has found numerous applications in labeled and label-free subdiffraction-limited pathogen detection. It also provides unique opportunities for nanoparticle classification upon detection. More specifically, the nanoparticle defocus images result in a particle-specific response that can be of great utility for nanoparticle classification, particularly based on type and size. In this work, we combine a model-based supervised learning algorithm with a wide-field common-path interferometric microscopy method to achieve accurate nanoparticle classification. We verify our classification schemes experimentally by blindly detecting gold and polystyrene nanospheres, and then classifying them in terms of type and size. © 2017 Optical Society of America

OCIS codes: (180.3170) Interference microscopy; (150.1135) Algorithms.

<https://doi.org/10.1364/AO.56.004238>

1. INTRODUCTION

Interferometric nanoparticle imaging in a common-path configuration has gained significant attention for its ability to detect subdiffraction-limited low-index biological nanoparticles, and its simple, cost-effective, and high-throughput setup [1]. It enables highly sensitive detection of nanoscale particles by providing means to enhance the signal through the interference between the scattered and reflected reference fields. To do so, it uses a layered sensor typically comprised of a thin layer of SiO₂ atop a Si substrate in a common-path interferometry configuration. It achieves enhanced scattering of nanoparticles in the light collection direction given the optimized thickness of this glass layer, similar to engineering a dipole antenna directivity [2–4]. Its signal is also affected by the polarizability of the particle, amplitude of the reference field, and the phase lag between them, as discussed in more detail in [2].

The common-path interferometric nanoparticle imaging has been mainly demonstrated for label-free virus detection and sizing [5,6]. In these earlier studies, the particle detection and subsequent sizing relied on the intensity reading at a single focal plane. Under ideal circumstances, where the sample only contains a certain kind of nanoparticles (i.e., optical properties known *a priori*), this type of blind digital detection and sizing of nanoparticles based on a single focal plane image could yield reliable and repeatable results. However, polydispersity in nanoparticles and morphological variations on the sensor

surface can lead to inaccurate detection and sizing in a single focal plane image. In such a case, a more robust detection and particle characterization is deemed necessary, and can be realized by the defocus data stack acquisition strategy [2]. In fact, interferometric nanoparticle imaging can provide unique opportunities for the detection and visualization of weakly scattering subdiffraction-limited particles, as well as for their classification upon detection. Recently, with the development of a rigorous physical model [2], it has been suggested that a more robust approach toward nanoparticle detection and discrimination can be achieved [7]. In particular, the distinctive defocus signature of the interferometric signal can be of great utility for classifying nanoparticles, as the additional phase introduced by the axial shift results in a particle-specific signature of nanoparticle response, as shown in Fig. 1(b) [2]. Here, the defocus is defined as the axial shift of the top sensor surface from the focal plane of the objective lens, as illustrated in Fig. 1(a).

A robust nanoparticle classification can especially be of great significance for high-throughput interferometric biological-particle imaging, where thousands of particles are detected simultaneously within a single field-of-view in a labeled and/or label-free fashion. For instance, when a multiplex detection of label-free viruses and antigens labeled with metallic particles takes place on the same sensor surface [8], an accurate classification of the particles in the image becomes extremely important for accurate quantitative measurements. In another

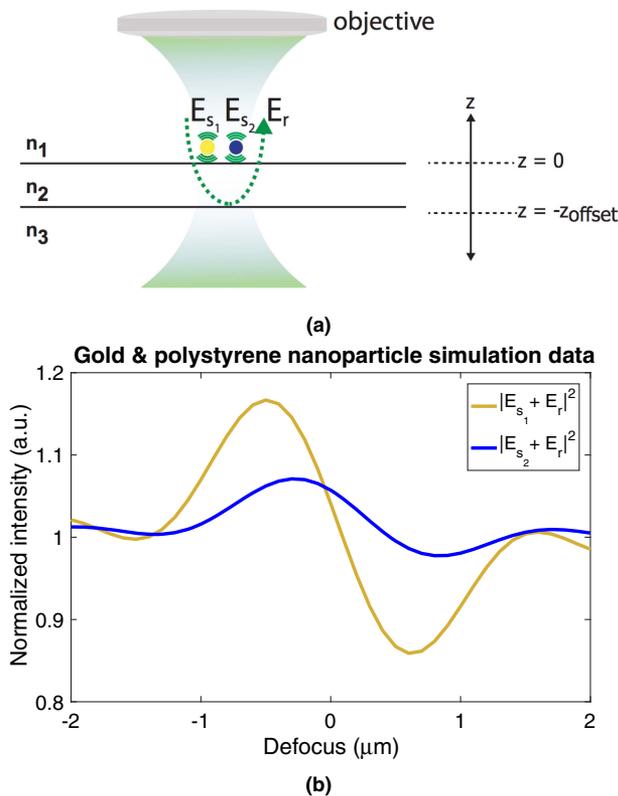


Fig. 1. Defocus in common-path wide-field interferometric nanoparticle imaging: (a) schematic of the setup illustrating the case when sensor surface is in the same plane as the focal plane of the objective ($z = 0$) (adapted from [2]); the defocus takes place by moving this layered substrate in the axial direction (z), (b) the simulated interferometric signal for gold nanosphere in 30 nm radius and polystyrene nanosphere in 37 nm radius.

case, an accurate size discrimination of a polydisperse pathogen captured on the sensor surface can be of great importance in certain diagnostic applications. Therefore, it is clear that upon detecting the target nanoparticles with interferometric microscopy, there is a need for a reliable nanoparticle classification, particularly in terms of type and size. In the light of this need, we propose a way to classify interferometrically detected nanoparticles based on size and type using an experimentally obtained interferometric signal and the physical model detailed in [2]. A direct fit approach using the analytical model might, at first, seem to be the way for this type of classification; however, doing so quickly becomes computationally expensive given the large parameter space and the high throughput of our optical system, where up to 10^6 nanoparticles can be detected simultaneously. The computational burden of having to fit each particle response for every single detection event can be circumvented by adapting a machine learning approach, where a simulation-based data set can be used as a training set in advance (offline). One advantage of using a simulation-based training data set is that large training data sets can be generated, eliminating the need for manual effort for a training data set and the “human error” that comes with it. The defocus signatures for nanoparticles of different types and sizes are utilized to realize a machine-learning-based nanoparticle classification

scheme using a support vector machine classifier (SVM). While there exist many supervised learning classifiers (e.g., bagging decision trees, boosted trees, k -nearest neighbors algorithm, etc.) that can be applied to the problem at hand, we choose SVM as our classification algorithm, owing to its well-established utility in various aspects of image classification, ranging from face recognition to cell sorting [9–14]. SVM classifiers essentially construct hyperplanes that linearly separate the labeled feature vectors in the training data set.

The training data set in this technique is provided by the simulated interferometric signals for different type and size nanoparticles. Once the supervised learning with simulated data is carried out, the experimental observations are fed into the algorithm, where classification takes place in two steps: first the nanoparticle type is determined in terms of its dielectric characteristics, i.e., resonant (e.g., gold) or nonresonant (e.g., polystyrene), and second, given the type, the nanoparticle size is determined. We experimentally validate our technique by imaging a sample that has a mixture of gold and polystyrene nanospheres, and then classifying the detected nanoparticles in terms of type and size. Our method has a potential application in multiplexed interferometric microscopy experiments, where classification of nanoparticles can not only provide further information about the target, but also eliminate the false count of spurious signal (noise) due to nonspecific binding events.

2. CLASSIFICATION ALGORITHM

A support vector machine (SVM), a supervised learning algorithm, can realize the hyperplane with the largest margin that optimizes the separation of a given labeled training data set of (\mathbf{x}_i, y_i) , with \mathbf{x}_i being a vector with N features and y_i being its label, resulting in the classification function $f(\mathbf{x}) = y$ [13]. The hyperplane is in the N -dimensional vector space (i.e., R^N) of N features, which is the total number of the data points of the vector \mathbf{x}_i . Moreover, the hyperplane in an SVM classifier realizes the linear separation of the corresponding labels of the input vectors. In this context, we employ an SVM on nanoparticle classification in terms of type and size distribution.

As discussed previously, in wide-field common-path interferometric microscopy, particles exhibit unique defocus signatures that help us realize their size and type. This study relies on the characteristics of those defocus signatures to propose a classification scheme in terms of nanoparticle type and size by learning from the simulations, making use of an SVM classifier. The training data set contains 41 features corresponding to the interferometric nanoparticle response at 41 defocus points whose range is determined by the depth of field of the optical system. The simulations span the following parameter space for gold and polystyrene nanoparticles: $15 \text{ nm} \leq r \leq 75 \text{ nm}$, $520 \text{ nm} \leq \lambda \leq 530 \text{ nm}$, where r is the nanoparticle radius and λ is the illumination wavelength.

For sake of convenience, we use the built-in Matlab SVM classification functions, *fitcsvm* and *fitcecoc*, which can support binary and multiple classes, respectively. The SVM classifier is set to a nonlinear kernel function (e.g., *Gaussian*) to handle the complexity due to the nonlinear features of the vectors [14], providing more accurate predictions with the test data set on the trained classifier. The supervised learning for particle

type determination considers only the dielectric characteristics of the nanoparticle as a class label, whereas the supervised learning for particle size determination considers only the radius (r) as a class label for a given type. We compartmentalize the classification into two steps: we first classify the type, and then proceed with size determination for each population. This two-step approach provides more accuracy as opposed to training and classifying the type and the size in one step.

A. Particle Type Classification

In common-path interferometric microscopy, the resonant particles (e.g., gold nanoparticles) differ from the low-index ones (e.g., polystyrene nanoparticles) by exhibiting an equally strong positive and negative peaks in their defocus signals, owing to their negative dielectric characteristics [2]. This nanoparticle signature difference between resonant gold and nonresonant polystyrene nanoparticles can be observed from Figs. 2(a) and 2(c). This phenomenon, in turn, renders resonant nanoparticles distinguishable from the low-index nanoparticles in common-path interferometric microscopy. Therefore, we are able to identify the particle types by using the built-in Matlab SVM function (*fitcsvm*), associated with binary-class model classification upon training with the simulated defocus curves that are single-labeled according to the particle type.

B. Particle Size Classification

For a given particle type, size distribution of the particle can be determined owing to its radius-dependent defocus signal. The amplitude of the defocus signal is size dependent such that it scales with the particle polarizability, which is a function of

particle volume. For instance, the peak-to-peak normalized intensities for polystyrene spheres of radius 25, 30, and 35 nm are 2.5%, 3.4%, and 6.5%, respectively. Thus, SVM can realize the size distribution of a given particle population from the size-dependent peak-to-peak intensity value of the defocus signal. The built-in Matlab SVM function, *fitcecoc*, is capable of multiclass model training. We use this function to train the classifier with the simulated data set for a particular type of nanoparticles of various sizes ($15 \text{ nm} \leq r \leq 75 \text{ nm}$) with the labels corresponding to the particle size.

3. EXPERIMENTAL RESULTS

In this section, we experimentally verify our nanoparticle classification algorithm. The experiment uses a sample that consists of gold nanospheres of 29 nm nominal radius and polystyrene nanospheres of 37 nm nominal radius spin-coated on a SiO_2/Si sensor chip with 100 nm thick silicon dioxide. The experimental defocus scan (consisting of 41 defocus planes) of the sample data is taken using the wide-field interferometric microscopy setup detailed in [3], and the data is processed with a custom Matlab code that performs the difference of Gaussian spatial filtering to find nanoparticles in the image [7]. It is important to note that the code finds the locations of the diffraction-limited nanoparticles in the images regardless of their type and signal strength. The exposure time and number of frames averaged for each defocus image is set to 100 ms and 10, respectively, to ensure data acquisition in the shot-noise-limited regime. Following this initial step of nanoparticle detection in

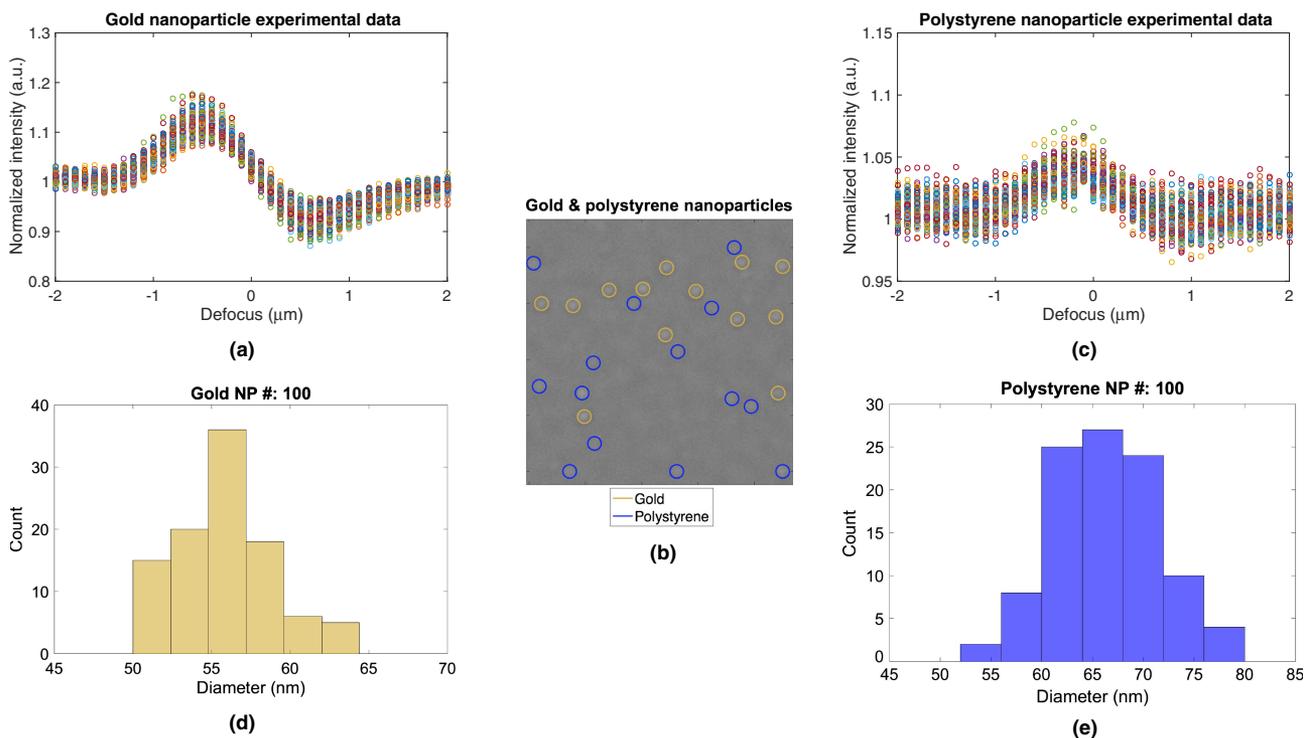


Fig. 2. (a) Experimentally obtained defocus curves of gold nanospheres upon supervised-learning-based type classification. (b) Image of gold and polystyrene nanospheres circled in yellow and blue, respectively. (c) Experimentally obtained defocus curves of polystyrene nanospheres upon supervised-learning-based type classification. (d) and (e) Nanoparticle size histograms of the type-classified gold and polystyrene nanospheres, respectively.

the images, for each nanoparticle, we obtain its defocus response. A total of 200 particles are detected from a single field-of-view. Each defocus datum is then first processed with the classification algorithm for the type determination. Figure 2(b) shows a zoomed-in cropped image from a single field-of-view, where gold and polystyrene nanospheres are distinguished from one another using this classification algorithm. Notice that the type classification works well due to the unique defocus curves that gold and polystyrene nanospheres exhibit. That is to say, resonant metallic nanoparticles exhibit an interferometric defocus signal that has strong positive and negative peaks within an approximately micron defocus range, whereas the nonresonant low-index nanoparticles only exhibit a strong positive peak, as illustrated in Fig. 1(b). Based on its assigned type, each detected nanosphere defocus response is demonstrated in Figs. 2(a) and 2(c), showcasing the discrimination performance with 100% accuracy using the artificial sample that only contains gold and polystyrene nanoparticles for the reader's reference.

The second part of the classification algorithm focuses on the size determination of the particles, taking into account their previously determined types. Since the normalized signal strength in the defocus data is size dependent, different sized nanoparticles of the same type can easily be distinguished among themselves. Therefore, by dividing the classification algorithm in two steps, not only do we simplify the classification problem that is normally nonlinear, but also ensure high accuracy. We experimentally verify the second step of this algorithm by further processing the previously detected and type-classified gold and polystyrene nanoparticle signals. We obtain the size histograms of the detected nanoparticles as presented in Figs. 2(d) and 2(e), which are in reasonable agreement with their nominal sizes provided by their manufacturers. Apart from the innate variations in size and shape during the nanoparticle synthesis, perhaps the slight deviations from the nominal sizes of each particle type can also be attributed to the spectral width of the light source ($\Delta\lambda \approx 40$ nm), and imperfections in the experimental setup such as misalignments, back reflections, and vibrations.

4. DISCUSSION AND CONCLUSION

In this study, we have demonstrated nanoparticle classification in a wide-field interferometric microscopy scheme by combining its powerful nanoparticle imaging capability with a model-based supervised learning algorithm. The proposed classification method accomplishes high accuracy, as experimentally verified in Fig. 2. It is imperative to note that the nanoparticle signal in common-path interferometric microscopy depends on various parameters, as examined in [2]. The multivariable dependency of the interferometric signal poses a challenging problem to decipher its constituents. However, this multi-parameter-dependent nanoparticle signal also opens up possibilities in terms of target particle analysis, following its detection. Certain *a priori* information/assumptions are therefore necessary, not only to simplify this inverse classification problem, but also to maintain accuracy. One of the limitations imposed by this particle classification scheme is its size range and shape.

In the scope of this study, we only assume particles of spherical shape, although the polarizability calculations can be extended to incorporate the classification of prolate particles. The size constraint, on the other hand, stems from the fact that the dipole approximation that is used in the physical model starts to break down as the particle size gets closer to the illumination wavelength—i.e., the electric dipole model to represent the scattering from nanoparticles holds true only when the size is much smaller than the wavelength. It is imperative to note that the dielectric characteristics of the particles are predetermined when carrying out the size classifications. Typically, the biological particles that are of great interest in biomedical studies exhibit similar nonresonant dielectric characteristics as polystyrene nanoparticles.

In conclusion, we have successfully demonstrated a nanoparticle classification scheme in two steps: first determining the particle type, and then realizing its size distribution. We have shown the utility of our model-based classification method in differentiating resonant gold nanoparticles from nonresonant polystyrene nanoparticles, as well as in sizing of the type-classified gold and polystyrene particles whose nominal radiuses are given as 29 nm and 37 nm, respectively. Our classification model has a potential impact in nanoparticle detection using interferometric microscopy, where it allows for simultaneous use of labeled (e.g., with gold nanoparticles) and label-free detection modalities, as well as sensitivity improvements by accurately eliminating the count of nonspecifically bound particles, revealing further information about target particles in terms of their type and size.

Acknowledgment. O. A. gratefully acknowledges support from I/UCRC Center for Biophotonic Sensors and Systems (CBSS).

REFERENCES

1. O. Avci, N. L. Ünlü, A. Y. Ozkumur, and M. S. Ünlü, "Interferometric reflectance imaging sensor (IRIS)—a platform technology for multiplexed diagnostics and digital detection," *Sensors* **15**, 17649–17665 (2015).
2. O. Avci, R. Adato, A. Y. Ozkumur, and M. S. Ünlü, "Physical modeling of interference enhanced imaging and characterization of single nanoparticles," *Opt. Express* **24**, 6094–6114 (2016).
3. O. Avci, M. I. Campana, C. Yurdakul, and M. S. Ünlü, "Pupil function engineering for enhanced nanoparticle visibility in wide-field interferometric microscopy," *Optica* **4**, 247–254 (2017).
4. L. Novotny and B. Hecht, *Principles of Nano-Optics* (Cambridge University, 2012).
5. G. G. Daaboul, C. A. Lopez, J. Chinnala, B. B. Goldberg, J. H. Connor, and M. S. Ünlü, "Digital sensing and sizing of vesicular stomatitis virus pseudotypes in complex media: a model for ebola and marburg detection," *ACS Nano* **8**, 6047–6055 (2014).
6. S. M. Scherr, G. G. Daaboul, J. Trueb, D. Sevenler, H. Fawcett, B. Goldberg, J. H. Connor, and M. S. Ünlü, "Real-time capture and visualization of individual viruses in complex media," *ACS Nano* **10**, 2827–2833 (2016).
7. J. Trueb, O. Avci, D. Sevenler, J. Connor, and M. S. Ünlü, "Robust visualization and discrimination of nanoparticles by interferometric imaging," *IEEE J. Sel. Top. Quantum Electron.* **23**, 394–403 (2017).
8. M. S. Ünlü, "Digital detection of nanoparticles: viral diagnostics and multiplexed protein and nucleic acid assays," *MRS Online Proc. Libr.* **1720** (2015).

9. O. Chapelle, P. Haffner, and V. N. Vapnik, "Support vector machines for histogram-based image classification," *IEEE Trans. on Neural Networks* **10**, 1055–1064 (1999).
10. B. Heisele, P. Ho, and T. Poggio, "Face recognition with support vector machines: global versus component-based approach," in *Proceedings 8th IEEE International Conference on Computer Vision (ICCV)* (2001), Vol. **2**, pp. 688–694.
11. N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* (2005), Vol. **1**, pp. 886–893.
12. A. Subasi, "Classification of EMG signals using PSO optimized SVM for diagnosis of neuromuscular disorders," *Comput. Biol. Med.* **43**, 576–586 (2013).
13. T. S. Furey, N. Cristianini, N. Duffy, D. W. Bednarski, M. Schummer, and D. Haussler, "Support vector machine classification and validation of cancer tissue samples using microarray expression data," *Bioinformatics* **16**, 906–914 (2000).
14. I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene selection for cancer classification using support vector machines," *Int. J. Mach. Learn. Cybern.* **46**, 389–422 (2002).