Bringing into play automated electron microscopy data processing for understanding nanoparticulate electrocatalysts’ structure–property relationships
Ana Rebeka Kamšek¹,², Francisco Ruiz-Zepeda¹, Andraž Pavlišič³, Armin Hrnjic¹,⁴ and Nejc Hodnik¹,⁴

Abstract
Observing and quantifying information at the atomic scale plays an essential role in assessing the structure–property relationships in electrocatalysis. Particularly, when studying supported nanoparticulate fuel cell and electrolyzer electrocatalysts, resolving nanoparticles’ structural features at the atomic scale and their evolution as a response to external stimuli is of great relevance. Atomically resolved electron micrographs of identical locations before and after induced changes are a still vastly unexplored resource of quantifiable data that can be used to elucidate structure–activity and–stability relationships of studied materials. In this short review, we highlight the recent approaches and opportunities in processing electron microscopy images and the development of their analysis algorithms enabling the acquisition of unprecedented structural information, focusing on systems of metallic nanoparticles.

Introduction
Understanding true structure–property relationships of electrocatalysts requires atomistic insights into their structure and dynamics. A typical representative is a proton exchange membrane fuel cell catalyst used for oxygen reduction reaction, consisting of supported Pt or Pt-alloy nanoparticles. Despite many efforts, comprehensive prediction of their properties remains an unresolved challenge [1–4].

Real-world batches of fuel cell electrocatalysts contain unimaginably large numbers of Pt-alloy nanoparticles that exhibit a wide variety of atomic-scale differences in sizes, morphologies, compositions, and crystal structures. More interestingly, a diverse range of defects, crystal facets, twin boundaries, concave sites, steps, kinks, and other anomalies influence the nanoparticles’ catalytic behavior [5]. Additionally, as the locations of atoms often deviate from perfect lattice positions, and as the composition in the surface layers can differ from the one in the core, the resulting strain also affects the catalytic activity [6]. It is crucial to recognize that those parameters can not only be interconnected but can also change during operation and that therefore degradation mechanisms can have a large impact on catalyst performance [1,7–9]. Such degradation phenomena occur during operation and are linked to multiple parameters besides the electrocatalyst structure and composition, such as potential, pH value, temperature, purity of fuel, start-up/shut-down states and other operation conditions [7]. Figure 1 depicts a schematic representation of the complexity of real-world systems and the multifaceted approach required to take on the challenge of explaining their structure–property relationships.

Advanced transmission electron microscopy (TEM) characterization techniques can reveal the structure and chemical state of nanoscale materials, shedding light on the investigation of the structure–property relationships. Particularly, aberration-corrected scanning transmission electron microscopy (AC-STEM) can provide very local and highly precise surface and near-surface...
In order to accurately track structural changes, it is essential to gather additional data regarding the dynamic changes occurring during a reaction instead of just still images of chosen sample regions, since not all nanoparticles exhibit the same changes despite being exposed to the presumably same external stimuli [1]. The temporal evolution of materials is often studied with in-situ methods; however, a better time resolution in TEM means sacrificing atomic spatial resolution besides enduring beam damage effects [8,12,13].

Identical location STEM (IL-STEM) offers a bottom-up approach to study such systems by imaging the same location before and after electrochemically induced changes to draw reliable conclusions about restructuring events [12]. Atomically resolved identical location images are a rich resource of quantitative information and can offer a more detailed perspective of the studied phenomena [14], which calls for new custom-made image analysis algorithms to explore and take full advantage of it. The goal of data-oriented (S)TEM is to analyze the data in a way that enables objectively and accurately observing meaningful connections within datasets [15]. Additionally, as (S)TEM follows today’s trends in generating big data, computer algorithms are not only a smart way to lower the amount of time that would be needed for manual analysis but might sometimes also be the only way to analyze all available data [15]. Thus, automated image analysis refers to sequences of image processing steps that are performed consecutively and successfully without any need for human intervention. This short review provides a summary of recent efforts of the scientific community to improve the quality and gain physically meaningful knowledge from atomically resolved electron micrographs of nanoparticulate electrocatalysts, focusing on extracting information about their structure—property relationships and the importance of implementing these advances in an automated manner.

**Image analysis algorithms**

There exist multiple approaches for automatic image processing, spanning from various classical algorithms to novel tools involving artificial intelligence [10,15,16]. When dealing with atomically resolved STEM imaging, the goal is to extract quantifiable information about the imaged structure. Nonetheless, we must keep in mind certain challenges that come with data analysis, such as the interpretability, the level of confidence in the prior knowledge, and the inherent characteristics of the analyzed sample, in order to correctly describe the information from the observed location and connect it to the information from other sources, which is averaged across the entire specimen [10].

Since a higher signal-to-noise ratio (SNR) of an image improves our ability to identify and retrieve relevant information about the structure and composition of the studied material, it is desirable to improve the quality of...
the image before starting any further analysis. This can be effectively done by different processing approaches.

Conventional and machine learning algorithms are two paths one can take when attempting to quantify information from electron microscopy datasets. Despite a multitude of analytical approaches to complete such a task, they often lack the efficiency compared to deep learning methods and only achieve a performance similar to that of an experienced human operator [10]. The main conceptual difference between conventional and machine learning approaches lies in how they approach problem solving. To yield the desired output, conventional algorithms require input data and rules that govern relationships between the input and output data. On the contrary, machine learning algorithms do not require those rules, as their main task is to figure them out and make predictions about new data based on the knowledge gained from the training data. Taking this inherent difference between the two groups of algorithms into account, machine learning is often better suited for solving problems where rules are difficult to accurately establish, but gathering a sufficient amount of training data does not pose a problem. Among the machine learning algorithms, supervised methods work with labeled data while unsupervised methods work with unlabeled data. They are becoming easier for domain experts to use due to a range of freely available collections of pre-trained models and archived datasets [16].

Since high-throughput advances in STEM have enabled a rapid generation of vast datasets, containing more information than ever, it is our belief that applying computer algorithms from other fields to electron micrographs of electrocatalysts presents an opportunity worth exploiting. Extracting a physically meaningful signal that provides results about the microstructure and crystallography of the investigated sample can be, in our opinion, greatly augmented. By harnessing the power of tools, currently less familiar to the electrochemistry community, it is possible to establish the structure–property relationships of electrocatalysts. Finally, the advantage of developing a reliable algorithm for a specific task and dataset lies in its possibility to reuse it as many times as needed, saving precious researcher time and energy.

As the number of different implementations of similar algorithms is very high, we chose to focus on those that could be the most beneficial for atomic-scale studies of nanoparticulate electrocatalysts. Here, it is worthwhile to mention that while not all chosen algorithms were implemented for STEM imaging of individual nanoparticles, the ideas behind them are often (fully or partially) transferrable, and possess great potential to be successfully employed in alternative purposes.

**Improving image quality**

Raw bright field or annular dark field STEM images often suffer from noise and distortions, and more particularly when images are acquired in an attempt to limit the radiation dose to prevent subsequent beam damage to the sample. Aberrations, statistical and scanning noise, along with possible sample drift and beam damage, are artifacts that lower the image quality [16]. In atomically resolved micrographs of nanoparticles, it is not only important to reduce the impact from the listed factors, but also to emphasize the signal from the atomic columns of the investigated nanoparticle and reduce the signal from the other parts of the image, coming either from the support or from other nanoparticles in the vicinity that are not in focus and thus at atomic resolution. However, successful denoising and deblurring of atomically resolved images are challenging and sometimes conventional algorithms cannot offer acceptable results, impacting on subsequent analytical steps [17]. For very noisy STEM images, where low-contrast details are difficult to recognize with a bare eye, deep-learning-based methods are often more robust than conventional methods such as edge detection and thresholding [16]. This is especially important to take into account when the sample thickness is not uniform, leading to low-contrast atomic columns on nanoparticle edges, a common challenge when imaging carbon-supported Pt-based nanoparticulate catalysts [17].

**Denoising**

Algorithms for improving the signal-to-noise ratio and drift distortions in electron micrographs are plentiful and have become extraordinarily successful even when dealing with lower quality data [18,19]. A rigid registration approach was proposed for correcting a series of low quality cryo-STEM images including possible unit cell misalignment [20]. Similarly, scan artifacts in sequences of serially acquired STEM images can be compensated via a non-rigid registration approach [21,22], and via a Gaussian process-based regression method to separate drift and random distortions [23].

While registration-based methods are immensely useful when sequences of images are available, deep learning comes in handy when analyzing individual noisy images. Deep convolutional neural networks proved useful for restoring low-dose images of metal clusters on lighter support films [24] and an encoder-decoder-type deep learning model was developed for noise reduction and atomic column localization of different crystal structures, shown in Figure 2a [17].

Similarly, the concepts are transferrable to other types of signal imaging, hence complementing the traditional STEM imaging. A non-rigid registration approach
followed by principal component analysis was used to denoise energy dispersive X-ray spectroscopy (EDS) data, often acquired simultaneously with annular dark field (ADF) STEM images [25]. Autoencoders, a type of unsupervised neural network, were used together with a classifier to denoise and classify electron energy loss spectra [26]. Tensor singular value decomposition was used to denoise a dataset comprised of EDS spectra and atomically resolved 4D STEM images [27], the latter being a large collection of full 2D diffraction patterns, recorded at each pixel position of the probe [28]. A frame averaging and Fourier filtration approach was used in graphene liquid cell transmission electron microscopy to subtract the graphene background and improve the quality of serially acquired images [29]. Since the currently available denoising algorithms are versatile and at the same time conveniently ready for adaptation, they are of great value when analyzing nanoparticle images because of the possible overlap of the atomically resolved information with the rest of the signal, and because of the presence of low-contrast columns on the edges. Successfully holding all information about the nanoparticle(s) under investigation is a step towards reliable tracking of dynamic changes occurring during a reaction.

Determining atomic column positions

Converting atomically resolved electron micrographs into quantifiable information such as precise locations of atomic columns is a necessary step in advanced image analysis. Since catalytic reactions take place on the nanoparticle surface, it is crucial to correctly determine the positions of identified columns on the nanoparticle edges. A golden standard in establishing atomic column positions is by modelling the image with 2D Gaussian peaks [30]. Using a statistical parameter estimation approach, atomic column locations were determined also for images including light-element atomic columns in the presence of heavier ones in annular bright and dark field imaging [31] and for images featuring single atoms [32]. Complementary approaches to this task include utilizing an encoder-decoder-type deep learning model [17] and an algorithm based on structural similarity [33], both able to surpass the 2D Gaussian fit method.
Extracting information from atomically resolved electron micrographs

A precise determination of the observed structure by converting atomically resolved images or movies into a list of atomic coordinates or trajectories is a good starting point for subsequent data analysis [37]. Structural information that can be learned from atomically resolved images of nanoparticles includes crystallographic phase and defect identification, strain analysis and 3D reconstruction [1,15].

Extracting crystallographic information

Crystallographic analysis of nanoparticulate electrocatalysts is often focused on investigating the presence of different crystallographic phases in the sample, since the presence of an intermetallic crystal structure is one of the parameters influencing its catalytic properties [1]. For instance, being able to, automatically determine the mass ratio of ordered and disordered alloy in a sample, would enable a more quantitative understanding of this aspect of electrocatalysts’ structure—property relationships. A machine learning approach exploring different algorithms and a deep learning approach were developed to segment high angle annular dark-field (HAADF) STEM images into separate phases as shown in Figure 2b [34,38]. Furthermore, a deep convolutional neural network was trained to determine the Bravais lattice symmetry in an image based on classifying 2D fast Fourier transformations [39].

In a typical population of nanoparticles, different crystal defects such as twin boundaries, stacking faults, impurities and vacancies can be found, manifesting themselves in HAADF STEM images as local deviations in the contrast and lattice periodicity. A number of studies tackled the challenge of automatic defect detection, including those utilizing a support vector machine as an unsupervised machine learning method [40], geometric graph theory as shown in Figure 2c [35], a weakly supervised approach with a deep neural network [41], and a convolutional neural network [42].

Strain analysis

For Pt-alloy nanoparticles, the formation of a core—shell structure, where the outermost atomic layers are richer in Pt than the nanoparticle core, the resulting compressive surface strain can improve the catalyst performance. Understanding the strain dependence on the distribution of atoms can be achieved by precisely evaluating atomic column displacements on the nanoparticle surface and close to structural defects [1,6].

Displacement and strain analysis in atomically resolved images can be performed by measuring shifts in individual column positions and comparing the measured positions with those of an ideal lattice to reveal facet contraction and expansion [1,22,30,36]. Fourier-space geometric phase analysis as shown in Figure 2d [36], a localised method where an effective lattice parameter is determined for each atomic column based on its first neighbors [43], and a method for determining local strain by using multiple references with different lattice symmetries [44] were all used as methods for strain analysis.

Accessing 3D information

When imaging nanoparticles with STEM, it is important to keep in mind that images are 2D representations of 3D objects. Therefore, obtaining information in three dimensions is inherently more difficult, however under certain conditions accessible [11]. As changes to the nanoparticle structure and composition occur in 3D, obtaining that information results in a more comprehensive overview of the structure—property relationships.

Electron tomography is a powerful tool for exploring the 3D morphology and composition of materials, including but not limited to studying crystal defects such as dislocations, steps, kinks, grain boundaries, chemical order, and strain. Both experimental tomography methods and reconstruction algorithms for small species investigation are undergoing fast development and today offer acquiring information also at the atomic scale [45,46]. Typically, a tilt series of projections is acquired and then reconstructed into a 3D model. However, conventional electron tomography sample holders do not always offer a full tilt range of 180° with small tilt increments due to the limited space inside the objective lens in a microscope and the limited number of projection images that can be taken before the electron beam damage may take place [46].

Nonetheless, reconstructions can still be accurately and promptly determined for tilt ranges of ±75° [37,47]. To minimize possible artifacts, arising during the reconstruction of a large and compact assembly of metallic nanoparticles, an improved reconstruction method was proposed by acquiring and merging two tilt series [48]. Also, information from both real and reciprocal space was used to determine a 3D reconstruction of objects from a limited number of projections [49]. Retrieving the coordinates of the corresponding atomic sites in 3D means going one step further, as shown in Figure 3a, where chemical ordering and grain boundaries in an alloyed nanoparticle were accurately determined [50]. Likewise, atomic-scale 3D reconstruction and strain mapping was achieved for elemental nanocrystals, imaged with graphene liquid cell transmission electron microscopy, as shown in Figure 3b [29].

Thanks to modern algorithm development, a number of studies successfully reconstructed the 3D model of an object from only a few 2D projections. Since the image contrast in HAADF STEM imaging is dependent on the
composition and thickness of the studied material, pixel intensities can be used to determine the composition and count the number of atoms in an atomic column [30,51]. Small differences in local image intensity cannot be distinguished visually but play an important role in measuring dynamic changes, which can be studied also using a hidden Markov model, as shown for Pt nanoparticles [52]. Furthermore, the atom counting approach was combined with depth sectioning to reconstruct the 3D morphology of nanoparticles [53], while combining it with molecular dynamics algorithms enabled a 3D nanoparticle reconstruction from a single projection as shown in Figure 3c [54,55].

It is worth noting that there remain certain limitations to atomic-level tomography. A limited amount of projections lowers the reconstruction reliability in the missing wedge direction [48,49]. This can to some extent be solved with compressive sensing by filling in the missing data [16], however subtle variations in the structure can still remain uncovered [49]. It was found that the results are more accurate when the nanoparticle in question is isotropic and without any voids [49,54]. Additionally, multiple scattering may affect the quality of the reconstruction [49], especially on the particle surface and when reconstructing larger particles [29]. Provided that the atomic columns in a nanoparticle are made up of different chemical elements, this increases the number of free parameters and makes the 3D reconstruction more complex, however still achievable [29,50]. Once again, we note that the presented algorithms are not limited to a specific material, but are rather adjustable to diverse nanoparticulate systems.
Studying structure–property relationships of electrocatalysts

Applying previously discussed image analysis algorithms to sets of identical location or in-situ images of nanoparticles, undergoing structural changes, is beneficial in trying to explain the mechanisms affecting both catalyst activity and stability since the atomic configuration of a nanoparticle can evolve during operation [1]. This section focuses on presenting several studies featuring automated image processing to quantify the information pertaining to nanoparticles’ structural characteristics or their dynamic changes in various environments.

Atomic scale particle evolution and crystal phase analysis for PtCo nanoparticles were performed by Hrnjic et al. on images acquired by IL-STEM before and after electrochemical activation as shown in Figure 4a [56]. With this study, the authors showed specifically which facets of a particle suffered from dissolution and/or redeposition. In another investigation by Xin et al., restructuring and strain evolution of PtCo nanoparticles were tracked with environmental TEM during their reaction with oxygen and hydrogen gases [57]. Using fast dynamic STEM combined with a denoising algorithm, Henninen et al. revealed trends in the nucleation of Pt clusters [58]. Finally, thermal motion and coalescence dynamics of Au nanoparticles were tracked using graphene liquid cell electron microscopy in studies from Kang et al. and Bae et al. [59,60]. The mentioned studies took advantage of the quantifiable information, available in 2D atomically resolved micrographs, and associated it with the main processes driving nanoparticle dynamics. In such reports, image analysis algorithms are used as an efficient and reliable way to quantify the data, however, the interpretation remains in the domain of a human expert.

Going a step further, a number of studies elucidated atomic-scale restructuring in three dimensions. Atomic-scale 3D transformations of Pt nanoparticles were studied by Altantzis et al. under the flow of selected gases to observe the evolution of their faceted morphology as...
shown in Figure 4b [61]. Similarly, Liu et al. [62] and Albrecht et al. [63] investigated 3D structural evolution at high temperature for Au and Au/Pd nanoparticles to evaluate their thermal stability. In another study, the course of FePt nanoparticles’ early-stage nucleation was studied by Zhou et al. [64]. In-situ tomographic studies were done for investigating the oxidation mechanism and hollow/porous structures of Ni–Fe nanoparticles by Xia et al. [65], for studying the structural heterogeneity including defects and strains of Pt nanocrystals in solution by Kim et al. [66] and for studying the shape and order evolution of PtCu nanoframes during heating by Gong et al. [67]. Finally, Gong et al. created 3D maps of surface atoms’ coordination numbers for Pt nanocrystals to expose the link between local coordination structure and catalytic performance as shown in Figure 4c [68] and Lee et al. determined the 3D structure, the surface strain map and the ORR activity map of a twinned Pt nanoparticle [69]. Including atomically resolved 3D information enables a transition from identifying restructuring events to directly establishing links to the local 3D structure as well as providing an accurate structural model for subsequent calculations.

Conclusions and outlook

Nowadays, we are fortunate enough to take advantage of both advanced catalysis synthesis and advanced characterization techniques that generate large amounts of potentially valuable data. Our goal should be to extract as much information out of these datasets as possible by automated data analysis algorithms. The obtained information should work hand-in-hand with theoretical computational methods, such as kinetic Monte Carlo or density functional theory, which will, in turn, provide the necessary feedback for better elucidation of materials properties. Among such computational methods, kinetic Monte Carlo is, in our opinion, of especial value as it can be employed to investigate the evolution of the atomic structure using information, extracted from electron micrographs as an initial condition for the simulation. If the IL-STEM and kinetic Monte Carlo results match, this can serve as an indication of which processes contribute the most to dynamic structural changes. Understanding the high complexity and interrelations between different structural changes, such as dissolution and facet development, can help substantially in explaining the structure–property relationships of electrocatalysts.

An approach of particular interest is coupling atomically resolved identical location and/or in-situ (S)TEM with the development of novel 3D reconstruction techniques and machine learning algorithms to identify changes in morphology, chemical composition, strain, crystal and electronic structure of electrocatalysts. Furthermore, automated data analysis for large datasets from, for example, 4D STEM is no longer an option, but a necessity, making such techniques an ideal playground for using advanced image analysis including machine learning algorithms to resolve functionalities of materials at the atomic level.

In conclusion, recent progress in advanced algorithms for the extraction of information in electron micrographs has led to new insights and brings a great opportunity for breakthroughs in the field of electrocatalysis, where the understanding of structure–property relationships is a fundamental step for improving the materials design. We believe that by combining domain expertise from different fields spanning from materials science to data science, a highway towards better electrocatalyst understanding and thus more effective development is emerging.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

Acknowledgement

Authors gratefully acknowledge prof. dr. Goran Dražić for his valuable comments and fruitful discussions about the subject of this work. We would like to thank the Slovenian research agency (ARRS) program P2–0393; the projects N2–0106, NC-0007, NC-0016; NATO Science for Peace and Security Program under grant G5729; and European Research Council (ERC) Starting Grant 128STABLE (Grant agreement ID: 852208) for funding the study. A.R.K. would like to acknowledge the Janko Jarmikan Doctoral Scholarship.

References

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest
** of outstanding interest


The authors compiled a large training library of simulated STEM images along with a deep learning model for denoising, deblurring, and precise atom segmentation. The model performs well also on images featuring atomic columns with a large thickness variation.


The authors developed a reliable framework for the analysis of graphene liquid cell electron micrographs of nanocrystals. They implemented particle tracking, background subtraction, identification of low-quality images, alignment and averaging, and atomic-resolution 3D reconstruction including strain analysis.


Using a semi-supervised machine learning framework, the authors of this study successfully segmented STEM images of oxide material systems. The method is robust against noise and does not require a
large dataset, which makes it useful for rapid image classification and feature mapping.


The authors of this study presented a methodology for defect recognition using geometric graph theory. The authors tracked the evolution of the 3D atomic structure of Pt nanoparticles in a variable gaseous environment. They quantified the facet evolution under the flow of selected gases and provided direct morphological information in the reactive environment.


