



Article

Statistically Representative Metrology of Nanoparticles via Unsupervised Machine Learning of TEM Images

Haotian Wen ^{1,*}, José María Luna-Romera ², José C. Riquelme ², Christian Dwyer ³ and Shery L. Y. Chang ^{1,4,*}¹ School of Materials Science and Engineering, University of New South Wales, Sydney, NSW 2052, Australia² Software and Computing Systems, Universidad de Sevilla, 41004 Seville, Spain; jmluna@us.es (J.M.L.-R.); riquelme@us.es (J.C.R.)³ Electron Imaging and Spectroscopy Tools, Sydney, NSW 2219, Australia; dwyer@eistools.com⁴ Mark Wainwright Analytical Centre, Electron Microscope Unit, University of New South Wales, Sydney, NSW 2052, Australia

* Correspondence: haotian.wen@unsw.edu.au (H.W.); shery.chang@unsw.edu.au (S.L.Y.C.)

Abstract: The morphology of nanoparticles governs their properties for a range of important applications. Thus, the ability to statistically correlate this key particle performance parameter is paramount in achieving accurate control of nanoparticle properties. Among several effective techniques for morphological characterization of nanoparticles, transmission electron microscopy (TEM) can provide a direct, accurate characterization of the details of nanoparticle structures and morphology at atomic resolution. However, manually analyzing a large number of TEM images is laborious. In this work, we demonstrate an efficient, robust and highly automated unsupervised machine learning method for the metrology of nanoparticle systems based on TEM images. Our method not only can achieve statistically significant analysis, but it is also robust against variable image quality, imaging modalities, and particle dispersions. The ability to efficiently gain statistically significant particle metrology is critical in advancing precise particle synthesis and accurate property control.

Keywords: nanoparticles; image analysis; machine learning



Citation: Wen, H.; Luna-Romera, J.M.; Riquelme, J.C.; Dwyer, C.; Chang, S.L.Y. Statistically Representative Metrology of Nanoparticles via Unsupervised Machine Learning of TEM Images. *Nanomaterials* **2021**, *11*, 2706. <https://doi.org/10.3390/nano11102706>

Academic Editor: Nikos Boukos

Received: 30 July 2021

Accepted: 6 October 2021

Published: 14 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Nanoparticles are widely used in applications such as bio-medicine [1–4], sensors [5,6], nanoscale thermal metrology [7], energy [8], environmental protection [9], optics [10–13] and electronics [14,15]. The size, shape, and dispersion of the particles play an important role in their properties. Examples include the enhancement of the catalytic activities of nanoparticles by controlling their shape [16,17]; the particle size- and shape-dependent drug molecule absorption in drug delivery [18]; the strong size/shape correlation with their optical properties [19,20]; and control of the size and dispersion of nanoparticles to optimize their photoacoustic effect in the field of bioimaging [21–23]. However, current studies generally either lack the accurate particle shape analysis or do not provide a statistically representative number of particles in their measurements. This is particularly problematic as many-nanoparticle syntheses can have large fluctuations within the same batch or across batches [24–26]. Therefore, there is an urgent need for an accurate and statistically significant analysis of the size and shape distribution of nanoparticles to aid the control in nanoparticle synthesis, thereby optimizing their desired properties.

Transmission electron microscopy (TEM) offers high spatial resolution imaging and can provide structural information of nanoparticles down to the atomic level [27–29]. The technique uses a beam of medium-energy (~80–300 keV) electrons to image a thin sample. It is well suited for an accurate analysis of nanoparticle shapes and sizes [30]. However, manually analyzing and classifying the shape of particles in TEM images is laborious and challenging, especially when a statistically meaningful number of particles is needed. Some existing image analysis software [31,32] can partially solve this problem by carrying out

particle size distribution (PSD) analysis of nanoparticles. Still, nearly all of the existing software lacks the ability to effectively analyze and classify the shape of the particles.

Machine learning (ML) methods have been shown to accomplish efficient classification of large data sets in many fields, such as text mining and biology [33,34]. Currently, deep learning methods such as Convolutional Neural Network (CNN) algorithms [35,36] and Artificial Neural Network (ANN) have been used for particle shape classification [37,38]. Although this approach achieves good results for specific particle shapes, it is dependent on *a priori* knowledge of the latter to generate training datasets. In order to achieve generalization in data analysis, an unsupervised ML algorithm, which does not require prior input and data training, is more desirable. Recently published studies are showing that unsupervised ML methods can efficiently analyze the size and shape information of the nanoparticles; however, they are limited to nanoparticles with well distinguished shapes with strong image contrast (such as gold nano-rods) and good dispersions (isolated particles) [39,40]. In reality, it is difficult to guarantee a clear contrast and a homogeneous background in EM images due to a wide range of nanoparticle compositions, dispersions and imaging conditions.

Here, we demonstrate a robust and efficient unsupervised ML method for nanoparticle shape classification and morphological distribution analysis. The method first uses computer vision algorithms to pre-process images and obtain particle morphology information, followed by a hierarchical-clustering based unsupervised ML algorithm to classify particle shapes and ultimately output statistical particle morphology analysis results. This method is highly automated and it is applicable to convex nanoparticle shapes (this limitation will become clear in Section 2.4), dispersion density and particle composition (which determines the image contrast). Our method is demonstrated using three carefully chosen examples that cover nanoparticles with different shapes, particle dispersions, and TEM imaging modes. By using our presented method, we can overcome the shortcomings of the conventional manual TEM data analysis methods, which analyze an insufficient number of particles, giving poor statistics. Moreover, this method also fills a gap in the particle shape classification in the common large-number particle analysis software.

2. Methodology

The schematic diagram of our ML method workflow is shown in Figure 1. The method is divided into two parts: The first part is pre-processing, which converts the raw image data into nanoparticle contour datasets that are then applied to the ML algorithm. The second part is the application of the ML algorithm to classify the shapes and the evaluation of the ML output to obtain the relevant information, in this case, the shape distribution.

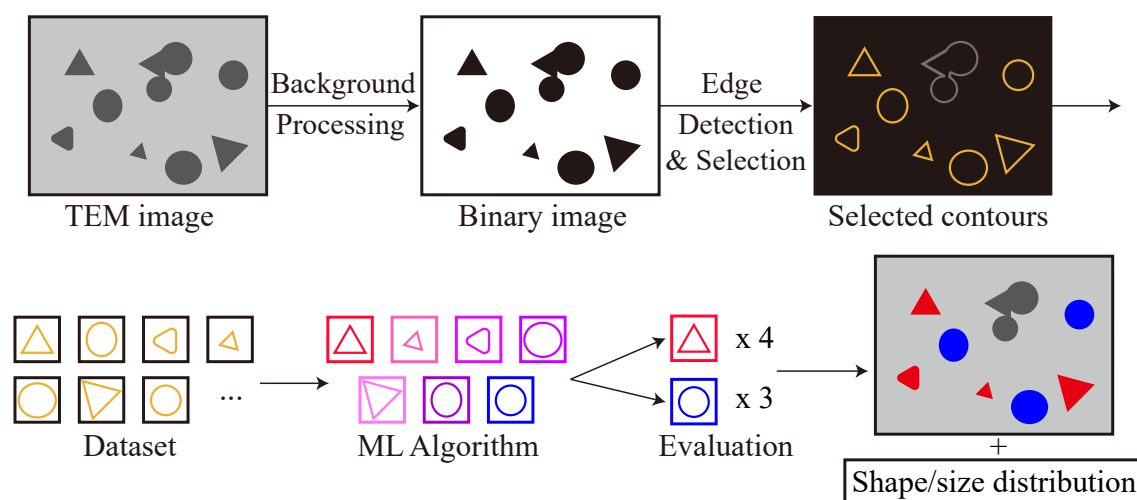


Figure 1. Schematic workflow of the methodology. The upper panel shows the imaging pre-processing steps and the lower panel illustrates the particle shape classification and evaluation via the ML algorithm.

2.1. Sample Preparation and TEM Imaging

The TEM samples of nanoparticles were prepared by drop-casting nanoparticle dispersions onto the holey carbon film coated Cu TEM grids.

The nanoparticle TEM samples were imaged using JEOL F200 (Akishima, Tokyo, Japan), operated at 200 kV, using either the bright-field TEM (BF-TEM) mode (where the objective aperture was placed around the transmitted beam to form an image) or the annular-dark field scanning-TEM (ADF-STEM) mode (where a small probe formed by STEM mode was raster through the specimen and the scattered electrons were collected by the ADF detector to form an image). Note that these modalities in the TEM generate images of the sample as seen in projection, and hence the shapes of nanoparticles extracted using the steps described below are two-dimensional projections.

2.2. Pre-Processing: Background Removal

The first step of pre-processing is to separate the particles from the background in the image. One common way to achieve this is to take advantage of the difference in brightness between the particles and the background in the TEM image by setting a global brightness threshold to filter out the background in the image [41]. However, one challenge in relying on this approach to obtain reliable background-removed images is the inhomogeneity of the background. Such inhomogeneity can arise from variations in the thickness, or surface unevenness of the sample background matrix, or from non-uniform electron microscope illumination. The inhomogeneous background can lead to the regional background being over- and under-filtered in their removal process when a global filtering threshold is applied. As a result, the process either does not properly isolate the particles or it introduces artifacts, which cause a misrepresentation of the shapes. Such inadequate background removal can interfere with the accuracy of particle shape information.

To overcome this, we adopt a more sophisticated approach of dynamic thresholding. This method measures the local background intensities within optimal sub-regions in the TEM image, then adjusts the background in different small sub-regions of the image to be consistent with the optimal before the final background removal [42]. As an example shown in Figure 2a, due to the uneven background intensity in the image, with a brighter background in the upper right region and a darker background in the lower-left region, the darker region indicated in red dashed lines is under-filtered when a global threshold is applied (Figure 2b).

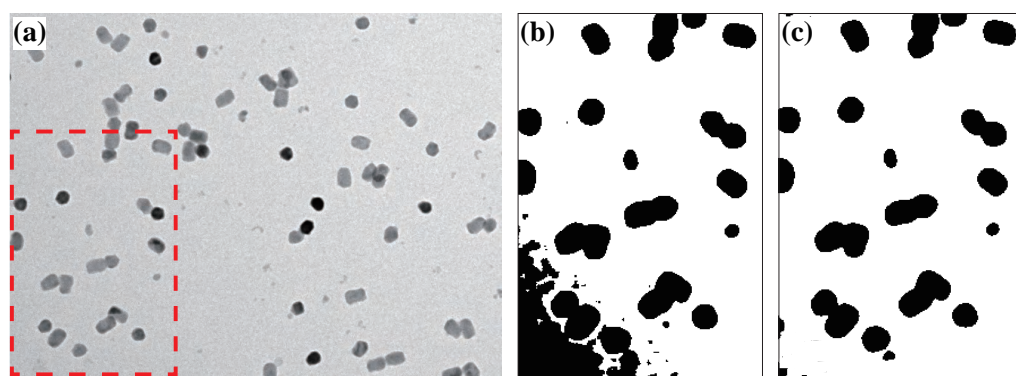


Figure 2. Background pre-processing and removal. (a) Raw TEM image with uneven background. (b) After global thresholding. The area shown here is highlighted in red dashed line in (a). (c) After dynamic thresholding of the same area in (b).

After using the dynamic thresholding method, the particles that are buried by the under-filtered background are now clearly distinguished (Figure 2c). We will show that this method is effective and robust in Section 3.

2.3. Pre-Processing: Particle Edge Identification

After removing the background and converting the images into binary form, the particle's edges (shapes) are identified by applying the Canny Edge Detection algorithm [43]. This algorithm is based on the first-order derivative of the noise-reduced, smoothed image in the X, Y directions, and the edges of the particles are located at the pixels with maximum values of the first-order derivative along the gradient directions. This is a widely used method due to its simplicity and robustness [44,45].

2.4. Pre-Processing: Overlapping Particles Filtering

One limitation of the Canny Edge method is that it cannot correctly identify the particle edges when they are aggregated or overlapped.

Numerous methods, such as watershed segmentation [39], and dilate and flood filling [46], have been developed to identify individual particles from aggregates. These methods work well when the particles are just "touching" but they fail when particles overlap each other. Recently, deep learning methods have been shown to achieve accurate identification and segmentation of highly overlapped particles through data training [47]. However, again, these deep learning methods are not generally applicable without *a priori* knowledge.

Here we use a straight-forward convexity filtering method to simply filter out the aggregated/overlapping particles. This approach assumes that the nanoparticles that are convex and aggregated tend to have concave shapes. By identifying the convexity of the particle contours, aggregates can be detected [48] and then removed from the dataset. This approach ensures meaningful classification and, at the same time, can still retain statistically significant numbers of nanoparticles being analyzed by using mode images. As the convexity filtering places the assumption that the isolated particles are convex shaped, this simple filtering method therefore sets a limitation to our method. Considering concave shaped nanoparticles are relatively uncommon [49], this convexity filtering method can still be applied to most types of nanoparticles. While the limitation of the method is set by the aggregate filtering here, the particle shape classification method (see next section) is not limited to convex shaped particles.

2.5. Classification

Having converted the TEM image data into the particle contour dataset, we then apply an unsupervised ML classification algorithm for the particle shape classification. Unsupervised ML algorithms, unlike *supervised* ML methods, do not require prior classified information for data training, thus allowing our method to be directly applied to a wide range of nanoparticles. More discussions regarding the differences between the supervised and unsupervised ML methods can be found in References [50,51].

In the classification process, we firstly parameterized the shape of each particle contour using Hu moments, which are a set of numbers that depend only on the particle shape and are independent of the particle size, position and orientation [52]. Then we applied the hierarchical agglomerative clustering method with the average linkage [53] to classify the parameterized particle shapes. Briefly speaking, a hierarchical clustering algorithm groups together data points (in our case, the particle contours) with similarities between them. This is achieved by using a measurement metric (in our case, the "distance" between the pair of data points) and a linkage criteria, which specifies the similarity of data sets as a function of pairwise distances of observations in the sets (more detail can be found in the Supplementary Information) [54]. This clustering method builds a hierarchy of clusters and therefore does not require a pre-determined optimum numbers of clusters. Such method has been previously applied in other fields [55–57]. Such method is different from the commonly used k-means method [58] where users need to decide the number of clusters in advance. Thus, the hierarchical clustering method allows full automation. Finally, the optimum numbers of clusters are determined automatically by applying internal Cluster Validity Indexes (CVIs). The three internal CVIs we choose are Silhouette [59], Davies–

Bouldin [60], and Calinski–Harabaz [61], and the optimum numbers of clusters are the common local extrema of each CVI. A detailed description of the internal CVIs can be found in the Supplementary Information.

We emphasize here that our method is an unsupervised machine learning algorithm (as distinct from, e.g., “deep learning” methods), which does not require initial simulations and is therefore generally applicable to any nanoparticle samples without the need of “training data.” In addition, with some initial setup, the method can automatically analyze sets of TEM images containing hundreds of nanoparticles within a few seconds using a personal laptop, which demonstrates its efficiency.

3. Results and Discussions

Here, we demonstrate the effectiveness of our method using three examples, which cover different particle shapes, dispersion, and TEM imaging modalities.

3.1. Well Dispersed, Multiple Shape Upconversion Nanoparticles

Upconversion nanoparticles (UCNPs) are a unique class of optical nanomaterials that can up-convert two or more low energy photons into one higher energy photon [62]. They have been used in a broad range of applications such as bio-imaging [63], photodynamic theranostics [64,65], and anti-counterfeiting [66]. Lanthanide ion (Ln) doped UCNPs have been demonstrated as one of the most efficient upconversion fluorescent nanoparticles due to their low phonon vibration energy [67,68]. The size and shape of UCNPs play an important role in modifying the optical and electromagnetic properties [69].

Figure 3a shows bright-field (BF)-TEM images of NaGdF₄: 49%Yb, 1%Tm nanoparticles. This image represents a typical situation of a relatively low magnification TEM image showing well dispersed nanoparticles. This type of image is commonly used to evaluate the particle synthesis process, as lower magnification images have the advantage of capturing more particles in the field of view, thus providing better statistical accuracy for a given amount of experimental data and time. The trade-off is that lower magnification images ultimately imply fewer pixels for describing a given particle contour. Therefore, the deliberately chosen low magnification image is served to test the robustness of our method in capturing the particle shapes.

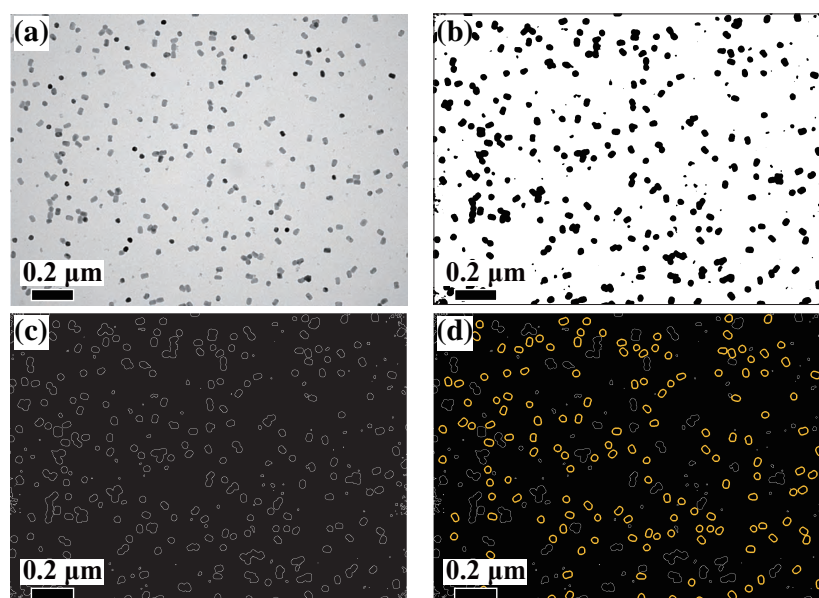


Figure 3. Image pre-processing of the UCNPs TEM image. (a) A low-mag, BF-TEM image for well dispersed UCNPs sample. (b) Binary image after the dynamic background processing. (c) Particle edges detected by the Canny Edge detection method. (d) Convexity filtered particle contours (yellow), where all the non-convex particle aggregates (grey) are filtered out.

Figure 3b–d shows the successful processing by our pre-processing method, where 156 isolated particle shape contours were extracted. The optimal number of clusters, k , is determined to be two by the CVIs. The classified particles overlaid with the original TEM image are shown in Figure 4a. It can be seen that one particle shape cluster, k_0 (blue), is polygonal (hexagons and pentagons) and the other cluster, k_1 (green), is rod-shaped. The particle size distribution and the aspect ratio distribution for each cluster are plotted in Figure 4b,c, respectively. The two shape clusters are clearly separated in the aspect ratio distribution plot, with the polygon cluster, k_0 , having a mean value close to 1 and a higher value of about 1.4 for k_1 . In addition, the particles in the k_0 cluster mostly have a smaller size with a mean diameter around 34 nm whereas those in the k_1 cluster have a slightly larger mean diameter around 38 nm. Table 1 provides a summary of the statistics for the UNCP size and shape analysis.

These results demonstrate the accuracy of our method in particle contour determination and shape classification, even at relatively low image magnification.

Table 1. Clustering summary for UCNP shape and size analysis. k_n is the cluster number, Σ and rate are the number and the fraction of particles for each cluster. Size and aspect ratio represents the mean size (diameter) and the mean aspect ratio of the particles for each cluster.

k_n	Σ	Rate	Size (nm)	Aspect Ratio
k_0	75	48.1%	33.6 ± 2.0	1.07 ± 0.06
k_1	81	51.9%	38.1 ± 1.7	1.37 ± 0.08

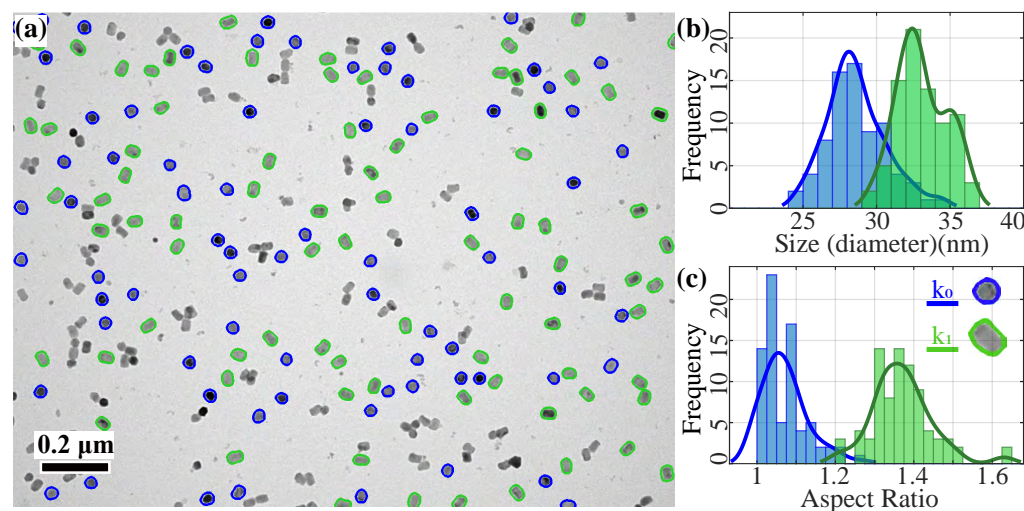


Figure 4. Shape classification results for the TEM image of UCNPs. (a) The two different shape classification particle contours are marked with two colors. (b) Size distribution (approximate diameter assuming the particles are spherical) of the two clusters, k_0 and k_1 . (c) Aspect ratio distribution of the two clusters.

In this case, we also implemented manual particle shape analysis as a validation of our method. For the same selected convex particles, the manual analysis results are 76 (polygonal) vs. 80 (rod-shaped), with an agreement rate of 98.7%.

3.2. High Packing Density Semiconductor Quantum Dots

Semiconductor quantum dots (QDs) are nanocrystals typically in the size ranges of 1–10 nm, exhibiting quantum confinement effects in their optical and electronic properties [70]. Such properties give semiconductor QDs great promise in applications such as bio-medicine [71], and light-emitting diodes [72]. The CdS/ZnSe used in this case are typical core-shell structure semiconductor QDs, with a shell layer that gives the particles better photoluminescence quantum yields and optical stability compared to bare cores [73].

It has been shown that by controlling the size of such QDs [70,74], modulation of their band gap and optical properties can be achieved. In addition, the shape of the QD particles reflects the formation of its shell layer [26]. This highlights the importance of precise control over their size and shape in their synthesis. Here we intentionally selected a batch of the CdS/ZnSe sample particles whose size was controlled but with non-distinct particle shapes.

As seen from the original BF-TEM image in Figure 5a, the QDs have a very high packing density, filling nearly the entire image where the background only occupies a very small proportion of the image. In addition, the QD contrast is relatively low compared to the background. Such complex background condition presents a challenge for background removal. Moreover, most of the particles have non-distinct shapes, making manual classification of the shape nearly impossible. Therefore, this case is deliberately selected and served to test the reliability of our method for images with low contrast, low background area and irregular particle shapes.

Figure 5b–d demonstrates the effectiveness of our pre-processing method even when little background area is present. A total of 482 isolated QDs in the image (Figure 5d, yellow) were successfully extracted for classification.

Our method determines the optimal number of clusters, $k_n = 5$. The classified particles overlaid with the original image are shown in Figure 5e. The corresponding size distribution (Figure 5f,g for high and low population clusters, respectively) and eccentricity distributions (Figure 5h,i for high and low population clusters, respectively) are plotted. Table 2 summarizes the detailed QD size and shape analysis. We can see that, as expected, all shape clusters have very similar average sizes, ranging from 11 to 12 nm.

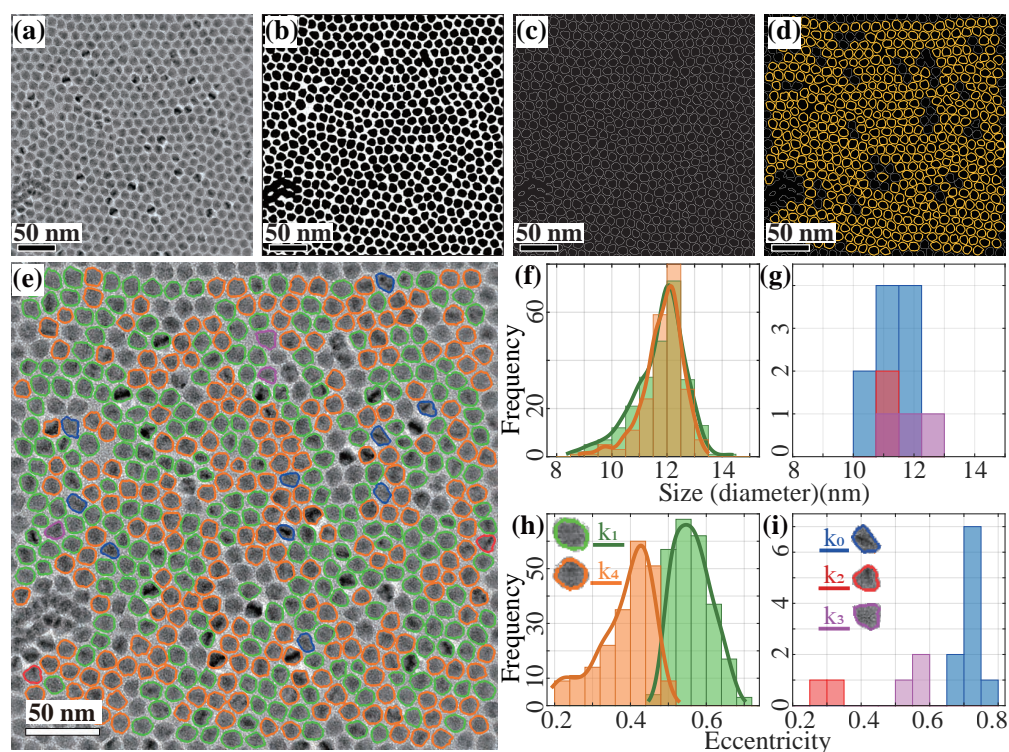


Figure 5. Image pre-processing and shape classification of the QDs. (a) BF-TEM image of the CdS/ZnSe QDs sample. (b) Binary image after the dynamic background processing. (c) Particle edges detected by the Canny Edge detection method. (d) Convexity filtered particle contours (yellow), where all the non-convex particle aggregates (grey) are filtered out. (e) Classified particle contours labeled with different colors overlaid with the BF-TEM image shown in (a). (f) Size distribution of the two highest population clusters. (g) Size distribution of the three lowest population clusters. (h) Eccentricity distribution of the high population clusters of k_1 and k_4 . (i) Eccentricity distribution of the low population clusters.

Table 2. Clustering summary for the QD size and shape analysis. k_n is the cluster number, Σ and rate are the number and the fraction of particles for each cluster. Size, aspect ratio and eccentricity represent the mean size (diameter), the mean aspect ratio and the mean eccentricity of the particles for each cluster.

k_n	Σ	Rate	Size (nm)	Aspect Ratio	Eccentricity
k_0	10	2.1%	11.2 ± 0.6	1.34 ± 0.11	0.71 ± 0.02
k_1	248	51.5%	11.8 ± 0.9	1.14 ± 0.08	0.56 ± 0.05
k_2	2	0.4%	11.5 ± 0.0	1.05 ± 0.03	0.29 ± 0.04
k_3	3	0.6%	11.7 ± 0.7	1.20 ± 0.11	0.57 ± 0.03
k_4	219	45.4%	11.9 ± 0.7	1.06 ± 0.04	0.38 ± 0.08

The two shape clusters, k_1 (green) and k_4 (orange), with the highest population fraction (51.5% and 45.4%) have almost identical size distributions. Their aspect ratio distributions (see Supplementary Information Figure S2) also have a large degree of overlap with the k_4 (orange) shape cluster, which has a narrower aspect ratio distribution, close to 1, whereas the k_1 (green) shape cluster has a more dispersed aspect ratio distribution ranging from 1 to 1.4, peaking at 1.15. The apparent differences between these two shape clusters are in their eccentricity distributions (Figure 5h) with the former peaked at 0.38 and the latter at 0.56.

The other three lower population shape clusters also exhibit their own characteristics, such as the significantly larger aspect ratio of cluster k_0 (blue), the triangular shape of cluster k_2 (red) with narrow size distribution, and the fan shape of cluster k_3 (purple). Their eccentricity distributions (Figure 5i) also exhibit clear differences from each other.

Such subtle differences in particle shapes would have been difficult to distinguish by human manual classification, but our ML algorithm demonstrates its effectiveness in this case.

3.3. Iron Nanocubes from ADF-STEM Images

Iron nanocubes combine advantages of small particle size and large magnetization, which make them an ideal tool for bio-imaging, especially magnetic particle imaging [75,76]. In the synthesis of such magnetic nanoparticles, both their shape and size can significantly affect their magnetic properties [49,77]. Therefore, the accurate identification and statistical analysis of its morphological information is crucial to realize its potential for bio-medical and thermal therapy applications. In this case study, the shapes the Fe-core/Fe₂O₃-shell nanoparticles are engineered to be cube-shaped to optimize their magnetic properties for biomedical imaging applications.

In this section, we use images acquired in ADF-STEM mode to demonstrate the versatility of our method in different TEM modalities. Due to the high sensitivity of the ADF-STEM mode in terms of the mass–thickness contrast, as well as the high spatial resolution it can achieve, it is often used for nanoparticles with higher atomic number compositions, as well as for high-resolution particle structure imaging. Compared to BF-TEM, ADF-STEM images generally have weaker background noise, which reduces the difficulty of background pre-processing. In addition, the contrast of ADF images provide local composition and/or particle thickness information. However, ADF-STEM imaging mode has some limitations, as it often takes longer image acquisition time than TEM mode and generally has a smaller field of view (FoV).

The ADF-STEM image of the iron nanocubes (Figure 6a) shows that the background is more homogeneous and particle-background contrast is sharper (compared to BF-TEM images). The iron nanocubes in the image also have a high packing density, with some of the particles touching each other forming aggregates. Although the iron nanocubes generally reflect their cubic shape, it is clear that there are variations in the shapes of the particles, which are also difficult to classify manually.

Again, our image pre-processing method works well for the ADF-STEM images, with 110 contours of iron nanocubes extracted (Figure 6d). Our method automatically classifies

the shape of the particles into three clusters, where k_0 (blue) represents elongated shapes, k_1 (green) regular cube shapes and k_2 (red) slightly distorted cube shapes, as shown in Figure 6e. Although the shape clusters do not reflect a clear distinction from their particle size distributions (Figure 6f), and the difference in their aspect ratio distributions is subtle (see Supplementary Information Figure S3), their apparent shape differences can be found in their eccentricity (Figure 6g and Table 3), where the eccentricities of the three shape clusters are concentrated at 0.4, 0.6 and 0.7, respectively.

Particles in the two highest population shape clusters accounted for over 96% of all particles, indicating that overall, the particles were close to cubic-shaped in the sample synthesis. On the other hand, the lowest population k_0 (blue) cluster has an aspect ratio around 1.3 (Table 3), a larger eccentricity (Figure 6g) and larger particle size (Figure 6f), making it easily distinguishable from the other two clusters. This cluster represents the few particles that did not form a cubic shape in the synthesis.

Table 3. Clustering summary for the iron nanocubes size and shape analysis. k_n is the cluster number, Σ and rate are the number and the fraction of particles for each cluster. Size, aspect ratio and eccentricity represent the mean size (diameter), the mean aspect ratio and the mean eccentricity of the particles for each cluster.

k_n	Σ	Rate	Size (nm)	Aspect Ratio	Eccentricity
k_0	4	3.6%	12.3 ± 1.6	1.33 ± 0.07	0.70 ± 0.02
k_1	81	73.6%	10.4 ± 1.1	1.06 ± 0.04	0.39 ± 0.09
k_2	25	22.7%	10.1 ± 1.3	1.15 ± 0.08	0.57 ± 0.03

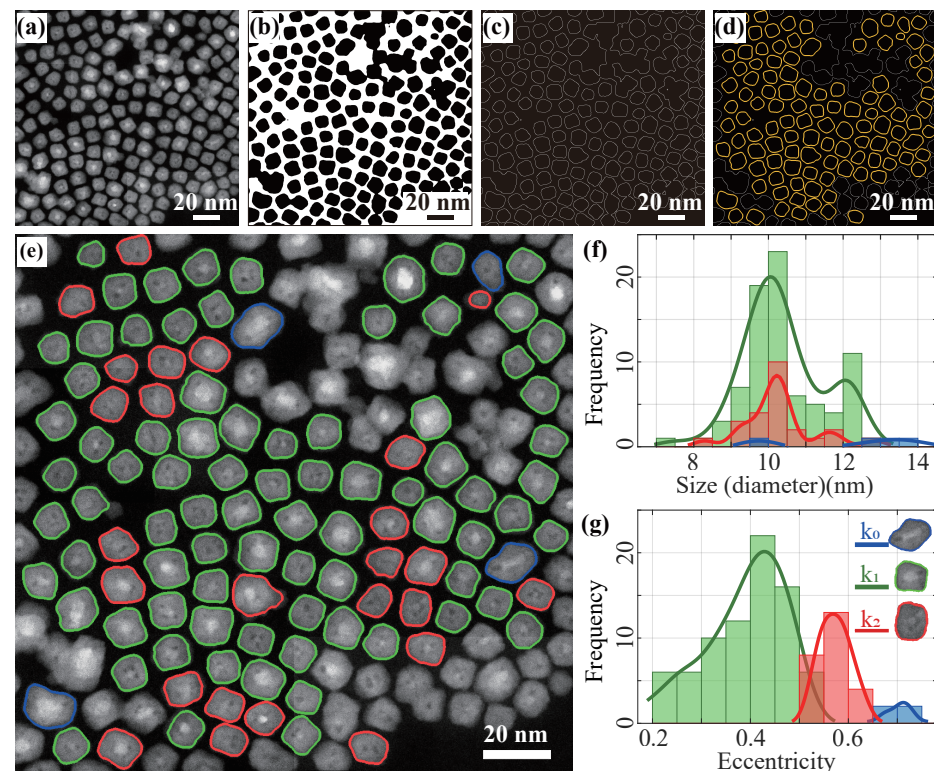


Figure 6. Image pre-processing of the iron cubes ADF-STEM image. (a) A high magnification, closely packed ADF-STEM image for the core-shell iron nanocubes sample. (b) Binary image after the dynamic background processing. (c) Particle edges detected by the Canny Edge detection method. (d) Non-convex particle contours (yellow) data collected as a dataset. (e) The three different shape classification particle contours are marked with three colors. (f) Size distribution (approximate diameter assuming the particles are spherical) of the three clusters. (g) Eccentricity distribution of the three clusters.

4. Conclusions

In this work, we have demonstrated an efficient, robust and highly automated, unsupervised ML method for analyzing the nanoparticle size and shape distributions from TEM images. Our method offers several advantages over the existing methods: Firstly, it utilizes the dynamic thresholding method that can effectively remove inhomogeneous background even with low nanoparticles contrast. Secondly, it uses a hierarchical clustering method for particle shape classification, which does not require any potentially biased, user pre-determined numbers of particle shape clusters. Lastly, the use of the Cluster Validation Indexes (CVIs) contributes to determining the optimum number of particle shape clusters and offers a fully-automated analysis method without any a priori input.

The effectiveness of our method is demonstrated on three distinct examples that cover the scenarios of different particle shapes, image magnification, image background qualities, and imaging modalities. Our results show that the unsupervised ML method that we have developed provides highly automated and accurate analysis of size and shape distribution over large numbers of nanoparticles. Moreover, even for irregularly shaped particles, our method is sufficiently accurate and robust to extract and correctly classify their different shape features.

The ability to statistically analyze the nanoparticle key performance parameters of the size and shape from TEM images will certainly have a significant contribution in advancing the control of nanoparticle synthesis, thereby giving better control of particle performance. Moreover, while our methodology was only demonstrated on TEM images, it should be directly applicable to other types of microscopy images and therefore provide a broader impact to other fields of microscopy.

Supplementary Materials: The following are available at <https://www.mdpi.com/article/10.3390/nano11102706/s1>, Figure S1: Divisive and agglomerative Hierarchical clustering representation, Figure S2: Shape classification result of High packing density QDs, Figure S3: Shape classification result of Iron nanocubes ADF-STEM images.

Author Contributions: Conceptualization, S.L.Y.C., J.M.L.-R. and H.W.; methodology, H.W. and J.M.L.-R.; software, H.W. and J.M.L.-R.; validation, H.W.; formal analysis, H.W.; investigation, H.W.; resources, S.L.Y.C. and J.C.R.; data curation, H.W.; writing—original draft preparation, H.W.; writing—review and editing, S.L.Y.C., C.D. and H.W.; visualization, H.W.; supervision, S.L.Y.C., C.D. and J.C.R.; project administration, S.L.Y.C.; funding acquisition, S.L.Y.C. All authors have read and agreed to the published version of the manuscript.

Funding: Funding support from Australia Research Council (ARC) under Grant number IC210100056 and the Spanish Ministry of Economy and Competitiveness under Projects TIN2014-55894-C2-R and TIN2017-88209-C2-2-R.

Data Availability Statement: The data are available from the corresponding author upon reasonable request.

Acknowledgments: H.W. acknowledges the support of TFS scholarship by UNSW. All authors thank Xiaoxue Xu, Soshan Cheong, Shen-Chuan Lo and Jung-Hsuan Chen for providing TEM images for the UCNP, iron nanocubes and QDs. This work used facilities supported by Microscopy Australia at the Electron Microscope Unit at UNSW.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wang, F.; Tan, W.B.; Zhang, Y.; Fan, X.; Wang, M. Luminescent nanomaterials for biological labelling. *Nanotechnology* **2005**, *17*, R1. [[CrossRef](#)]
2. Rosenblum, L.T.; Kosaka, N.; Mitsunaga, M.; Choyke, P.L.; Kobayashi, H. In vivo molecular imaging using nanomaterials: general in vivo characteristics of nano-sized reagents and applications for cancer diagnosis. *Mol. Membr. Biol.* **2010**, *27*, 274–285. [[CrossRef](#)]
3. Chen, Y.S.; Frey, W.; Kim, S.; Kruijzinga, P.; Homan, K.; Emelianov, S. Silica-coated gold nanorods as photoacoustic signal nanoamplifiers. *Nano Lett.* **2011**, *11*, 348–354. [[CrossRef](#)]

4. Hu, J.; Lawrence, M.; Dionne, J.A. High quality factor dielectric metasurfaces for ultraviolet circular dichroism spectroscopy. *ACS Photonics* **2019**, *7*, 36–42. [[CrossRef](#)]
5. Rocha, U.; Jacinto da Silva, C.; Ferreira Silva, W.; Guedes, I.; Benayas, A.; Martinez Maestro, L.; Acosta Elias, M.; Bovero, E.; van Veggel, F.C.; Garcia Sole, J.A.; et al. Subtissue thermal sensing based on neodymium-doped LaF₃ nanoparticles. *ACS Nano* **2013**, *7*, 1188–1199. [[CrossRef](#)]
6. Kalytchuk, S.; Poláková, K.; Wang, Y.; Froning, J.P.; Cepe, K.; Rogach, A.L.; Zbořil, R. Carbon dot nanothermometry: Intracellular photoluminescence lifetime thermal sensing. *ACS Nano* **2017**, *11*, 1432–1442. [[CrossRef](#)]
7. Gandolfi, M.; Crut, A.; Medeghini, F.; Stoll, T.; Maioli, P.; Vallée, F.; Banfi, F.; Del Fatti, N. Ultrafast thermo-optical dynamics of plasmonic nanoparticles. *J. Phys. Chem. C* **2018**, *122*, 8655–8666. [[CrossRef](#)]
8. Sun, X.; Zhang, Y.; Chen, G.; Gai, Z. Application of nanoparticles in enhanced oil recovery: A critical review of recent progress. *Energies* **2017**, *10*, 345. [[CrossRef](#)]
9. Savage, N.; Diallo, M.S. Nanomaterials and water purification: Opportunities and challenges. *J. Nanoparticle Res.* **2005**, *7*, 331–342. [[CrossRef](#)]
10. Chen, W.T.; Zhu, A.Y.; Sanjeev, V.; Khorasaninejad, M.; Shi, Z.; Lee, E.; Capasso, F. A broadband achromatic metalens for focusing and imaging in the visible. *Nat. Nanotechnol.* **2018**, *13*, 220–226. [[CrossRef](#)] [[PubMed](#)]
11. Celebrano, M.; Rocco, D.; Gandolfi, M.; Zilli, A.; Rusconi, F.; Tognazzi, A.; Mazzanti, A.; Ghirardini, L.; Pogna, E.A.; Carletti, L.; et al. Optical tuning of dielectric nanoantennas for thermo-optically reconfigurable nonlinear metasurfaces. *Opt. Lett.* **2021**, *46*, 2453–2456. [[CrossRef](#)]
12. Liu, S.; Vabishchevich, P.P.; Vaskin, A.; Reno, J.L.; Keeler, G.A.; Sinclair, M.B.; Staude, I.; Brener, I. An all-dielectric metasurface as a broadband optical frequency mixer. *Nat. Commun.* **2018**, *9*, 1–6. [[CrossRef](#)]
13. Ronchi, A.; Sterzi, A.; Gandolfi, M.; Belarouci, A.; Giannetti, C.; Del Fatti, N.; Banfi, F.; Ferrini, G. Discrimination of nano-objects via cluster analysis techniques applied to time-resolved thermo-acoustic microscopy. *Ultrasonics* **2021**, *114*, 106403. [[CrossRef](#)]
14. Chuang, P.H.; Lin, C.C.; Liu, R.S. Emission-tunable CuInS₂/ZnS quantum dots: Structure, optical properties, and application in white light-emitting diodes with high color rendering index. *ACS Appl. Mater. Interfaces* **2014**, *6*, 15379–15387. [[CrossRef](#)]
15. Nayak, L.; Mohanty, S.; Nayak, S.K.; Ramadoss, A. A review on inkjet printing of nanoparticle inks for flexible electronics. *J. Mater. Chem. C* **2019**, *7*, 8771–8795. [[CrossRef](#)]
16. Xie, C.; Niu, Z.; Kim, D.; Li, M.; Yang, P. Surface and interface control in nanoparticle catalysis. *Chem. Rev.* **2019**, *120*, 1184–1249. [[CrossRef](#)] [[PubMed](#)]
17. Köhler, J.; Abahmane, L.; Wagner, J.; Albert, J.; Mayer, G. Preparation of metal nanoparticles with varied composition for catalytical applications in microreactors. *Chem. Eng. Sci.* **2008**, *63*, 5048–5055. [[CrossRef](#)]
18. Gaumet, M.; Vargas, A.; Gurny, R.; Delie, F. Nanoparticles for drug delivery: The need for precision in reporting particle size parameters. *Eur. J. Pharm. Biopharm.* **2008**, *69*, 1–9. [[CrossRef](#)] [[PubMed](#)]
19. Banerjee, A.; Chattopadhyay, K. Size-dependent optical properties of sputter-deposited nanocrystalline p-type transparent Cu Al O₂ thin films. *J. Appl. Phys.* **2005**, *97*, 084308. [[CrossRef](#)]
20. Huang, M.H. Facet-Dependent Optical Properties of Semiconductor Nanocrystals. *Small* **2019**, *15*, 1804726. [[CrossRef](#)]
21. Pang, G.A.; Laufer, J.; Niessner, R.; Haisch, C. Photoacoustic signal generation in gold nanospheres in aqueous solution: Signal generation enhancement and particle diameter effects. *J. Phys. Chem. C* **2016**, *120*, 27646–27656. [[CrossRef](#)]
22. Gandolfi, M.; Banfi, F.; Glorieux, C. Optical wavelength dependence of photoacoustic signal of gold nanofluid. *Photoacoustics* **2020**, *20*, 100199. [[CrossRef](#)] [[PubMed](#)]
23. Pang, G.A.; Poisson, F.; Laufer, J.; Haisch, C.; Bossy, E. Theoretical and experimental study of photoacoustic excitation of silica-coated gold nanospheres in water. *J. Phys. Chem. C* **2019**, *124*, 1088–1098. [[CrossRef](#)]
24. Reineck, P.; Trindade, L.F.; Havlik, J.; Stursa, J.; Heffernan, A.; Elbourne, A.; Orth, A.; Capelli, M.; Cigler, P.; Simpson, D.A.; et al. Not all fluorescent nanodiamonds are created equal: A comparative study. *Part. Part. Syst. Character.* **2019**, *36*, 1900009. [[CrossRef](#)]
25. Cui, J.; Beyler, A.P.; Marshall, L.F.; Chen, O.; Harris, D.K.; Wanger, D.D.; Brokmann, X.; Bawendi, M.G. Direct probe of spectral inhomogeneity reveals synthetic tunability of single-nanocrystal spectral linewidths. *Nat. Chem.* **2013**, *5*, 602–606. [[CrossRef](#)]
26. Lavorato, G.; Alzamora, M.; Contreras, C.; Burlandy, G.; Litterst, F.J.; Baggio-Saitovitch, E. Internal structure and magnetic properties in cobalt ferrite nanoparticles: Influence of the synthesis method. *Part. Part. Syst. Character.* **2019**, *36*, 1900061. [[CrossRef](#)]
27. Gontard, L.C.; Chang, L.Y.; Hetherington, C.J.; Kirkland, A.I.; Ozkaya, D.; Dunin-Borkowski, R.E. Aberration-corrected imaging of active sites on industrial catalyst nanoparticles. *Angew. Chem.* **2007**, *119*, 3757–3759. [[CrossRef](#)]
28. Shery, L. Size dependent surface reconstruction in detonation nanodiamonds. *Nanoscale Horizons* **2018**, *3*, 213–217.
29. Chang, L.Y.; Barnard, A.S.; Gontard, L.C.; Dunin-Borkowski, R.E. Resolving the structure of active sites on platinum catalytic nanoparticles. *Nano Lett.* **2010**, *10*, 3073–3076. [[CrossRef](#)]
30. Evans, J.E.; Jungjohann, K.L.; Browning, N.D.; Arslan, I. Controlled growth of nanoparticles from solution with in situ liquid transmission electron microscopy. *Nano Lett.* **2011**, *11*, 2809–2813. [[CrossRef](#)] [[PubMed](#)]
31. Abramoff, M.D.; Magalhães, P.J.; Ram, S.J. Image processing with ImageJ. *Biophotonics Int.* **2004**, *11*, 36–42.
32. Woehrle, G.H.; Hutchison, J.E.; Özkar, S.; Finke, R.G. Analysis of nanoparticle transmission electron microscopy data using a public-domain image-processing program, image. *Turk. J. Chem.* **2006**, *30*, 1–13.
33. Tarca, A.L.; Carey, V.J.; Chen, X.W.; Romero, R.; Drăghici, S. Machine learning and its applications to biology. *PLoS Comput. Biol.* **2007**, *3*, e116. [[CrossRef](#)] [[PubMed](#)]

34. Khan, A.; Baharudin, B.; Lee, L.H.; Khan, K. A review of machine learning algorithms for text-documents classification. *J. Adv. Inf. Technol.* **2010**, *1*, 4–20.
35. Cid-Mejías, A.; Alonso-Calvo, R.; Gavilán, H.; Crespo, J.; Maojo, V. A deep learning approach using synthetic images for segmenting and estimating 3D orientation of nanoparticles in EM images. *Comput. Methods Programs Biomed.* **2021**, *202*, 105958. [[CrossRef](#)]
36. Asari, Y.; Terada, S.; Tanigaki, T.; Takahashi, Y.; Shinada, H.; Nakajima, H.; Kanie, K.; Murakami, Y. Deep convolutional neural network image processing method providing improved signal-to-noise ratios in electron holography. *Microscopy* **2021**. [[CrossRef](#)] [[PubMed](#)]
37. Hughes, A.; Liu, Z.; Raftari, M.; Reeves, M.E. *A Workflow for Characterizing Nanoparticle Monolayers for Biosensors: Machine Learning on Real and Artificial SEM Images*; Technical Report; PeerJ PrePrints: London, UK, 2014.
38. Pellegrino, F.; Isopescu, R.; Pellutiè, L.; Sordello, F.; Rossi, A.M.; Ortel, E.; Martra, G.; Hodoroaba, V.D.; Maurino, V. Machine learning approach for elucidating and predicting the role of synthesis parameters on the shape and size of TiO₂ nanoparticles. *Sci. Rep.* **2020**, *10*, 1–11. [[CrossRef](#)] [[PubMed](#)]
39. Lee, B.; Yoon, S.; Lee, J.W.; Kim, Y.; Chang, J.; Yun, J.; Ro, J.C.; Lee, J.S.; Lee, J.H. Statistical Characterization of the Morphologies of Nanoparticles through Machine Learning Based Electron Microscopy Image Analysis. *ACS Nano* **2020**, *14*, 17125–17133. [[CrossRef](#)]
40. Wang, X.; Li, J.; Ha, H.D.; Dahl, J.C.; Ondry, J.C.; Moreno-Hernandez, I.; Head-Gordon, T.; Alivisatos, A.P. AutoDetect-mNP: an unsupervised machine learning algorithm for automated analysis of transmission electron microscope images of metal nanoparticles. *JACS AU* **2021**, *1*, 316–327. [[CrossRef](#)]
41. Kapur, J.N.; Sahoo, P.K.; Wong, A.K. A new method for gray-level picture thresholding using the entropy of the histogram. *Comput. Vision, Graph. Image Process.* **1985**, *29*, 273–285. [[CrossRef](#)]
42. Davies, E. Machine Vision: Theory. *Algorithms Pract.* **1990**, *2*, 105–129.
43. Canny, J. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, **1986**, *8*, 679–698. [[CrossRef](#)]
44. Sharifi, M.; Fathy, M.; Mahmoudi, M.T. A classified and comparative study of edge detection algorithms. In Proceedings of the International Conference on Information Technology: Coding and Computing, Las Vegas, NV, USA, 8–10 April 2002; pp. 117–120.
45. Kalbasi, M.; Nikmehr, H. Noise-Robust, Reconfigurable Canny Edge Detection and its Hardware Realization. *IEEE Access* **2020**, *8*, 39934–39945. [[CrossRef](#)]
46. Liang, Z.; Nie, Z.; An, A.; Gong, J.; Wang, X. A particle shape extraction and evaluation method using a deep convolutional neural network and digital image processing. *Powder Technol.* **2019**, *353*, 156–170. [[CrossRef](#)]
47. Masubuchi, S.; Watanabe, E.; Seo, Y.; Okazaki, S.; Sasagawa, T.; Watanabe, K.; Taniguchi, T.; Machida, T. Deep-learning-based image segmentation integrated with optical microscopy for automatically searching for two-dimensional materials. *NPJ 2D Mater. Appl.* **2020**, *4*, 1–9. [[CrossRef](#)]
48. Park, C.; Huang, J.Z.; Ji, J.X.; Ding, Y. Segmentation, inference and classification of partially overlapping nanoparticles. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *35*, 1. [[CrossRef](#)] [[PubMed](#)]
49. Kwizera, E.A.; Chaffin, E.; Shen, X.; Chen, J.; Zou, Q.; Wu, Z.; Gai, Z.; Bhana, S.; O'Connor, R.; Wang, L.; et al. Size- and shape-controlled synthesis and properties of magnetic-plasmonic core-shell nanoparticles. *J. Phys. Chem. C* **2016**, *120*, 10530–10546. [[CrossRef](#)]
50. Murphy, K.P. *Machine Learning: A Probabilistic Perspective*; MIT Press: Cambridge, MA, USA, 2012.
51. Bishop, C. *Pattern Recognition and Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2006.
52. Hu, M.K. Visual pattern recognition by moment invariants. *IRE Trans. Inf. Theory* **1962**, *8*, 179–187.
53. Jain, A.K. Data clustering: 50 years beyond K-means. *Pattern Recognit. Lett.* **2010**, *31*, 651–666. [[CrossRef](#)]
54. Murtagh, F.; Contreras, P. Algorithms for hierarchical clustering: An overview. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2012**, *2*, 86–97. [[CrossRef](#)]
55. Luna-Romera, J.M.; Núñez-Hernández, F.; Martínez-Ballesteros, M.; Riquelme, J.C.; Ibáñez, C.U. Analysis of the evolution of the Spanish labour market through unsupervised learning. *IEEE Access* **2019**, *7*, 121695–121708. [[CrossRef](#)]
56. Albert, W.; Tullis, T. *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics*; Newnes: Boston, MA, USA, 2013.
57. Lee, S.; Crawford, M.M. Unsupervised multistage image classification using hierarchical clustering with a Bayesian similarity measure. *IEEE Trans. Image Process.* **2005**, *14*, 312–320.
58. Kodinariya, T.M.; Makwana, P.R. Review on determining number of Cluster in K-Means Clustering. *Int. J.* **2013**, *1*, 90–95.
59. Rousseeuw, P.J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **1987**, *20*, 53–65. [[CrossRef](#)]
60. Davies, D.L.; Bouldin, D.W. A cluster separation measure. *IEEE Trans. Pattern Anal. Mach. Intell.* **1979**, *1*, 224–227. [[CrossRef](#)]
61. Caliński, T.; Harabasz, J. A dendrite method for cluster analysis. *Commun. Stat. Theory Methods* **1974**, *3*, 1–27. [[CrossRef](#)]
62. Wen, S.; Zhou, J.; Zheng, K.; Bednarkiewicz, A.; Liu, X.; Jin, D. Advances in highly doped upconversion nanoparticles. *Nat. Commun.* **2018**, *9*, 1–12. [[CrossRef](#)]
63. DaCosta, M.V.; Doughan, S.; Han, Y.; Krull, U.J. Lanthanide upconversion nanoparticles and applications in bioassays and bioimaging: A review. *Anal. Chim. Acta* **2014**, *832*, 1–33. [[CrossRef](#)] [[PubMed](#)]

64. Chen, G.; Qiu, H.; Prasad, P.N.; Chen, X. Upconversion nanoparticles: Design, nanochemistry, and applications in theranostics. *Chem. Rev.* **2014**, *114*, 5161–5214. [[CrossRef](#)] [[PubMed](#)]
65. Xu, J.; Xu, L.; Wang, C.; Yang, R.; Zhuang, Q.; Han, X.; Dong, Z.; Zhu, W.; Peng, R.; Liu, Z. Near-infrared-triggered photodynamic therapy with multitasking upconversion nanoparticles in combination with checkpoint blockade for immunotherapy of colorectal cancer. *ACS Nano* **2017**, *11*, 4463–4474. [[CrossRef](#)] [[PubMed](#)]
66. You, M.; Zhong, J.; Hong, Y.; Duan, Z.; Lin, M.; Xu, F. Inkjet printing of upconversion nanoparticles for anti-counterfeit applications. *Nanoscale* **2015**, *7*, 4423–4431. [[CrossRef](#)]
67. Boyer, J.C.; Cuccia, L.A.; Capobianco, J.A. Synthesis of colloidal upconverting NaYF₄: Er³⁺/Yb³⁺ and Tm³⁺/Yb³⁺ monodisperse nanocrystals. *Nano Lett.* **2007**, *7*, 847–852. [[CrossRef](#)] [[PubMed](#)]
68. Lin, M.; Zhao, Y.; Wang, S.; Liu, M.; Duan, Z.; Chen, Y.; Li, F.; Xu, F.; Lu, T. Recent advances in synthesis and surface modification of lanthanide-doped upconversion nanoparticles for biomedical applications. *Biotechnol. Adv.* **2012**, *30*, 1551–1561. [[CrossRef](#)]
69. Wang, G.; Peng, Q.; Li, Y. Lanthanide-doped nanocrystals: Synthesis, optical-magnetic properties, and applications. *Accounts Chem. Res.* **2011**, *44*, 322–332. [[CrossRef](#)] [[PubMed](#)]
70. Vasudevan, D.; Gaddam, R.R.; Trinchì, A.; Cole, I. Core-shell quantum dots: Properties and applications. *J. Alloys Compd.* **2015**, *636*, 395–404. [[CrossRef](#)]
71. Zhou, J.; Yang, Y.; Zhang, C.Y. Toward biocompatible semiconductor quantum dots: From biosynthesis and bioconjugation to biomedical application. *Chem. Rev.* **2015**, *115*, 11669–11717. [[CrossRef](#)]
72. Park, N.M.; Kim, T.S.; Park, S.J. Band gap engineering of amorphous silicon quantum dots for light-emitting diodes. *Appl. Phys. Lett.* **2001**, *78*, 2575–2577. [[CrossRef](#)]
73. Adegoke, O.; Nyokong, T.; Forbes, P.B. Deposition of CdS, CdS/ZnSe and CdS/ZnSe/ZnS shells around CdSeTe alloyed core quantum dots: Effects on optical properties. *Luminescence* **2016**, *31*, 694–703. [[CrossRef](#)]
74. Ganguli, N.; Acharya, S.; Dasgupta, I. First-principles study of the electronic structure of CdS/ZnSe coupled quantum dots. *Phys. Rev. B* **2014**, *89*, 245423. [[CrossRef](#)]
75. Gloag, L.; Mehdipour, M.; Ulanova, M.; Mariandry, K.; Nichol, M.A.; Hernández-Castillo, D.J.; Gaudet, J.; Qiao, R.; Zhang, J.; Nelson, M.; et al. Zero valent iron core-iron oxide shell nanoparticles as small magnetic particle imaging tracers. *Chem. Commun.* **2020**, *56*, 3504–3507. [[CrossRef](#)] [[PubMed](#)]
76. Gloag, L.; Mehdipour, M.; Chen, D.; Tilley, R.D.; Gooding, J.J. Advances in the application of magnetic nanoparticles for sensing. *Adv. Mater.* **2019**, *31*, 1904385. [[CrossRef](#)] [[PubMed](#)]
77. Dhavale, V.M.; Singh, S.K.; Nadeema, A.; Gaikwad, S.S.; Kurungot, S. Nanocrystalline Fe-Fe₂O₃ particle-deposited N-doped graphene as an activity-modulated Pt-free electrocatalyst for oxygen reduction reaction. *Nanoscale* **2015**, *7*, 20117–20125. [[CrossRef](#)] [[PubMed](#)]