Vision, Modeling, and Visualization (2012)
M. Goesele, T. Grosch, B. Preim, H. Theisel, and K. Toennies (Eds.)

Automatic Detection and Recognition of Engineered Nanoparticles in SEM Images

S. Kockentiedt[1,2], K. Toennies[1], E. Gierke[2], N. Dziurowitz[2], C. Thim[2] and S. Plitzko[2]

[1]Department of Simulation and Graphics, Faculty of Computer Science, University of Magdeburg, Germany
[2]Federal Institute for Occupational Safety and Health, Germany

Abstract

Engineered nanoparticles have gained importance in recent years and will do so in the future, but their potential toxicity remains an open question. To better understand their effects on the human body, it is necessary to determine their concentration in ambient air. We propose a method to automatically detect nanoparticles in SEM images and differentiate engineered particles from other particles common in ambient air. The method reached G-means of 0.985, 0.779 and 0.820 for the classification against non-engineered particles of silver, titanium dioxide and zinc oxide respectively. This is comparable to manual classification.

Categories and Subject Descriptors (according to ACM CCS): I.4.8 [Image Processing and Computer Vision]: Scene Analysis—Object recognition

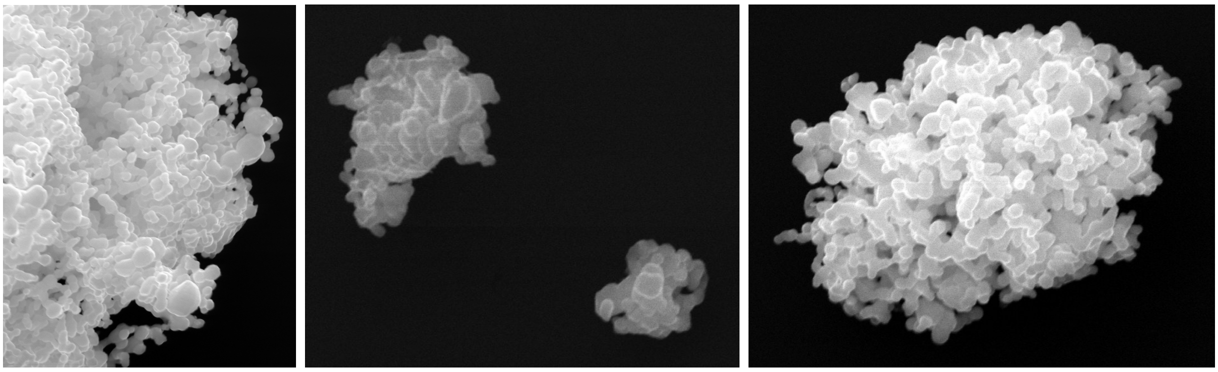
1. Introduction

Nanoparticles are particles with diameters ranging from 1 to 100 nm. These can occur naturally, e.g., produced by volcanic eruptions or forest fires. However, major sources of nanoparticles are automobile exhausts and factory fumes. In the past few years, a new type of nanoparticles, engineered nanoparticles, has gained importance. For example, silver nanoparticles are used in deodorants, toothpaste and fabrics for their antimicrobial activity, zinc oxide is used in sun cream and titanium dioxide is used in self-cleaning surfaces.

The potential toxicity of these nanoparticles has not been sufficiently explored [OMC*09]. Therefore, it is desirable to do further research and to introduce exposure limits for the handling of nanoparticles. For that, it is necessary to measure a person's exposition to different types of nanoparticles. A measurement system has to differentiate the engineered nanoparticles from other particles of the same scale which also occur in the ambient air, called background particles. Current automatic techniques to measure the concentration of nanoparticles in the air, however, are not sensitive to the composition of the measured particles [MHG09,NIO09]. This way, current automatic measurement methods cannot distinguish between engineered particles and background particles [SAN*10]. Therefore, to classify the nanoparticles based on size and composition, the particles are gathered and images of them are taken using transmission electron microscopy (TEM) or scanning electron microscopy (SEM) [MHG09, NIO09]. This method has the advantage that it is possible to only count the engineered nanoparticles and that the device used to gather the particles can be very small and can be carried by a worker to directly measure his/her exposure. Currently, however, it is necessary to manually locate and identify the type of each particle, which can be very time consuming depending on the number of particles [MHG09].

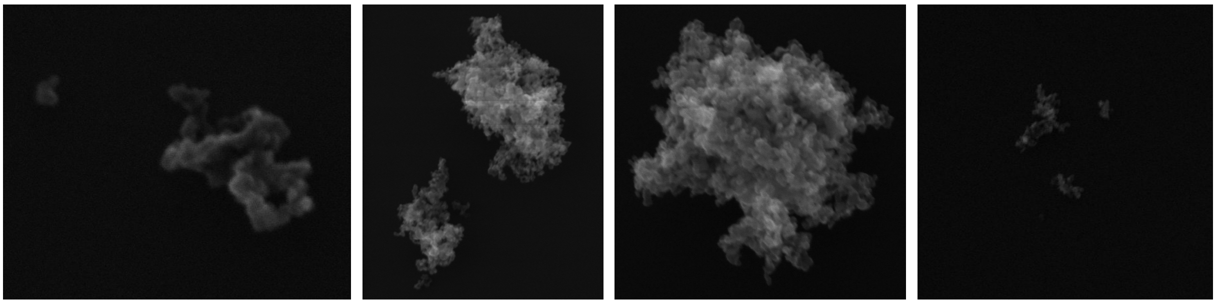
Our goal is to remove the need to manually locate and identify these particles and to find an automatic solution to detect as many particles as possible and to distinguish the particles of one engineered particle type at a time from other particles occurring in ambient air and in industrial scenarios in SEM images. The type of engineered particle is known beforehand so that our method can be trained to find that particular particle type. The constraint that only one type of engineered nanoparticles occurs is realistic because in a production environment, only one particle type will be produced at a time. Thus, the concentration of the currently produced particle type will be significantly higher than other types potentially produced before, which can be neglected. Therefore, it is no disadvantage to only consider the two-class case. Equally, the type of product particle type to detect is known in advance because it is the one currently produced at the industrial site.

DOI: 10.2312/PE/VMV/VMV12/023-030



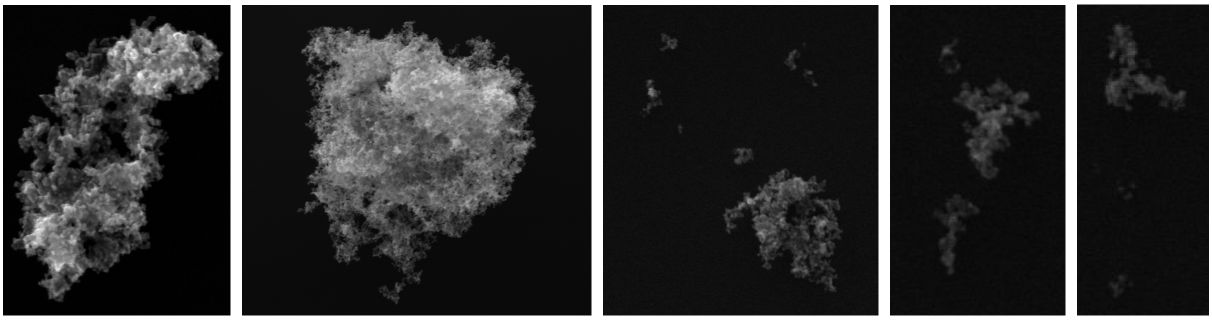
(a) Cropped Ag agglomerate **(b)** **(c)**

Figure 1. Examples of silver particles (not true to scale)



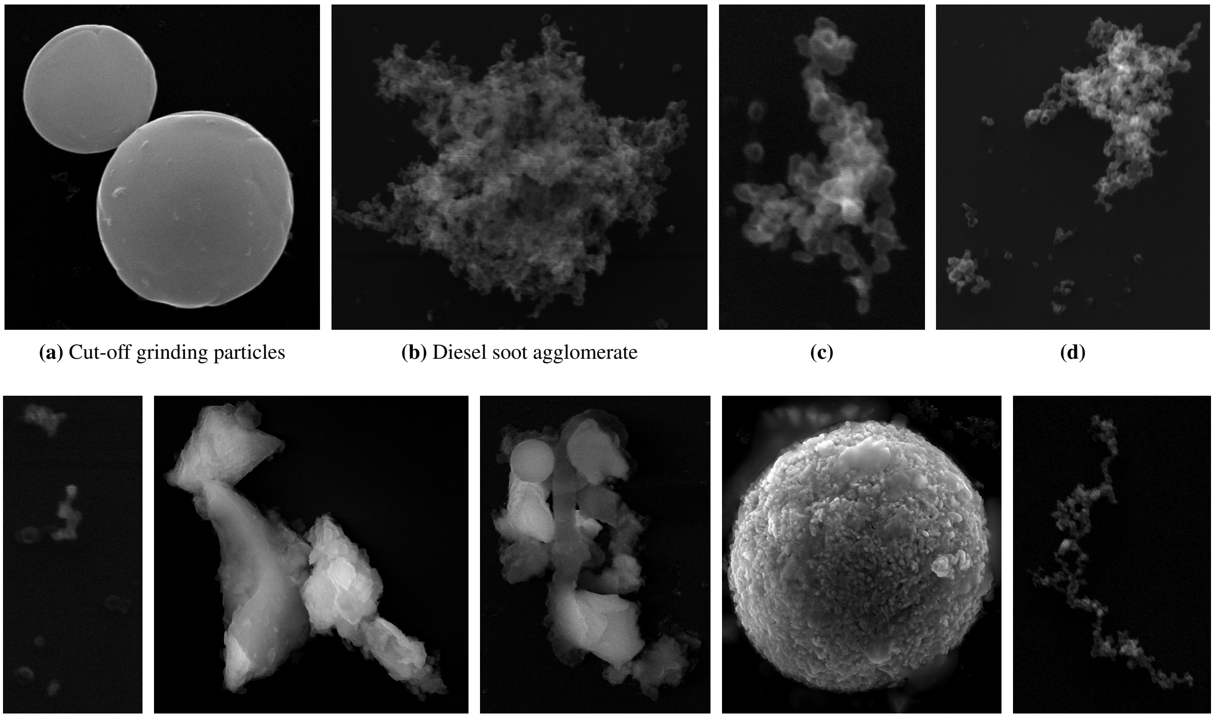
(a) **(b)** **(c)** **(d)**

Figure 2. Examples of titanium dioxide particles (not true to scale)



(a) **(b)** **(c)** **(d)** **(e)**

Figure 3. Examples of zinc oxide particles (not true to scale)



(a) Cut-off grinding particles **(b)** Diesel soot agglomerate **(c)** **(d)**

(e) **(f)** **(g)** **(h)** **(i)**

Figure 4. Examples of background particles (not true to scale)

In particular, we wanted to recognize silver (Ag), titanium dioxide (TiO_2) and zinc oxide (ZnO) particles with average diameters of 200 nm, 25 nm and 10 nm respectively. We chose these particle types because they are on the OECD's list of representative manufactured nanomaterials for safety testing [OEC10] and are commonly used in commercial products. Strictly speaking, the Ag particles are larger than nanoparticles, but we wanted to test how our methodology performs at different scales. TiO_2 and ZnO were also chosen as they are very similar to accidentally created nanoparticles such as diesel soot. They often have similar primary particle sizes and form agglomerates in a similar way (e.g., see Fig. 2c and 4b or Fig. 3c and 4d). A manual classification experiment containing large ZnO and diesel soot particles led to an accuracy of 72%. From now on, we will refer to both single particles and agglomerates simply as particles. If we specifically want to refer to single particles, we will call them primary particles.

2. Related Work

In contrast to most non-nanoscale objects, nanoparticles form agglomerates. Therefore, our method has to recognize agglomerates instead of single particles. This means that a common assumption of object classification does not apply: that the object to recognize has a well defined shape and size. E.g., the detection of biological particles [RTH*07] and pollen grains [RDCFFD06] depends on it. Instead, agglomerates can be small or big, jagged or smooth, elongated or round. They only differ in the size of their primary particles and, thus, in their local structure.

There is little literature on image-based nanoparticle classification. To the best of our knowledge, [FRAA03] and [Ost10] are the only works which cover the automatic classification of engineered nanomaterials via image analysis. Flores et al. [FRAA03] try to recognize different types of very small engineered metal nanoparticles in TEM images using co-occurrence matrices and polygon-based contour approximation.

Oster [Ost10] differentiates carbon nanotubes (CNT) from diesel soot and quartz particles in SEM images using a combination of isoperimetric quotient, band-pass filter responses and a nearest neighbor classifier achieving a precision of 100% and a recall of 83%. It must, however, be mentioned that only 18 CNT agglomerates and 36 other samples were used to evaluate the method.

Kindratenko et al. [KvET96] classify silver halide crystals and aerosol particles in SEM images using the fractal dimension and Fourier coefficients of the contour.

Flores et al.'s methods are not well applicable to our

Figure 5. Our workflow of detecting and recognizing engineered nanoparticles

problem because SEM images have different properties than TEM images and the particles we considered were larger and formed agglomerates. The problem covered by Oster is not realistic because the particles were hand-picked by their types and sizes as opposed to a real-world scenario where many other types and sizes can occur. In addition, all used particles are centered in the image so that they are fully visible and not clipped at the image border. Kindratenko et al. make use of additional information about the composition of the particles, which is only feasible for a limited number of particles.

It remains an open question, how well an automatic system to distinguish engineered nanoparticles from all types of other particles common in ambient air can perform. It also remains unanswered, which particle features are most suitable to be used in such a classification.

3. Materials and Methods

Here we describe the particles of our experiments and the three phases of our workflow, which is illustrated in Fig. 5.

3.1. Particle Gathering

The particles for our experiments were either gathered using a thermal precipitator [AWAS*09] or an electrostatic precipitator [CF98]. The samples of the engineered nanoparticles were produced by introducing material samples into a closed chamber and gathering these particles with a precipitator. Thus, each image contains only one type of particle. This allowed a reliable manual classification of the particles and agglomerates in the images preventing background particles being labeled as engineered particles and vice versa. SEM images of the gathered particles were then taken. The images had magnifications of 5,000 or 20,000 and a resolution of 4000×3200 pixels. That corresponds to pixel sizes of 5.1 nm and 1.3 nm respectively. The magnifications are a compromise between resolution and number of particles per image, as the image taking process is time-consuming. To keep the properties of different images consistent, guidelines for the image taking process were introduced. Table 1 lists for all sample types the number of microscopic images and number of particles or agglomerates in the images. The high particle counts for some types underline that a great amount of the particles are very small.

Table 1. The number of images and particles/agglomerates for the engineered particle types Ag, TiO_2 and ZnO and the background particles

Type	Images	Agglomerates
Ag	16	70
TiO_2	14	68
ZnO	10	56
Construction Dust	13	89
Diesel Soot	14	5471
Composite Material Dust	6	177
Cut-Off Grinding Dust	8	1173
Welding Smoke	12	814
Industrial Dust	7	891
Ambient Air	12	170

3.2. Particle Detection

As a characteristic of the image acquisition method, all engineered particles and almost all types of background particles are brighter than the image background and can be separated by thresholding (e.g., see Fig. 1–4). We tested three different threshold finding methods: The heuristic method by Zack et al. [ZRL77], fitting a mixture of two Gaussian distributions to the image's intensity distribution and the method by Kittler and Illingworth [KI86], which was found to perform best in a comparison with 40 other methods [SS04].

Testing these threshold finding methods by comparing them to manually determined ground truth thresholds showed that Zack's method produced the least segmentation errors. It is due to the fact that the background accounts for a very high percentage of the pixels in most images and the method by Zack et al. is mainly influenced by the background and makes only few assumptions about the foreground distribution which is a much weaker signal than the background.

The particle detection on the SEM images performs the following steps:

1. Smoothing using a Gaussian filter with $\sigma = 1.5$
2. Thresholding using Zack's method
3. Find 4-connected components in the resulting binary image
4. Discard any connected components that are smaller than 80 pixels

The engineered nanoparticles used in our experiments mostly occur in agglomerates with the smallest observable particle being $\sim$ 370 nm^2. Hence, our method classifies structures only if they occupy at least 80 pixels (corresponds to $\sim$ 35% of the smallest particle size at 20,000×). Setting a minimum object size reduces the risk to detect parts of the background as particles in noisy images.

3.3. Feature Extraction

As a result of the particle detection step, the positions of the particles are known. For statistical classification, every particle has to be represented by features which best capture the typical differences of the involved classes. The feature vector we used consisted of shape features capable of representing low and high frequency components of the contour and histogram-based intensity features.

The simplest features we used were the microscope's *Magnification* setting m for the corresponding image, the *Area* (A) of the particle in nm^2, the *Outer Contour Length* (L_O), which is the length of the contour which surrounds the particle, and the *Sum of Outer and Inner Contour Lengths* (L_T) calculated as L_O plus the sum of lengths of the contours of the holes inside the particle.

Another simple feature we used was the *Isoperimetric Quotient* (Q_I). It is defined as the ratio of the particle's area A and the area of a circle having a perimeter equal to L_T. It is calculated as $Q_I = 4\pi A/L_T^2$. If the particle has a circular shape, its value is 1 and it becomes smaller as the shape becomes more complex. This complexity can either stem from global features (e.g., a very elongated shape) or local features (e.g., a ragged shape). These features were selected as they capture some of the characteristics of nanoparticles visible at the boundary of the agglomerates.

3.3.1. Mean Contour Angle Wavelet Response

The aforementioned features are suitable to distinguish round shapes from irregular ones or small objects from large ones. However, the agglomerates often only differ in the size distribution of the primary particles. E.g., the agglomerates of engineered nanoparticles typically have small variance in primary particle sizes while background particles such as diesel soot agglomerates can contain primary particles from a wider size range. These size distributions are typically reflected in the texture and the contour of the agglomerates. Because the texture is greatly affected by the noise and blur in some images, we decided to concentrate on the contour.

The fractal dimension is a shape feature which is used in several publications [aGW01]. However, real objects do not really possess fractal dimensionality [KvET96], its computation is complex and, most importantly, it yields only one or two values, so that it isn't able to successfully discriminate complex shapes [DDF00]. In addition, the fractal dimension does not discriminate between particle sizes as it is scale-invariant.

Another popular tool are Fourier descriptors [ZL04]. However, these are more suitable for regular shapes such as crystals than for irregular contours [KvET96] like those of nanoparticle agglomerates because the approach tries to find regular global rather than changing local frequencies in the contour. This manifests in the fact that it cannot well handle changes in the phase of the waves which can easily happen in irregular contours.

Instead, we decided to use a scale-space approach which is able to capture local contour frequencies caused by the size distribution of the primary particles in an agglomerate. We used wavelets because they are able to detect these local frequencies. To apply a wavelet analysis, the contour shape has first to be converted into a 1D real-valued function. We chose the *Normalized Cumulative Angular Function* (also known as *Normalized Tangent-Angle Function*) originally proposed by Zahn and Roskies [ZR72] because it is suitable for non-convex shapes and the function's shape is invariant under translation and rotation. If we regard the outer contour of a detected particle or agglomerate as a series of points $(x(l), y(l)) = Z(l)$ as a function of arc length $l, 0 \le l \le L_O$, we define the *Cumulative Angular Function* $\phi(l)$ to be the change of the contour angle between the starting point and l. Note that the angle of the contour here is not defined relative to a particular point but relative to the global coordinate system. It follows that $\phi(0) = 0$ and $\phi(L_O) = -2\pi$. We define the *Normalized Cumulative Angular Function* $\phi^*(l) = \phi(l) + 2\pi l/L_O$ so that $\phi^*(0) = \phi^*(L_O) = 0$. This way, $\phi^* \equiv 0$ is true for a circular shape. Note that, in contrast to Zahn and Roskies, we do not set the domain of the function to the interval $[0, 2\pi]$ but instead to $[0, L_O]$ because we want to infer the real size of structures from the function. $\phi^*(l)$ can be understood as a periodic function with period L_O which is smooth if the original contour is smooth.

As wavelet functions, we used the real parts of Morlet wavelets $\psi_\lambda(t) = \mathrm{e}^{-\left(\frac{t}{\lambda}\right)^2} \cos\left(2\pi\frac{t}{\lambda}\right)$, where λ is the wavelength of the wavelet. To make the responses to wavelets with different wavelength more comparable, we normalized the wavelets' L^1-norm: $\psi_\lambda^*(t) = \frac{\psi_\lambda(t)}{||\psi_\lambda||_1} = \frac{1}{\lambda}\psi_1^*\left(\frac{t}{\lambda}\right)$.

The *Mean Contour Angle Wavelet Response* (W_λ) is defined as the mean absolute response of the contour function ϕ^* to the wavelet ψ_λ^* and is obtained by the convolution of the two functions:

$$W_\lambda = \frac{1}{L_O}\int_0^{L_O}\left|\int_{-\infty}^{\infty}\phi^*(l-t)\psi_\lambda^*(t)\,dt\right|dl. \qquad (1)$$

We computed the values of W_λ for the following wavelengths: 5, 10, 20, 50, 100, 200, 500, 1000 and 2000 nm.

3.3.2. Normalized Relative Histogram

Another distinguishing factor is the varying intensity of the particles. E.g., the Ag agglomerates often appeared brighter than many other particle types. To include information about the particles' intensity distributions, we computed a *Normalized Relative Histogram* (h_i) with 10 bins $h_0, \ldots, h_9$ for the pixels of every connected component. To compensate for contrast and brightness differences, we normalized the histograms so that the minimum intensity of each histogram was set to the modal intensity of the image which was also the most common background intensity. The maximum histogram intensity was set to the highest intensity present in the particle. While this could be prone to outliers, it compensates for varying contrast and brightness settings, which can be freely adjusted by the microscope operator and we never observed this to be a problem.

3.4. Classification

For the classification of the particles, we used the data mining software Weka [HFH*09]. Because the distributions of the applied features were unknown, a geometric classification approach using decision boundary construction was appropriate [JDM00]. We therefore used *Support Vector Machines* (SVMs) with polynomial (Poly) and radial basis function (RBF) kernels trained using sequential minimal optimization [Pla98] as classification algorithms. An advantage of SVMs is that they are less prone to class imbalance than other classification learning algorithms [SWK09]. Class imbalance means that one class is much more prevalent in the training data than the other class(es). In our data, the ratio of engineered nanoparticles to other particles was in the order of 1:100. Before training the classifiers, we normalized the feature values so they were in the interval $[0,1]$.

For both tasks of classifying TiO_2- and ZnO-particles against the background particles respectively, training the classifiers on all given samples did not work. In most cases, every sample was classified as a background particle. A good solution for that is to resample the data before training the classifier [SWK09]. For TiO_2 and ZnO respectively, we therefore oversampled the class of engineered particles using SMOTE [CBH02] and then randomly undersampled the class of background particles so that both classes had the same size.

Given a fixed number of training samples, the classification quality can decrease with the number of features. This is known as the *curse of dimensionality*. A common solution for it is feature selection [JDM00]. After determining the best classification parameters using all features, we used the genetic search algorithm described by Goldberg [Gol89, pp. 59–70] with a population size and generation count of 20, a crossover probability of 0.6 and a mutation probability of 0.033 to find the best feature set. Each feature set was evaluated with the parameters, which yielded the best results in a 10-fold cross-validation when using all features.

4. Evaluation and Results

The class imbalance problem needs to be considered in the choice of the classification performance measure which is used to evaluate the method. The commonly used *accuracy*, the ratio of correctly classified samples to the total number of samples, is a poor choice as classifying every sample as a part of the prevalent class (in our case, the background particles) leads to a relatively high accuracy of 0.99 for a ratio of 1:100. Sun et al. [SWK09] suggest *F-measure* and *G-mean* as appropriate measures in the case of class imbalance. We used G-mean because it only relies on the true positive rate and the true negative rate and, thus, in contrast to F-measure, normally doesn't change if the ratio between the classes changes. G-mean is defined as the geometric mean of the true positive rate and the true negative rate [KHM98]: $g = \sqrt{TP_r \cdot TN_r}$. To evaluate the classification quality, we used 10-fold cross-validation on the training data. Note that the SMOTE resampling was done separately for each step of the cross-validation guaranteeing that no training sample was created using the data of test samples.

The separation of particles from the background worked well, often detecting particles a person would overlook. Only very few artifacts occurred. The major problem was that some background particles which were hard to distinguish from the image background were divided into multiple parts by the algorithm. The classification results are summarized in Table 2. The polynomial kernel performed best for Ag and TiO_2 while ZnO was best classified by the RBF kernel. However, for all materials, the two kernel types yielded comparable results. For the classification of Ag particles and background particles, only 9 (Poly) and 7 (RBF) of all 8855 particles were incorrectly classified. In addition, 4 of the 5 false negatives that occurred were caused by the Ag particles being cut off at the image edges and, thus, not being fully visible (e.g., see Fig. 1a). Most of the false positives were agglomerates of few round particles with a size similar to the Ag primary particles. E.g., the two particles in Fig. 4a were created by cut-off grinding steel and are likely made of steel which melted and formed spheres. For TiO_2 and ZnO, almost all false positives were agglomerates of primary particles with a similar size to the engineered particles (e.g., see Fig. 2c and 4b). Most of these agglomerates were very small, making a manual classification almost impossible.

Every one of our 24 features was used by at least one of the classifiers. This suggests little redundancy between the features. For Ag and TiO_2, both kernel types show better results for a magnification of 5,000 while both kernel types for ZnO work better for a magnification of 20,000. That suggests that 5,000× is sufficient for Ag and TiO_2, which have larger primary particles (200 and 25 nm). The classification of ZnO with an average primary particle size of 10 nm can benefit from the increased resolution brought by a magnification of 20,000 making smaller details visible. The decrease in classification quality for the higher resolution for the other particle types could be explained by a smaller training set due to a smaller area being covered by the images and a higher probability of particles being cut off by image edges.

Table 2. The best parameter values and classification results for the different nanoparticles and SVM kernels. The following information is given: the kernel type, the soft margin parameter C, the exponent E of the polynomial kernel, the scaling parameter γ of the radial basis function kernel, the G-mean before the feature selection (g_{all}), the features selected by the genetic algorithm, the true positive rate TP_r, the true negative rate TN_r, the corresponding G-mean g and the G-means for the particles separate for magnifications of 5,000 ($g_{5\text{k}}$) and 20,000 ($g_{20\text{k}}$) respectively

Material	Ker.	C	E	γ	g_{all}	Selected Features	TP_r	TN_r	g	$g_{5\text{k}}$	$g_{20\text{k}}$
Ag 200	Poly	5	3	–	0.985	m, Q_I, W_5, W_{10}, W_{20}, W_{50}, W_{100}, W_{500}, W_{2000}, h_0, h_1, h_2, h_3, h_4, h_5, h_6, h_8	0.971	0.999	0.985	0.991	0.957
	RBF	10	–	1.0	0.971	m, Q_I, W_5, W_{20}, W_{50}, W_{100}, W_{200}, W_{500}, h_2, h_3, h_5, h_7	0.957	1.000	0.978	0.982	0.957
TiO_2 25	Poly	1	2	–	0.735	m, Q_I, W_5, W_{10}, W_{20}, W_{100}, W_{500}, h_0, h_2, h_3, h_5, h_8	0.750	0.809	0.779	0.785	0.746
	RBF	1	–	0.005	0.726	A, L_O, Q_I, W_5, W_{10}, W_{50}, W_{500}, W_{1000}, h_1, h_2, h_8, h_9	0.691	0.784	0.736	0.744	0.713
ZnO 10	Poly	300	1	–	0.795	A, L_T, Q_I, W_{20}, W_{50}, W_{200}, W_{500}, h_0, h_6, h_7, h_8	0.821	0.814	0.818	0.808	0.866
	RBF	10	–	0.1	0.797	L_T, W_{20}, W_{50}, W_{500}, h_0, h_2, h_3, h_6, h_8, h_9	0.804	0.838	0.820	0.808	0.885

5. Conclusions and Future Work

We have shown that a reliable automatic detection of engineered nanoparticles in SEM images is possible provided they are sufficiently distinct in shape. The approach is limited to cases where particles of different types do not overlap. However, this can be achieved by controlling the population density of particles on the sample carrier by reducing the gathering time of the used precipitator.

The classification of Ag nanoparticles against background particles yielded a true negative rate of 0.999 and a true positive rate of 0.971 while most of the false negatives were caused by particles only being partly visible. Our approach does not need human interaction and makes it feasible to search for specific particles in a large number of images.

The detection of the titanium dioxide and zinc oxide particles was more challenging. As mentioned before, this is caused by the high resemblance to accidentally created nanoparticles such as diesel soot due to similar primary particle sizes and the high amount of small agglomerates. The small particles only occupy a limited number of pixels. Together with the blur and relatively strong noise typical for SEM, this leads to a lack of distinct contour features. Given that the manual classification of these particles is very difficult, the obtained results represent a good outcome.

The classification of ZnO particles in images with a magnification of 20,000 resulted in a G-mean of 0.885 compared to 0.808 for 5,000. This indicates that higher magnifications could lead to better results for ZnO. However, this would require more images to capture the same amount of particles. Also, single agglomerates could become too large to be captured by a single image.

All used features seem to be appropriate to distinguish the nanoparticles as each of them was present in at least one set of selected features.

A simple way to improve the classification reliability is to gather more training samples, especially from engineered particles. Further improvement could be brought by adding more features. Since we currently do not take texture into account, we will investigate ways to compute such features in view of the fact that agglomerates greatly vary in size and, hence, the texture measures must scale well with size. Another possibility that we will investigate is the detection of circular patterns in the contours of agglomerates to infer the size of the primary particles. This will require the use of magnifications for which the primary particles are clearly visible. We will also have to deal with particles which are only partially visible, which is, however, fairly easy as it only requires to take image boundaries into account when computing the features.

References

[aGW01] AP GWYNN I., WILSON C.: Characterizing fretting particles by analysis of SEM images. *European cells & materials 1* (Jan. 2001), 1–11. 5

[AWAS*09] AZONG-WARA N., ASBACH C., STAHLMECKE B., FISSAN H., KAMINSKI H., PLITZKO S., KUHLBUSCH T. A. J.: Optimisation of a thermophoretic personal sampler for nanoparticle exposure studies. *Journal of Nanoparticle Research 11*, 7 (Aug. 2009), 1611–1624. 4

[CBH02] CHAWLA N. V., BOWYER K. W., HALL L. O.: SMOTE: synthetic minority over-sampling technique. *Journal Of Artificial Intelligence Research 16* (2002), 321–357. 6

[CF98] CHOI B., FLETCHER C.: Turbulent particle dispersion in an electrostatic precipitator. *Applied Mathematical Modelling 22*, 12 (Dec. 1998), 1009–1021. 4

[DDF00] DROLON H., DRUAUX F., FAURE A.: Particles shape analysis and classification using the wavelet transform. *Pattern Recognition Letters 21*, 6-7 (June 2000), 473–482. 5

[FRAA03] FLORES A. B., ROBLES L. A., ARIAS M. O., ASCENCIO J. A.: Small metal nanoparticle recognition using digital image analysis and high resolution electron microscopy. *Micron 34*, 2 (Feb. 2003), 109–118. 3

[Gol89] GOLDBERG D. E.: *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley, 1989. 6

[HFH*09] HALL M., FRANK E., HOLMES G., PFAHRINGER B., REUTEMANN P., WITTEN I. H.: The WEKA data mining software: an update. *ACM SIGKDD Explorations Newsletter 11*, 1 (Nov. 2009), 10–18. 6

[JDM00] JAIN A. K., DUIN R. P., MAO J.: Statistical pattern recognition: a review. *IEEE Transactions on Pattern Analysis and Machine Intelligence 22*, 1 (2000), 4–37. 6

[KHM98] KUBAT M., HOLTE R. C., MATWIN S.: Machine learning for the detection of oil spills in satellite radar images. *Machine Learning 30*, 2-3 (1998), 195–215. 6

[KI86] KITTLER J., ILLINGWORTH J.: Minimum error thresholding. *Pattern Recognition 19*, 1 (Jan. 1986), 41–47. 4

[KvET96] KINDRATENKO V. V., VAN ESPEN P. J. M., TREIGER B. A.: Characterisation of the shape of microparticles via fractal and fourier analyses of scanning electron microscope images. *Mikrochimica Acta suppl. 13* (1996), 355–361. 3, 5

[MHG09] METHNER M., HODSON L., GERACI C.: Nanoparticle emission assessment technique (NEAT) for the identification and measurement of potential inhalation exposure to engineered nanomaterials: part A. *Journal of Occupational and Environmental Hygiene 7*, 3 (2009), 127–132. 1

[NIO09] *Approaches to safe nanotechnology: managing the health and safety concerns associated with engineered nanomaterials*. National Institute for Occupational Safety and Health, 2009. 1

[OEC10] *List of manufactured nanomaterials and list of endpoints for phase one of the sponsorship programme for the testing of manufactured nanomaterials: revision*. OECD, 2010. 3

[OMC*09] OSTROWSKI A. D., MARTIN T., CONTI J., HURT I., HERR HARTHORN B.: Nanotoxicology: characterizing the scientific literature, 2000-2007. *Journal of nanoparticle research 11*, 2 (Feb. 2009), 251–257. 1

[Ost10] OSTER T.: *Erkennung von Nanofaseragglomeraten in REM-Bildern auf Basis von Texturinformationen*. Bachelor thesis, Otto-von-Guericke University of Magdeburg, 2010. 3

[Pla98] PLATT J. C.: Fast training of support vector machines using sequential minimal optimization. In *Advances in kernel methods: Support vector machines*, Schölkopf B., Burges C. J. C., Smola A. J., (Eds.). MIT Press, Cambridge, 1998, ch. 12. 6

[RDCFFD06] RODRIGUEZ-DAMIAN M., CERNADAS E., FORMELLA A., FERNANDEZ-DELGADO M.: Automatic detection and classification of grains of pollen based on shape and texture. *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews) 36*, 4 (July 2006), 531–542. 3

[RTH*07] RANZATO M., TAYLOR P., HOUSE J., FLAGAN R., LECUN Y., PERONA P.: Automatic recognition of biological particles in microscopic images. *Pattern Recognition Letters 28*, 1 (Jan. 2007), 31–39. 3

[SAN*10] SAVOLAINEN K., ALENIUS H., NORPPA H., PYLKKÄNEN L., TUOMI T., KASPER G.: Risk assessment of engineered nanomaterials and nanotechnologies—a review. *Toxicology 269*, 2-3 (Mar. 2010), 92–104. 1

[SS04] SEZGIN M., SANKUR B.: Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging 13*, 1 (2004), 146–165. 4

[SWK09] SUN Y., WONG A. K., KAMEL M. S.: Classification of imbalanced data: a review. *International Journal of Pattern Recognition and Artificial Intelligence 23*, 4 (2009), 687–719. 6

[ZL04] ZHANG D., LU G.: Review of shape representation and description techniques. *Pattern Recognition 37*, 1 (Jan. 2004), 1–19. 5

[ZR72] ZAHN C. T., ROSKIES R. Z.: Fourier descriptors for plane closed curves. *IEEE Transactions on Computers C-21*, 3 (Mar. 1972), 269–281. 5

[ZRL77] ZACK G. W., ROGERS W. E., LATT S. A.: Automatic measurement of sister chromatid exchange frequency. *Journal of Histochemistry & Cytochemistry 25*, 7 (July 1977), 741–753. 4