2021 IEEE 21st International Conference on Nanotechnology (NANO)
July 28-30, 2021. Montréal, Canada

Size Based Characterization of Gold Nano Particles using Machine Learning Approach

P. Senoamadi, S.Krishnannair, *Member,IEEE.,* and L.Sikhwivhilu

Abstract— **In drug delivery, there is a need for precision in reporting particles parameters. Studies have shown that absorbency of drugs in the blood stream depends on the size of the nanoparticle. The shape and size of nanoparticles (NPs) matter the most, hence the distribution of NP depends on the size and shape of NPs. By synthesizing and characterizing the NPs, we are able to cluster and get the amount of a certain type of morphology and accurate size determination. Moreover, the size distribution of a particle plays a more important goal as it possesses an increase in the usability of a diagnostic and therapeutic tool in medicine. The shape and size distribution of NPs is important for the delivery of drugs and for the cure or treatment of several chronic diseases such as cancer. Hence it is important to get the accurate size distribution of NPs for better results. Gold nano particles (AuNPs) where measured manually by the use of transmission electron microscope, hence, in most cases human error could play part in terms of inaccurate measurements. The digital images of AnNPs contain noise, making it difficult to get accurate measurements using the transmission microscope. AuNPs were measured in terms of their width and length. This study focused on the characterization of AuNPs collected by the transmission electron microscope using machine learning approaches. Image preprocessing and processing techniques are used for extracting the features (length and width) of AuNPs. In this study, filtering techniques such as Gaussian blur, Median and Mean filtering techniques are employed for noise removal to increase the precision in estimating the size of NPs. Unsupervised machine learning algorithm such as K-means and Otsu are used to perform image segmentation of the filtered nano images for the accurate extraction of particles' features such as length and width. The size measurements obtained using the machine learning approaches are compared with the measurements taken by the transmission electron microscope (TEM) for error estimation in the size distributions of NPs. The results showed that machine learning approaches provided accurate measurements of most of the NPs as compared to TEM. Therefore, it is recommended that machine learning approaches can be used to estimate the size of NPs so that the shapes can be described better and classified during the synthesis process.**

*Resrach supported by MINTEK.

P.Senoamadi is with the Department of Mathematical Sciences, University of Zululand, Kwadlangezwa, 3886 South Africa.

S. Krishnannair is with the Department of Mathematical Sciences, University of Zululand, Kwadlangezwa, 3886 South Africa. (Phone:0027359026235; e-mail: krishnannairs@unizulu.ac.za).

L.Sikhwivhilu is with Mintek, 200 Malibongwe Dr, Randburg, South Africa

I. Introduction

Nanoparticles (NPs) are microscopic particles that are at least one dimension less than 100 nanometers (nm). NPs have increasingly shown several applications in the field of nanotechnology lately. Those applications are drug delivery, bio-detection of pathogens, delivery of proteins, etc. NPs can be classified in terms of size, morphology, and physico-chemical properties as metal, ceramics, semi-conductor, polymer, and lipid-based nanoparticles. On the other hand nanostructures are classified as nano-spheres, liposomes, nano tubes, quantum dots, nano-rods (NRs) and other nano structures [1]. In most cases, nanoparticles are used to deliver drugs into cancer cells. The size of NPs matters most in drug delivery for the penetration of a human cell's interior without damaging it. The size, shape and other features affect the way the absorption of delivered drugs behaves in the human system.

The self-assembly and virtualization of these NPs can be done thorough the following microscopes: transmission electron microscope (TEM), atomic force microscope (AFM) and scanning electron microscope (SEM). These microscopes vary according to their capabilities and functionality. There are advantages and disadvantages of using electron microscopes. It has been elaborated by N. Marturi and co-workers [2], that electron microscopes like SEM produce noisy images that will lead to difficulties in performing image segmentation or analysis. One of the advantages of using gold nano particles (AuNPs), is its malleable characteristic to be able to shine when exposed to light, thus excellent optical properties. In modern days, AuNPs are widely used for cancer detection, by injecting AuNPs into the blood stream to target cancer cells. It has been proven scientifically that AuNPs could be used in the fight against cancer tumors, which has been shown to be safe and effective. According to the study by Y. Kherde and co-workers [3], it has been highlighted that AuNPs has the ability to be easily synthesized and are colloidally stable, and they can be conjugated with biological molecules in a straight forward way. AuNPs have been mostly used in the field of nano-technology due to their unique natureof being more versatile and permeable [4].

Section II describes the data used and the methods used for the characterization of AuNPs. Section III presents the results obtained by the use of the proposed method described in section II. Section IV presents the conclusion of the study.

II. METHODOLOGY

A. Data

The self-assemble of AuNPs was done at the DSI/MINTEK Nanotechnology Innovation Centre laboratory (Johannesburg, South Africa). The visualization of these AuNPs was done using the transmission electron microscope (TEM). Data used in this study is in the form of TEM images, which are digital images. Fig. 1 below represents an original image taken by means of TEM. Fig. 1 presents a digital image in a greyscale format. The analysis of the images in this study was conducted on some cropped images from the set of images taken by TEM as shown in Fig. 1.

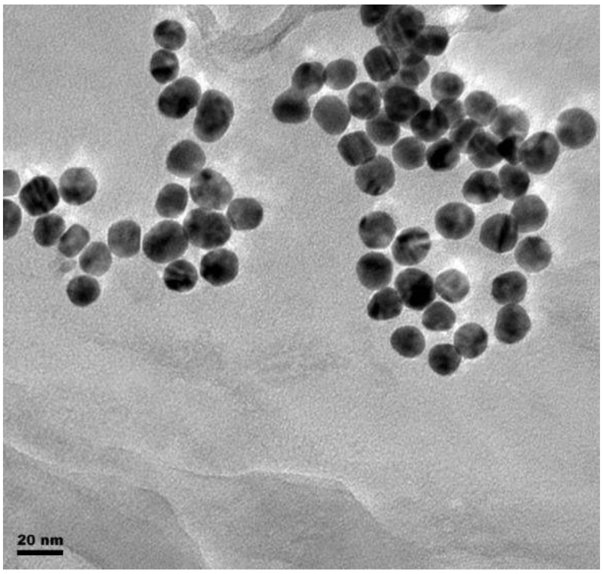


Figure 1. TEM image of AuNPs

Filtering techniques used in this study are mean, median and Gaussian blur filtering. These filters are used to reduce noise and to enhance the features so that segmentation and measurements can be estimated effectively. The main tasks for image processing may include image filtering, segmentation, binarization, depending on what specific task or the goal one is intending to perform [5]. In this study, the above mentioned filtering techniques were used in the AuNPs given in Fig. 1. Most particles were agglomerated, and watershed method was used to separate particles that are overlapping. This filter assumes that out-of-image pixels have a value equal to the nearest edge pixel. This gives higher weight to edge pixels than pixels inside the image, and higher weight to corner pixels than non-corner pixels at the edge [6].

Adaptive Median Filter (AMF) handles this type of problems easily without any blurring or removing important features on an image. According to the study by R.Roy and co-workers [7], adaptive thresholding method has shown great results by their capability for combining the edge and gray-level-information to get the best thresholding surface to validate the process. The dispersion of the pixels which is represented by the standard deviation did not make greater impact than the Gaussian filter. The standard deviation was reduced from 49.472 to 48.206 using AMF. Hence, noise was effectively removed using Gaussian filter. The filtered images are shown in Fig. 2.

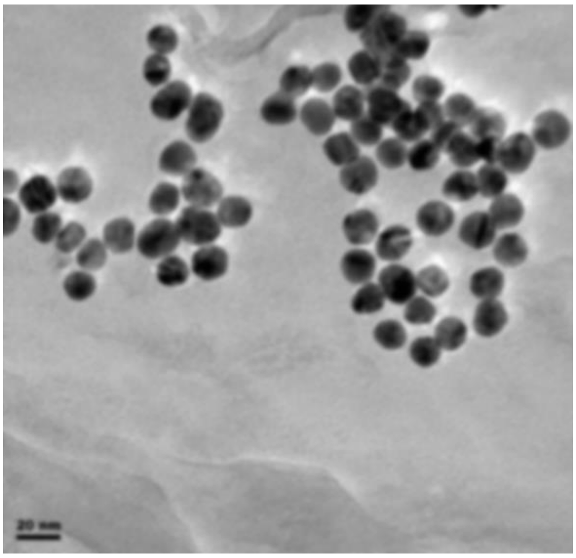


Figure 2. Images of AuNPs filtered using Gaussian filter

K-means is an unsupervised machine learning algorithm that is well known for pattern recognition. It is the most used algorithm for image segmentation. As reported by [8], using Otsu clustering and k-means clustering methods are useful in order to get a better segmentation of TEM images and get the best characterization of the gold nanoparticles (AuNPs). Given a digital image, x and y were let to represent the image pixels. Then a pixel of size $x * y$ was obtained. The image will be partitioned into k number of partitions. Then (x, y) will be taken as input pixels to be partitioned. Let Ck, be the cluster centers of partitions. Firstly k, the number of partitioned clusters, and the center are initialized. For each pixel of an image, the distance d was computed, between the center and each pixel [9]. The distance formula to be used is given below:

$$d = |p(x, y) - C_k| \tag{1}$$

Then, the data points to the closest center in relation to d was assigned. Hence, after all the pixels have been assigned to their center, the new center is re-computed using equation (2) given below and the process is iterated until the distance is reduced, that is d < ε, where ε is a small positive number.

$$C_k = \frac{1}{k} \sum_{y \in C_k}^{n} \sum_{x \in C_k}^{n} p(x, y) \tag{2}$$

The general idea behind the k-means clustering is to sketch a histogram with clusters being in the foreground and background. The value of k will have to be fixed at two, as the image into background and foreground is partitioned. The ε is given as a smaller number that will act as a tolerance value in the iteration of K-means clustering algorithm. The input dataset is the image pixels. Then, each pixel finds the center to which it is closest. For these new clusters, again a center is found using equation (2). This is repeated until no clusters change. The final clusters form the segmented image into a foreground (Black) and background (White) [9].

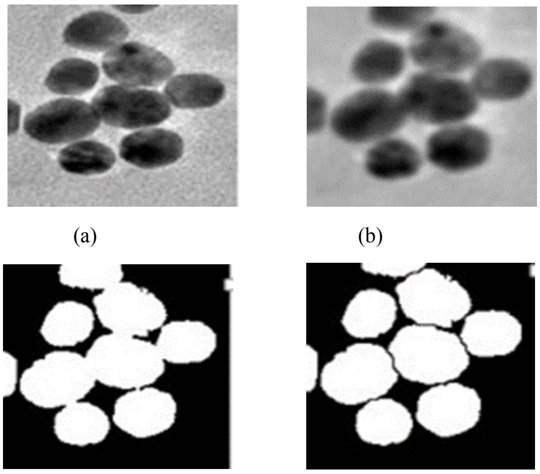
III. RESULTS

Standard deviation was used to assess how pixels spread after the filtering methods have been applied as shown in Table 1. In Table 1, F1, F2 and F3 present filters at different pixels levels R1, R2 and R3. Choosing the radius more than three pixels away from the center, the images have shown loss of quality and the images were too blur, hence particles lose more features. Table 1 below presents the standard deviations (S.D) of pixels taken after applying the filtering techniques.

TABLE I. STATISTICS OF DIFFERENT FILTERING TECHNIQUES

		F1		F2		F3	
Filters	*S.D*	*R1*	*S.D*	*R2*	*S.D*	*R3*	*S.D*
Median	49.5	1	49.01	2	48.2	3	47.2
Mean	49.5	1	48.99	2	47.4	3	46.1
Gaussian	49.5	1	48.25	2	45.9	3	40.2

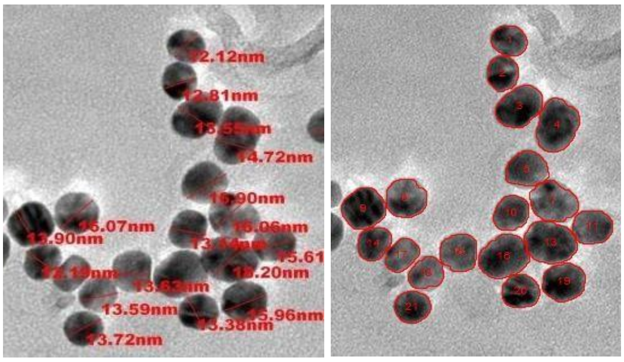
From Table 1, F2, the images have shown more improvement on the visual assessment and the standard deviation. To separate the overlapping of nanoparticles, we start by the input of an image followed by blurring an image using Gaussian blur method. This follows partitioning of an image using K-means clustering to black and white. Hence the particles were separated using a watershed method on a binary image. The above mentioned method finds the center of each object (using a morphological erode operation), then calculates a distance map from the object center points to the edges of the objects, then fills that "topological map" with imaginary water. Where two "Watersheds" meet, it builds a line that separate particles. Fig. 3 (a)-(d) below show the process followed to deal with overlapping particles in this study. The dispersion of the pixels which is represented by the standard deviation made no significant impact when compared to the Gaussian filter. The standard deviation was reduced from 49.472 to 48.206. Hence, noise was not effectively removed. Median and Mean filter have shown a great decrease in standard deviation from a standard deviation 49.472 (original image) to 48.2 and 47.4 respectively. However, Gaussian blur filter was able to decrease from 49.472 to 45.9, which has shown a great optimization of filtered image. Hence, noise was effectively removed. Therefore, the Gaussian filter image was selected to be the optimal solution with a reduced standard deviation of 45.9.



(a) (b) (c) (d)

Figure 3. (a) Original TEM Image, (b) blurred/filtered, (c) K-means Thresholder, (d) Watershed process (particle separation).

Fig. 3 (a) presents the original image used for the analysis. Fig. 3 (b) presents the blurred image followed by Fig. 3 (c) with the partitioned image in black and white. Fig. 3 (d) presents the separated particles as the results of the watershed method used. The gold nanoparticles used in this study takes the morphology of a spherical shape. A digital image is made up of rows and columns of pixels. A pixel in such an image can be specified by determining which column and which row it contains. In terms of coordinates, a pixel can be identified by a pair of integers giving the column number and the row number. Fig. 4 (a) below present the AuNPs and its measurements found using the TEM. While, Fig. 4 (b) shows the AuNPs from processed image with the proposed method. The procedure to get the end results is obtained after performing image filtering and thresholding. Table 2 below shows the estimation of the length of each image by the proposed method and TEM using the images in Fig. 4 (b). Table 2 below presents the measurements between the proposed method and TEM. The error of estimation in the last column in Table 2 is obtained by calculating difference between the measurements using TEM and the proposed approach. It is evident that there exists an error between TEM and proposed method measurements. In most of the cases the measurements obtained using the proposed approach were smaller than those obtained by TEM. This is due to TEM calibration error in the use of TEM for the estimation of the particle size. The mathematical approach used in this study managed to measure all the particle size, while the other AuNPs from the TEM images were not measured. Furthermore, the agglomeration of particles was managed to be dis-attached by the watershed method. Based on the results, by the implementation of K-means clustering for segmenting the AuNPs image and Gaussian blur for filtering the images, we can conclude that, the measurements taken by the TEM and the mathematical approach have shown an average error that is ~2.85 that concludes that there are human errors and systemic errors associated with manual TEM measurement collection. Thus the proposed method in the study can be used as an alternative strategy to measure the size of nano images taken by TEM to minimize the calibration error of TEM in the estimation of its size distribution. This could be extended to other electron microscopes like SEM and possibly beyond.



(a) (b)

Figure 4. (a) Measurements obtained by TEM, (b) Measurements estimated using proposed mathematical approach.

TABLE II. LENGTH OF NANO IMAGES OBTAINED BY TEM AND MACHINE LEARNING APPROACH

Label	Proposed Method	TEM	Error
1	12,17	13,84	1,67
2	10,21	12,43	2,22
3	12,51	14,18	1,67
5	12,34	14,36	2,02
6	12,49	14,56	2,07
7	12,59	15,41	2,82
8	14,85	16,22	1,37
9	11,53	13,57	2,04

IV. CONCLUSION

The main objective of this paper is to characterize gold nanoparticles' (AuNPs) images taken by the transmission electron microscope (TEM). Using mathematical and machine learning approaches, the images were processed and the lengths of each particle were estimated. The proposed approach was found to be effective in measuring all the particles including overlapping images while TEM had limited effectiveness in measuring all the particles in the images taken. The TEM method measured the size of the overlapping particles considering them as single particle of large size and hence incorrectly measured the longest length from one edge to other. On the other hand, the proposed method managed to detach the nanoparticles and hence take the measurements individually. Furthermore, the calibration error of TEM was investigated by calculating the measurement difference between TEM and the proposed method. This investigation is important for the accurate size distribution of nano particles for the manufacturing of drugs in the field of drug delivery. Furthermore, measurements taking microscope like TEM can be implemented with more extensive image processing methods to get better results. For further investigation, more agglomerated AuNPs can be used.

REFERENCES

[1] N. Kumar and S. Kumbhat, "*Concise Concepts of Nanoscience and Nanomaterials*", scientific publishers, 2018.

[2] N.Marturi, S.Dembele, and N.Piat, "Scanning electron microscope image signal-to-noise ratio monitoring for micro-nano manipulation, *J.Sc.Microsc*, vol. 36, no.4, pp.419-429, 2014.

[3] Y.Kherde, Z.P.Aguilar, L.L.C.Zystein, H.M.Rodgers and W.Hamilton, "Gold nanoparticles: various methods of synthesis and antibacterial applications", *Front.Biosci*,vol.19, pp.1320-1344.

[4] M.Oishi, A. Tamura, T. Nakamura and Y.Nagasaki, "A smart nanoprobe based on fluorescence-quenching PEGylated nanogels containing gold nanoparticles for monitoring the response to cancer therapy", *Adv.Funct.Mater*, vol.19, no.6, pp.827-834, 2009.

[5] J.Sauvola and M.Pietikainen, "Adaptive document image binarization", *Pattern.Recognit*, vol.33, no.2, pp.225-236.

[6] D.J.Groom, K.Yu, S.Rasouli, J. Polarinakis, A.C. Bovik and P.J.Ferreira, "Automatic segmentation of inorganic nanoparticles in BF TEM micrographs", *Ultramicroscopy*, vol.194, pp.25-34, 2018.

[7] P.Roy, S.Dutta, N.Dey, G.Dey, S.Chakraborty and R.Ray, "Adaptive thresholding: a comparative study", IEEE *ICCICCT*, pp.1182-1186, 2014.

[8] D.Liu and J.Yu, "Otsu method and K-means", IEE *9th HIS*, vol.1, pp.344-349, 2009.

[9] K.Anandan, "Image processing of nanostructures", *Ph.D.dissertation*, Univ. Georgia, 2009.