Original article

Journal of High School Science

Automated particle detection in Atomic Force Microscopy images of environmental films using two novel algorithms

Zang E

Submitted: October 7, 2025, Revised: version 1, November 13, 2025
Accepted: November 14, 2025

Abstract

Reliable particle identification in Atomic Force Microscopy (AFM) images is challenging, particularly for environmental films—surface layers that accumulate diverse nanoscale particles outdoors. The commonly used Igor Pro™ software requires extensive manual input, introducing user variability while producing merged clusters and zero-area artifacts. By implementing steps such as the white top-hat transform and Adaptive Gaussian Threshold, this work presents two fully automated algorithms for AFM particle detection. One emphasizes speed, reducing computational time by an order of magnitude, while the other prioritizes accuracy, improving the F1 score by over 40% compared to Igor Pro™. Both approaches eliminate common artifacts and allow for reproducible, high-throughput particle analysis. These results demonstrate that automated algorithms can significantly advance AFM-based studies of films, bridging the gap between complex, resource-intensive and simple, error-prone methods.

Keywords

Image processing algorithms, Particle identification, Atomic Force Microscopy, Environmental films, Automated image analysis, White top-hat transform, Adaptive Gaussian threshold, High-throughput algorithms, Computational efficiency, Materials science

Eric Zang, William P. Clements High School, 4200 Elkins Road, Sugar Land, TX 77479, USA.
eric.q.zang@gmail.com

https://doi.org/10.64336/001c.147429

1. Introduction

Environmental films are thin layers of heterogeneous material that accumulate on surfaces exposed to the environment over long periods of time (1). They are formed through a plethora of different processes, including the deposition of airborne particles, condensation of water vapor, and development of biological activity (2). These films commonly form on solar panels, buildings, vehicles, and other infrastructure, but are often seen as innocuous (3). However, their growth has been shown to have important practical consequences. For example, films on photovoltaic cells—the silicon surfaces that make up solar panels—drastically reduce light transmission and lower solar energy efficiency over time (4). On architectural structures, they can accelerate material weathering, discoloration, and degradation (5). Nevertheless, despite their prevalence, the mechanisms behind their growth remain poorly understood.

At the microscopic scale, environmental films consist of nanoscale and microscale particles of extremely diverse origin (6). Their composition, size distribution, and morphology play a pivotal role in determining how these films evolve (3). Fortunately, Atomic Force Microscopy (AFM) is particularly well-suited for this analysis. Since its invention in 1986, AFM has revolutionized surface characterization, allowing both conducting and insulating materials to be imaged with atomic-scale precision, producing 3D topographical maps (7, 8). Atomic Force Microscopes operate by using a piezoelectric scanner to move a sharp cantilever tip (typically <10 nm radius) across the sample surface, while a laser beam reflected onto a photodiode detects tiny tip deflections, which are then converted into high-resolution topographical data (9). AFM has been widely used in many different fields (9–16).

AFM differs from other imaging techniques, such as Optical Microscopy (OM), Scanning Electron Microscopy (SEM), and Scanning Tunneling Microscopy (STM) due to its ability to precisely measure surface topography without requiring conductive coatings or vacuum conditions. OM, SEM, and STM could all in theory provide information about surface morphology, but they each lack the combination of resolution, versatility, and sample compatibility needed for heterogeneous environmental films. OM offers fast, non-destructive imaging, but its diffraction-limited resolution prevents characterization of nanoscale particles (17). SEM provides high-resolution imaging and compositional information, but requires conductive coatings and high-vacuum operation, which can alter delicate surfaces (18). Finally, STM achieves atomic-scale resolution, but it can only be applied to conductive or semi-conductive surfaces and is highly sensitive to contamination (19). AFM's ability to provide nanoscale topographical maps of a wide variety of surfaces makes it ideal for studying environmental films.

A crucial step in the analysis of particles from AFM images is their identification. To output accurate data, algorithms must first correctly recognize particles before bounding and analyzing them. Manual identification of particles is extremely slow, while existing automated methods such as watershed segmentation and gradient-based techniques struggle with mixed particle populations and a variety of imaging artifacts (20). Previously, attempts have been made to overcome these challenges through approaches such as watershed refinements and texture-based segmentation, but they remain sensitive to image noise and tip-induced distortions and importantly, have not been tailored to environmental films (21, 22).

Recently, approaches based on machine learning have shown promise for AFM image analysis, offering improved adaptability and accuracy compared to prior methods (23, 24). However, these techniques often depend on large, curated datasets and significant computational resources, which can limit their accessibility and reproducibility in many applications (25). This creates a gap between highly sophisticated but resource-intensive methods and simple, widely used techniques that remain prone to bias and error. Only algorithms that are lightweight, reproducible, and capable of handling AFM imaging artifacts would be able to bridge the gap.

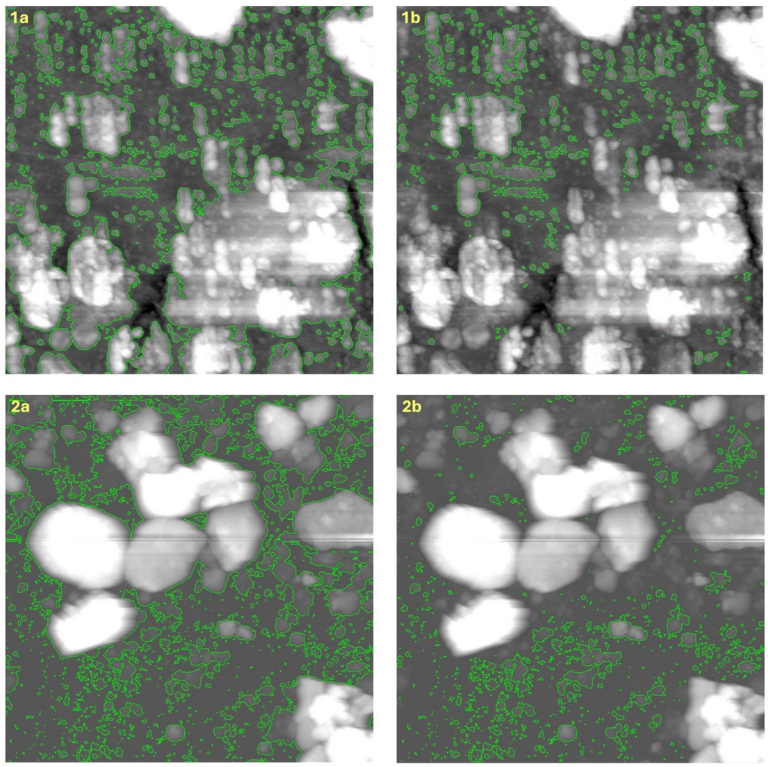
This work presents two fully automated algorithms designed to bridge this gap. The first achieves rapid execution while outperforming standard software, and the second attains the highest overall accuracy through the incorporation of a white top-hat transform. These lightweight, reproducible algorithms eliminate user bias, handle common imaging artifacts, and enable high-throughput AFM particle analysis, enhancing the study of environmental films and expanding the accessibility of accurate nanoscale particle identification.

2. Background

A wide variety of software platforms can be used to analyze AFM data. These include commercial packages such as Igor Pro™ (Sutter Instrument Corp., Novato, CA) and MountainsMap™ (Digital Surf, Besancon, France), open-source tools like Gwyddion (https://doi.org/10.2478%2Fs11534-011-0096-2), and custom analysis pipelines developed in MATLAB or Python (26–28). Among these, Igor Pro™ is particularly common in surface science research due to its straightforward image flattening, masking, and particle analysis functions (1, 3, 29).

However, despite its accessibility, Igor Pro™ contains notable limitations. First, it requires manual adjustment of masking and flattening parameters for each image, often requiring multiple adjustment cycles before the data is ready for analysis. This repetitive process is extremely time-consuming and unreproducible. Secondly, Igor Pro™'s particle detection accuracy is a more fundamental limitation. The software tends to merge large, adjacent particles into a single object, leading to undercounting. Additionally, Igor Pro™ frequently produces false detections in the form of undesired particles, which include; 1] Edge particles, whose incomplete representation leads to misleading properties (such as area, volume, and circularity). Igor Pro™ offers an option to exclude these automatically and 2] Zero-area particles, which have zero values for critical parameters such as area. These occur in significant numbers and must be manually removed.

Unfortunately, Igor Pro™ frequently merges large, clustered particles into a single object at the image boundary, as shown in Figures 1 and 2, causing them to be flagged as edge particles and thus excluded from analysis. Because large particles are inherently much rarer than smaller ones, their removal skews particle size distributions and biases subsequent analysis significantly. This limitation directly motivates the algorithms developed in this work.



Figures 1 and 2. Visualization of Igor Pro™ particle identification and the effect of edge particle removal. Each figure shows a pair of AFM images: (**a**) particles identified by Igor Pro™ before edge particles are removed, and (**b**) particles after edge particle removal. Note the large clusters discarded as edge particles by the Igor Pro™ software. Unless otherwise noted, all AFM images were acquired over 10 × 10 µm scan areas.

3. Methods

3.1 AFM data acquisition

The AFM images used in this work were captured using an Asylum Research Atomic Force Microscope (Oxford Instruments plc, UK) of the Tivanski Lab at the University of Iowa. They span several dozen outdoor environmental films collected from three distinct regions of the United States, each characterized by markedly different environments and climate, and encompass samples deposited on three different substrates across four distinct time scales. The images span root-mean-square roughness (R_q) values from 1.932 to 287.802 nm, contain particle populations varying from sparse to densely aggregated, and reflect the diversity in particle size, contrast, and morphology typical of environmental film samples reported in prior literature (1–3, 6).

All AFM images had dimensions of 10 × 10 µm and were acquired using Igor Pro™ (6.38B01) in alternating contact mode. Imaging parameters were calibrated according to standard Asylum Research procedures. The AFM tip radius was 2 nm, z-calibration was performed using the Asylum Research internal piezo standard, and the setpoint, scan speed, and pixel size were 3.0, 102.80 µm/s, and 19.5 nm, respectively. Images were saved through Igor Pro™ as Igor Binary Wave (.ibw) files, which store data in eight channels. Two channels were used extensively: HeightTrace (HtT) and AmplitudeTrace (AmT), shown in Figure 3. HtT provides the topographical height map of the surface, and AmT enhances contrast and facilitates separation of overlapping particles.

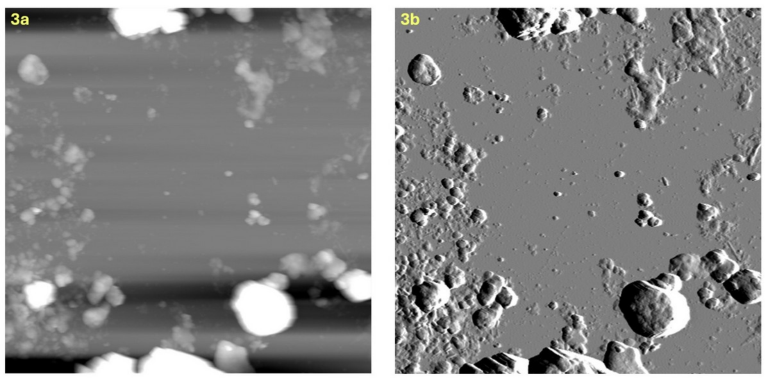


Figure 3. Representative AFM images showing two measurement channels: (**a**) HeightTrace, used for algorithmic particle analysis, and (**b**) AmplitudeTrace used for manual particle verification.

3.2 Novel particle detection algorithms

3.2.1 *White top-hat transform*

Top-hat transforms are morphological operations that enhance small, bright features under an empirically determined parameter while suppressing large-scale background variations (30). They are categorized as either white or black and have been used in a wide variety of studies (31–33).

The white top-hat transform (WTH) is defined as, $WTH(I)=I-(I \circ b)$, where I is the original image and $\circ$ denotes morphological opening with structuring element b (30).

In AFM images, WTH is especially effective in isolating particles in images with rough backgrounds. WTH suppresses uneven surface topography, reduces noise, and improves the accuracy of subsequent thresholding and contour detection down the particle identification line. The effect of the transform on a representative HtT AFM image is shown in Figure 4.

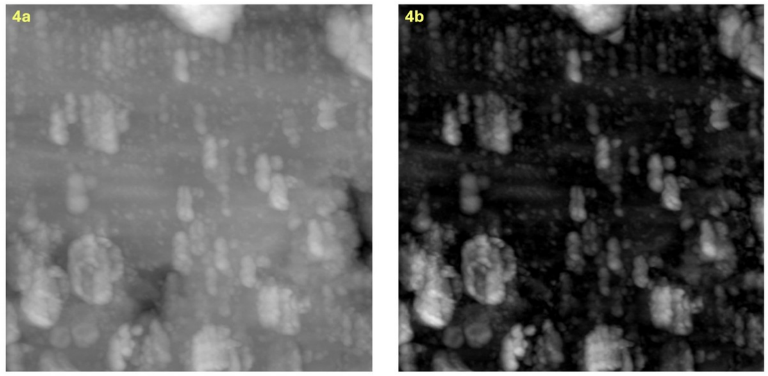


Figure 4. Effect of the white top-hat transform (WTH) on a representative AFM image. **(a)** Original image without WTH. **(b)** Image after WTH, which suppresses uneven surface topography, reduces background noise, and enhances particle contrast.

3.2.2 *Adaptive Gaussian Thresholding*

Adaptive Gaussian thresholding (AGT) is a local thresholding method for converting grayscale images into binary masks (34). AGT differs from traditional Gaussian thresholding because it compares each pixel to a locally defined threshold rather than a single global value. The method has been applied in a diverse range of studies (34–36). For each pixel *p*, the threshold is given by

$$T(p) = \frac{\sum_{q \in N(p)} I(q) \cdot w(q)}{\sum_{q \in N(p)} w(q)} - C$$

where *N(p)* is the local neighborhood of pixel *p*, *I(q)* is the intensity of pixel *q*, *w(q)* is a Gaussian weighting function, and *C* is a constant offset (37).

AGT is used in this paper to account for spatially varying backgrounds, ensuring proficient segmentation even in rough images. By adapting the threshold to local intensity variations, AGT minimizes the risk of particle loss in shadowed regions and prevents false detections in brighter areas, providing a more reliable foundation for subsequent algorithm steps.

3.2.3 *Contour-based particle identification*

A contour-tracing algorithm was used to identify candidate particles following binarization. Each contour consists of a closed curve deriving from connected pixels in the AGT binary mask. There are two purposes of contour extraction: (i) it defines particles geometrically, and (ii) it allows for particle visualization through overlaying on top of the AFM height image. The resulting image overlay can thus be used to analyze particle detection accuracy, as shown in Figure 5.

A final particle identification algorithm was then designed to discard both edge and zero-area particles. Importantly, it improves on Igor Pro™'s inability to discard particles with zero area.

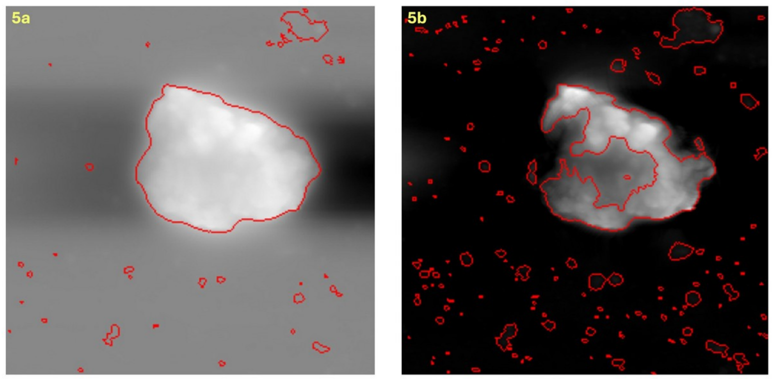


Figure 5. Particle contour extraction of a representative AFM image, shown **(a)** without the white top-hat transform (WTH) and **(b)** with the WTH. Closed curves are traced from connected pixels in the AGT binary mask and overlaid on the HeightTrace image.

3.2.4 *Parameter selection*

The radius of the structuring element used in the WTH, as well as the neighborhood size and offset parameters in the AGT, were selected using a systematic empirical procedure rather than ad hoc or image-specific tuning. To ensure broad applicability, parameter selection was performed using a diverse subset of 24 AFM images drawn from the full set of environmental films collected across three U.S. regions, three substrates, and four time scales. The nanometer R_q values for this subset had mean 92.384, median 67.215, standard deviation 78.946, minimum 2.431, and maximum 265.307, reflecting substantial heterogeneity in topography and particle morphology.

For each candidate parameter value, segmentation outputs were evaluated using three reproducible criteria: (i) suppression of large-scale background variation without distorting particle boundaries, (ii) preservation of both small and large particles, and (iii) minimization of over-segmentation in shadowed or low-slope regions. Comparison was performed using AmT-channel overlays but did not rely on knowledge of particle ground truth.

Disk structuring element radii from 20–80 px were tested for the WTH. A radius of 45 px provided the best tradeoff: smaller radii fragmented large particles and amplified noise, while larger radii flattened genuine broad features. For the AGT, neighborhood sizes from 21–101 px and offsets from –4.5 to +2.5 were evaluated; a 71 px neighborhood ensured adequate local statistics across both smooth and rough surfaces, and an offset of –2 avoided suppressing small particles while preventing false positives.

3.2.5 *Algorithm definitions*

This work presents two novel algorithms, labeled Algorithm A and Algorithm B. Each algorithm uses the HeightTrace (HtT) channel of Igor Binary Wave files. The difference between the two algorithms is the white top-hat transform within Algorithm B. The steps of each algorithm are shown in Table 1.

Table 1. Comparison of the image-processing steps used in Algorithms A and B. While both algorithms share the same preparation, thresholding, feature extraction, and filtering steps, Algorithm B incorporates an additional white top-hat transform during preprocessing.

Step	**Algorithm A**	**Algorithm B**
Preparation	HtT channel	HtT channel
Preprocessing	None	White top-hat transform *radius = 45 px*
Thresholding	Adaptive Gaussian thresholding *neighborhood size = 71 px, offset = -2*	Adaptive Gaussian thresholding *neighborhood size = 71 px, offset = -2*
Feature Extraction	Contour-based particle identification	Contour-based particle identification
Particle Filtering	Edge and zero-area particles discarded	Edge and zero-area particles discarded
Output	Particle count and visualization	Particle count and visualization

All image processing and analysis steps were implemented using Python (3.13.6), with random-number-dependent steps employing a fixed seed value of 42. Computations were performed on an AMD Ryzen 7 7730U running Windows 11 with 16 GB RAM; the Igor2 (0.5.12), NumPy (2.2.6), matplotlib (3.10.5), scikit-image (0.25.2), and OpenCV (4.12.0) libraries were used. Table 2 describes each library's role.

Table 2. Software libraries used in the implementation of Algorithms A and B and their respective functions.

Library	**Role**
Igor2	Extracted HtT channel and corresponding scale factors from Igor Binary Wave files
NumPy	Performed numerical operations and coordinate transformations
matplotlib	Produced visualizations of outputted images
scikit-image	Conducted WTH
OpenCV	Performed AGT and produced particle-bounding contours

3.3 Performance evaluation

To accurately evaluate the performances of Algorithm A, Algorithm B, and Igor Pro™, ground-truth particles were manually identified. Out of a repository of 84 AFM images, 16 were selected using simple random sampling to construct the ground-truth dataset. For the 16 images, the nanometer R_q values of the images had a mean of 80.951, a median of 55.626, a standard deviation of 72.558, and a range of 3.700 to 238.848, indicating that the selected subset captured the variability present in the full 84-image repository.

All manual particle identifications used the AmT channel of .ibw files for high contrast. Manual particle identification was performed in four 2 × 2 µm subregions within each 10 × 10 µm image, yielding 64 annotated regions across 16 images. From these regions, 2014 particles were labeled, corresponding to an average of 31.5 particles per subregion. The average particle number density was 7.87 ± 4.39 particles per square nanometer. This particle density is consistent with prior reports of environmental films imaged under similar conditions (2).

Although particle density varied across samples, sampling multiple subregions across all 16 images ensured representation of both sparse and dense regions. Cross-validation was not performed because manual particle labeling is time-intensive and cannot be scaled to multiple independent resampling rounds. Figure 6 illustrates the schematic for manual particle identification: true, identified particles are bound by ellipses for simplicity.

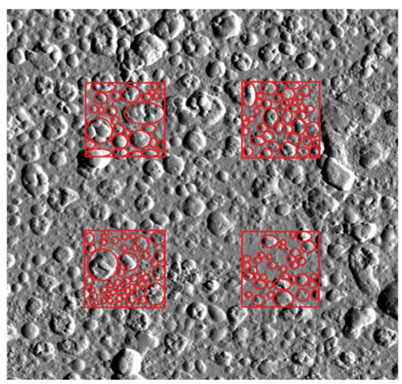


Figure 6. Schematic of manual particle identification using the AmplitudeTrace (AmT) channel. Particles were manually marked within designated 2 × 2 µm regions for subsequent comparison with algorithmic detection.

Subsequently, the images were analyzes using Algorithm A, Algorithm B, and Igor Pro™. For each image, manual particle count (N_{man}) was recorded along with runtime and algorithm particle count (N_{alg}) for each algorithm. Additionally, Igor Pro™'s zero-area particle count (N_0) was measured, and the root-mean-square roughness (R_q) was obtained from Igor Pro™. Particles were classified as true positives (TP), false positives (FP), or false negatives (FN) by comparing algorithm results to the manual ground truth. Performance metrics were then computed.

$$Precision = \frac{TP}{TP+FP}, \; Recall = \frac{TP}{TP+FN}$$

$$F1\,Score \;=\; 2\times\frac{Precision\times Recall}{Precision+Recall},\;\; Zero-area\;Rate \;=\; \frac{N_0}{N_{alg}}$$

Precision measures detection quality, recall measures detection quantity, and the F1 score provides a balanced single metric of performance. The zero-area rate quantifies invalid particle detections.

A one-way repeated-measures (RM) ANOVA was then applied to assess whether the metrics computed by the different algorithms were different. Then, post hoc pairwise comparisons were conducted with Bonferroni correction to account for multiple comparisons. All statistical analyses were performed using the Python library *pingouin*, and effect sizes were computed using Hedges' g. Performance was compared across Algorithm A, Algorithm B, and Igor Pro™.

4. Results

Overall, Algorithms A and B outperformed Igor Pro™, with Algorithm A demonstrating faster processing and Algorithm B achieving higher detection accuracy. Table 3 summarizes the performance of each algorithm across these metrics. Figure 7 further shows representative outputs from Algorithm A, and Algorithm B, and Igor Pro™. The following subsections describe each key metric in detail.

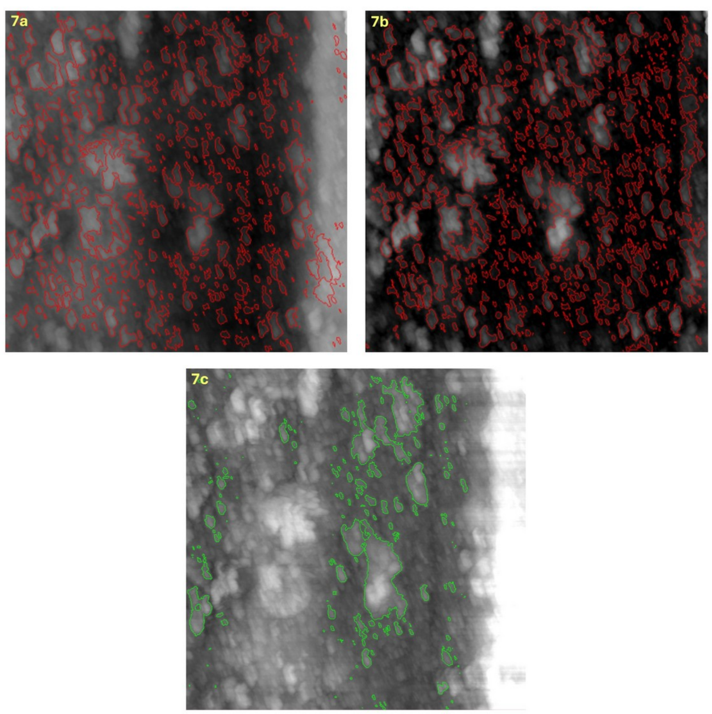


Figure 7. Representative particle identification outputs from (**a**) Algorithm A, (**b**) Algorithm B, and (**c**) Igor Pro™, illustrating differences in detection.

Table 3. Performance of Algorithms A and B compared with Igor Pro™ across five evaluation metrics. Reported values represent the mean ± standard deviation.

Metric	**Algorithm A**	**Algorithm B**	**Igor Pro™**
Runtime (s)	1.49 ± 0.01	3.22 ± 0.24	43.6 ± 16.5
Precision	0.78 ± 0.15	0.77 ± 0.14	0.60 ± 0.22
Recall	0.54 ± 0.18	0.61 ± 0.14	0.46 ± 0.26
F1 Score	0.60 ± 0.16	0.66 ± 0.10	0.46 ± 0.19
Zero-area Rate	N/A	N/A	0.34 ± 0.14

4.1 Runtime

Figure 8 shows a scatterplot of runtime for all three algorithms.

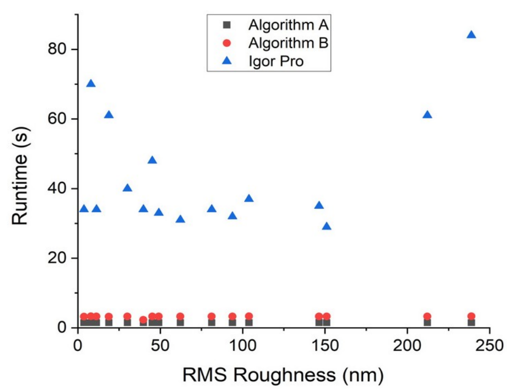


Figure 8. Scatterplot of runtime for Algorithm A, Algorithm B, and Igor Pro™ as a function of RMS roughness.

RM ANOVA revealed a statistically significant effect of algorithm on runtime, $F(2,30) = 100.49$, $p < 0.001$, $\eta_G^2 = 0.817$, with a violation of sphericity (Greenhouse–Geisser $\varepsilon = 0.50$). Post hoc tests showed that Algorithm A performed significantly faster than both Algorithm B ($p < 0.001$, $g = 9.91$) and Igor Pro™ ($p < 0.001$, $g = 3.52$), and Algorithm B performed significantly faster than Igor Pro™ ($p < 0.001$, $g = 3.38$).

Algorithm A was more than twice as fast as Algorithm B and 25 times faster than Igor Pro™, corresponding to a 96.6% reduction. This disparity is due to Igor Pro™'s manual masking and flattening cycles, whereas Algorithms A and B are fully automated. Additionally, Algorithm B implements a WTH that Algorithm A does not, slightly increasing runtime. Algorithm A's speed makes it well-suited for high-throughput imaging experiments where large datasets must be processed rapidly. This distinction is critical in practical settings, as datasets containing hundreds of AFM images can be processed in a fraction of the time, making real-time analysis feasible in the laboratory. In addition, lower runtime requirements reduce computational costs and broaden accessibility to researchers using standard hardware rather than specialized computing resources, facilitating broader adoption of AFM-based film analysis.

Afterward, a quadratic regression was conducted to examine the relationship between RMS roughness and Igor Pro™ runtime, as shown in Figure 9. To mitigate multicollinearity between R_q and R_q^2, R_q was mean-centered prior to model fitting. The regression was statistically significant, $F(2,13) = 11.43$, $p = 0.0014$, and explained approximately 64% of the variance in runtime. Variance Inflation Factors for both predictors were less than 2, confirming minimal multicollinearity. The quadratic term remained significant ($p = 0.001$), indicating a U-shaped dependence of runtime on surface roughness.

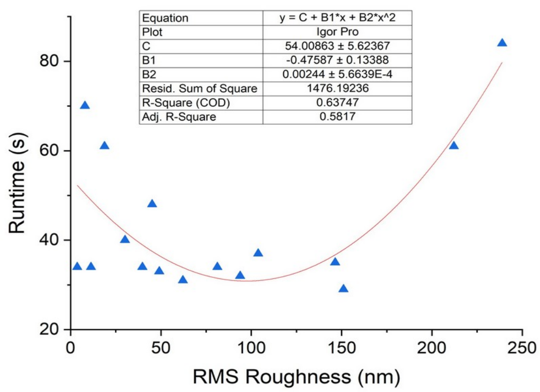


Figure 9. Quadratic regression of Igor Pro™ runtime as a function of RMS roughness.

The minimum runtime of ~ 30 seconds was predicted near $R_q = 100$ nm. This trend likely reflects two competing effects. For smoother images, the higher particle density—mainly due to an increased number of smaller particles—requires extensive manual masking and flattening, increasing runtime. In contrast, for very rough images, the irregular surface complicates the masking and flattening process, elevating runtime as well.

4.2 Precision

Figure 10 shows box-and-whisker plots of precision for all three algorithms. RM ANOVA revealed a statistically significant effect of algorithm on precision, $F(2,30) = 7.72$, $p = 0.002$, $\eta_G^2 = 0.197$. Post hoc tests showed that Igor Pro™ was significantly outperformed by both Algorithm A ($p = 0.042$, $g = 0.94$) and Algorithm B ($p = 0.040$, $g = 0.92$), and Algorithms A and B performed identically ($p = 1.00$).

Igor Pro™ displayed a considerably larger false positive rate, primarily due to its inability to discard zero-area particles—a limitation addressed by both automated algorithms. Higher precision in the automated algorithms ensures that fewer false positives are recorded, reducing particle overcounting and improving the reliability of particle identification.

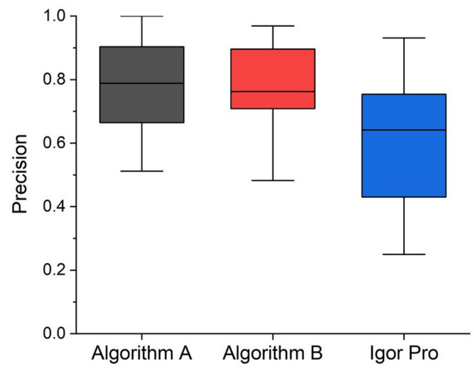


Figure 10. Box-and-whisker plot of precision for Algorithm A, Algorithm B, and Igor Pro™. Default plot parameters were used.

4.3 Recall

Figure 11 shows box-and-whisker plots of recall for all three algorithms. RM ANOVA revealed a statistically significant effect of algorithm on recall, $F(2,30) = 4.14$, $p = 0.026$, $\eta_G^2 = 0.087$, with a violation of sphericity (Greenhouse–Geisser $\varepsilon = 0.50$). Post hoc tests showed no significant differences between any pairwise comparisons after Bonferroni correction, with all $p > 0.05$, though Algorithm B trended higher than both Algorithm A and Igor Pro™.

The lower recall of Igor Pro™ stems largely from its tendency to merge adjacent particles, resulting in missed detections. Higher recall in the automated algorithms ensures more accurate detection of true particles, which is especially important in dense or low-contrast images where undercounting could bias results.

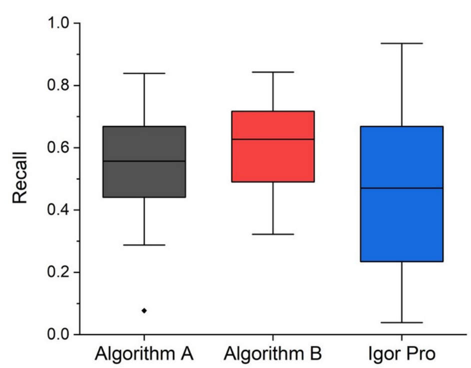


Figure 11. Box-and-whisker plot of recall for Algorithm A, Algorithm B, and Igor Pro™. Default plot parameters were used.

4.4 F1 Score

Figure 12 shows box-and-whisker plots of F1 score for all three algorithms. RM ANOVA revealed a statistically significant effect of algorithm on F1 score, $F(2,30) = 12.99$, $p < 0.001$, $\eta_G^2 = 0.238$. Post hoc tests showed that Igor Pro™ was significantly outperformed by both Algorithm A ($p = 0.021$, $g = 0.78$) and Algorithm B ($p < 0.001$, $g = 1.29$), and Algorithm B was marginally more effective than Algorithm A ($p = 0.268$).

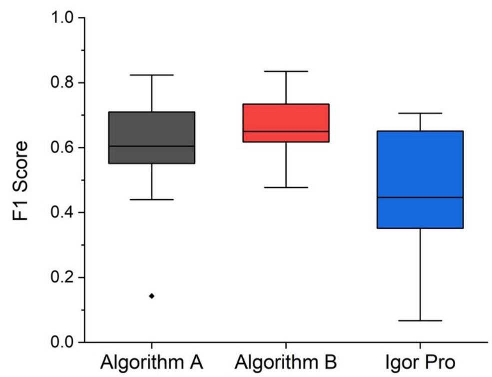


Figure 12. Box-and-whisker plot of F1 score for Algorithm A, Algorithm B, and Igor Pro™. Default plot parameters were used.

Algorithm B's F1 score of 0.66—43.5% higher than Igor Pro™'s 0.46—demonstrated its balanced precision and recall while also exemplifying the value of its WTH. Also, Algorithm A—employing only AGT and particle filtering—already provided meaningful improvements over Igor Pro™. In contexts where both undercounting and false positives are detrimental, such as when constructing particle-size distributions or quantifying surface coverage, a balanced performance is essential. By achieving higher F1-scores, Algorithm B reduced result uncertainty, leading to greater confidence in experimental conclusions. This makes the method particularly suitable for comparative studies where subtle differences between conditions could otherwise be masked by algorithmic variability.

4.5 Zero-area rate

No relationship was found between Igor Pro™'s zero-area rate and the R_q of images, indicating that the detection of zero-area particles is a fundamental limitation of the software, rather than an artifact of image quality. On average 34% ± 14% of Igor Pro™'s detections had zero reported area, inflating apparent particle counts, increasing false-positive rates, and necessitating manual removal. In contrast, Algorithms A and B explicitly included contour-based area checks and filtering rules that removed such artifacts automatically, which directly reduced false positives and yielded higher precision. Beyond count accuracy, eliminating zero-area artifacts stabilizes downstream particle property measurements and reduces between-image variance, an important improvement for studies that compare particle statistics across samples or time.

Taken together, these results indicate that Algorithms A and B were more streamlined, efficient, and reliable than Igor Pro™. The automated algorithms exhibited a lower standard deviation than Igor Pro™ across all measured metrics, indicating more stable and reproducible performance. Whereas Algorithm A is optimal when speed is the primary concern, Algorithm B provided the most accurate and balanced performance for comprehensive particle analysis. The consistent superiority of Algorithms A and B across precision, recall, and F1 metrics demonstrated their robustness under varied imaging conditions and across heterogeneous particle populations. Importantly, the results provided generalizable performance for a wide range of AFM datasets, strengthening the reproducibility of AFM-based research and nanoparticle analysis.

5. Discussion

Several prior approaches have addressed AFM particle detection and related segmentation problems. Many AFM image analyses use classical image-processing methods such as watershed segmentation and gradient-based techniques (21, 22). These methods are accessible but often fail with spatially varying backgrounds and the broad particle size distributions typical of environmental films. In addition, morphological transforms such as the white top-hat transform have been adapted in diverse imaging contexts to enhance localized features against uneven surfaces (30, 31). This work employed the WTH to directly address uneven background challenges in AFM images. Hessian-based detection algorithms have also been used for AFM particle detection (20), providing strong accuracy for well-defined biological particles. However, such methods require careful tuning of detection scales and struggle with particle heterogeneity or low-contrast backgrounds. Additionally, machine learning and deep learning approaches have gained traction in AFM and related nanoscale imaging (23, 24), demonstrating the potential of neural networks; however, their reliance on large, labeled datasets and high-performance computing makes them less accessible. Finally, established software platforms remain widely used, including Gwyddion and Igor Pro™. Both provide powerful general AFM toolsets but require extensive manual input and produce merged clusters or zero-area detections.

The present contribution differs by implementing a combination of morphological preprocessing and local thresholding while avoiding deep-learning dependence and user intervention. The improvements demonstrated via Algorithms A and B have several implications for AFM-based studies of environmental films and for nanoparticle characterization more broadly. The combination of automated preprocessing (WTH), adaptive local thresholding (AGT), and contour-based filtering substantially reduces manual workload. Removing manual masking and flattening cycles yields time savings and improves reproducibility across operators and datasets. The runtime advantage of Algorithm A makes high-throughput studies practicable even on modest hardware. By avoiding particle clusters and zero-area artifacts, the presented methods provide more reliable size distributions and shape statistics, reducing bias in downstream physical interpretations. Because the algorithms are training-free and implemented using standard, open Python libraries, they are readily shareable and extensible. In sum, the methods enable reproducible, efficient particle analysis and bridge a practical gap between simple but error-prone tools and high-accuracy, resource-intensive machine-learning systems.

Although this work presents two novel algorithms that outperform Igor Pro™, several limitations remain. Standard evaluation metrics such as precision, recall, and F1 score may not fully capture errors in particle boundary shape. While Algorithms A and B excelled at avoiding the merging of large, clustered particles—a significant improvement over Igor Pro™—they also showed a tendency to over-split large particles, most notably in Algorithm B due to the WTH flattening features larger than the structuring element. Algorithm A, which did not use a WTH, remained more susceptible to image artifacts and shadowed regions.

Algorithm performance depended on empirical parameter choices, including the WTH radius and the AGT neighborhood size and offset. However, these parameters were selected using a deliberately heterogeneous calibration dataset drawn from three climatically distinct U.S. regions, three substrate types per region, and four sampling time scales, spanning a reasonably large range of roughness, contrast, and particle-morphology variability in the full collection. This variability ensured that parameter selection was not narrowly tailored to a single film type or imaging condition and instead stress-tested the algorithms across heterogeneous surfaces. Since the parameter-selection dataset was intentionally constructed to include highly diverse films, the resulting parameters are expected to generalize well. Even when a new dataset requires adjustment, the calibration procedure is rapid and requires only a one-time sweep over a representative subset, which is far less labor-intensive than tuning parameters for each image individually.

Finally, at present, no established or validated simulator exists for generating realistic AFM AmplitudeTrace or HeightTrace images of heterogeneous environmental films with known particle-scale ground truth. Developing a physically accurate simulator lies outside the scope of this study; nevertheless, synthetic-image validation represents an important future direction once a reliable simulation framework becomes available.

6. Conclusion

This work presented and evaluated two automated algorithms for particle detection in AFM images of environmental films. Both Algorithm A and Algorithm B significantly outperformed the widely used Igor Pro™ software in accuracy, efficiency, and consistency, particularly in avoiding the merging of distinct, adjacent particles and the discarding of zero-area particles. Algorithm A exhibited the fastest runtime while outperforming Igor Pro™ in accuracy, and Algorithm B demonstrated the finest overall performance. Further work is underway to extract quantitative particle properties such as height, volume, and circularity to enable detailed characterization of environmental films. In addition, approaches such as dualthreshold methods may provide a more seamless integration of small and large particles, potentially improving detection accuracy of heterogeneous particle distributions. Overall, by removing the need for manual adjustments and raising segmentation accuracy, the methods developed here have the potential to accelerate research on environmental films and support more systematic investigations of particle morphology and dynamics. In the long term, these techniques may expand the use of AFM in particle analysis, where precise and efficient segmentation is critical.

Acknowledgments
The author would like to thank Professor Alexei Tivanski (Department of Chemistry, University of Iowa) along with his graduate students Arman Chowdhury and Ruwini Weerasinghe for generously allowing the use of AFM images from their lab.

Code availability
The algorithms developed in this work are publicly available at https://github.com/erzang/afm-particle-detection.git.

7. References

1. DeYoung, J. L.; Holyoake, E. A.; Shaw, S. K. What Are the Differences between Two Environmental Films Sampled 1 Km Apart? *ACS Earth and Space Chemistry* **2021**, *5* (12), 3407–3413. https://doi.org/10.1021/acsearthspacechem.1c00257.

2. DeYoung, J. L.; Uchechukwu G. A.; Cheng, Z.; China, S.; Vandergrift, G. W.; Anderton, C. R.; et al. Chemical Differences in Environmental Films Collected on Surfaces with Different Hydrophilicity. *ACS Earth and Space Chemistry* **2024**, *8* (12), 2411–2419. https://doi.org/10.1021/acsearthspacechem.4c00170.

3. Grant, J. S.; Zhu, Z.; Anderton, C. R.; Shaw, S. K. Physical and Chemical Morphology of Passively Sampled Environmental Films. *ACS Earth and Space Chemistry* **2019**, *3* (2), 305–313. https://doi.org/10.1021/acsearthspacechem.8b00158.

4. Yeom, S.; An, J.; Kang, H.; Hong, T. Analyzing the Optimal Visible Light Transmittance of Thin-Film Photovoltaic Using Experiment with Virtual Reality and Economic Assessment. *Energy and Buildings* **2023**, *296*, 113380. https://doi.org/10.1016/j.enbuild.2023.113380.

5. Kang, Y.; Jo, H. H.; Kim, S. Effects of UV Degradation on Building Materials with Emphasis on Microplastic Generation Potential. *Journal of Hazardous Materials* **2024**, *483*, 136521. https://doi.org/10.1016/j.jhazmat.2024.136521.

6. DeYoung, J. L.; Shaw, S. K. Host Surface Orientation Impacts Environmental Film Accumulations. *Chemosphere* **2022**, *307*, 135823. https://doi.org/10.1016/j.chemosphere.2022.135823.

7. Binnig, G.; Quate, C. F.; Gerber, Ch. Atomic Force Microscope. *Physical Review Letters* **1986**, *56* (9), 930–933. https://doi.org/10.1103/physrevlett.56.930.

8. Alldritt, B.; Hapala, P.; Oinonen, N.; Urtev, F.; Krejci, O.; Canova, F. F.; et al. Automated Structure Discovery in Atomic Force Microscopy. *Science Advances* **2020**, *6* (9), eaay6913. https://doi.org/10.1126/sciadv.aay6913.

9. Müller, D. J.; Dufrêne, Y. F. Atomic Force Microscopy: A Nanoscopic Neighborhood on the Cell Surface. *Trends in Cell Biology* **2011**, *21* (8), 461–469. https://doi.org/10.1016/j.tcb.2011.04.008.

10. Muller, D. J. AFM: A Nanotool in Membrane Biology. *Biochemistry* **2008**, *47* (31), 7986–7998. https://doi.org/10.1021/bi800753x.

11. Stylianou, A.; Kontomaris, S.-V.; Grant, C.; Alexandratou, E. Atomic Force Microscopy on Biological Materials Related to Pathological Conditions. *Scanning* **2019**, *2019*. https://doi.org/10.1155/2019/8452851.

12. Marrese, M.; Guarino, V.; Ambrosio, L. Atomic Force Microscopy: A Powerful Tool to Address Scaffold Design in Tissue Engineering. *Journal of Functional Biomaterials* **2017**, *8* (1). https://doi.org/10.3390/jfb8010007.

13. Yang, H.; Wang, Y.; Lai, S.; An, H.; Li, Y.; Chen, F. Application of Atomic Force Microscopy as a Nanotechnology Tool in Food Science. *Journal of Food Science* **2007**, *72* (4), R65-75. https://doi.org/10.1111/j.1750-3841.2007.00346.x.

14. Joshua, A. M.; Gan, C.; Ee Von Lau. Soft Matter Analysis via Atomic Force Microscopy (AFM): A Review. *Applied Surface Science Advances* **2023**, *17*, 100448. https://doi.org/10.1016/j.apsadv.2023.100448.

15. Ramirez-Aguilar, K. A.; Lehmpuhl, D. W.; Michel, A. E.; Birks, J. W.; Rowlen, K. L. Atomic Force Microscopy for the Analysis of Environmental Particles. *Ultramicroscopy* **1999**, *77* (3-4), 187–194. https://doi.org/10.1016/S0304-3991(99)00048-0.

16. Hiesgen, R.; Morawietz, T.; Handl, M.; Corasaniti, M.; Friedrich, K. A. Atomic Force Microscopy on Cross Sections of Fuel Cell Membranes, Electrodes, and Membrane Electrode Assemblies. *Electrochimica Acta* **2015**, *162*, 86–99. https://doi.org/10.1016/j.electacta.2014.11.122.

17. Balasubramanian, H.; Hobson, C. M.; Chew, T.-L.; Aaron, J. S. Imagining the Future of Optical Microscopy: Everything, Everywhere, All at Once. *Communications Biology* **2023**, *6* (1), 1–12. https://doi.org/10.1038/s42003-023-05468-9.

18. Vernon-Parry, K. D. Scanning Electron Microscopy: An Introduction. *III-Vs Review* **2000**, *13* (4), 40–44. https://doi.org/10.1016/s0961-1290(00)80006-x.

19. Hansma, P. K.; Tersoff, J. Scanning Tunneling Microscopy. *Journal of Applied Physics* **1987**, *61* (2), R1–R24. https://doi.org/10.1063/1.338189.

20. Marsh, B. P.; Chada, N.; Sanganna Gari, R. R.; Sigdel, K. P.; King, G. M. The Hessian Blob Algorithm: Precise Particle Detection in Atomic Force Microscopy Imagery. *Scientific Reports* **2018**, *8* (1), 978. https://doi.org/10.1038/s41598-018-19379-x.

21. Zaitoun, N. M.; Aqel, M. J. Survey on Image Segmentation Techniques. *Procedia Computer Science* **2015**, *65*, 797–806. https://doi.org/10.1016/j.procs.2015.09.027.

22. Wu, Y.; Li, Q. The Algorithm of Watershed Color Image Segmentation Based on Morphological Gradient. *Sensors* **2022**, *22* (21), 8202. https://doi.org/10.3390/s22218202.

23. Ziatdinov, M.; Dyck, O.; Maksov, A.; Li, X.; Sang, X.; Xiao, K.; Unocic, R. R.; Vasudevan, R. K.; Jesse, S.; Kalinin, S. V. Deep Learning of Atomically Resolved Scanning Transmission Electron Microscopy Images: Chemical Identification and Tracking Local Transformations. **2017**, *11* (12), 12742–12752. https://doi.org/10.1021/acsnano.7b07504.

24. Ovchinnikova, O. S.; Tsai, W.-Y.; Korolkov, V. V.; Balke, N.; Kalinin, S. V. Machine Learning-Based Multidomain Processing for Texture-Based Image Segmentation and Analysis. **2020**, *116* (4), 044103. https://doi.org/10.1063/1.5135328.

25. Sokolov, I. On Machine Learning Analysis of Atomic Force Microscopy Images for Image Classification, Sample Surface Recognition. *Physical Chemistry Chemical Physics* **2024**, *26* (15), 11263–11270. https://doi.org/10.1039/d3cp05673b.

26. Nečas, D.; Klapetek, P. Gwyddion: An Open-Source Software for SPM Data Analysis. *Open Physics* **2012**, *10* (1). https://doi.org/10.2478/s11534-011-0096-2.

27. Ţălu, Ş.; Bramowicz, M.; Kulesza, S.; Pignatelli, F.; Salerno, M. Surface Morphology Analysis of Composite Thin Films Based on Titanium-Dioxide Nanoparticles. *Acta Physica Polonica A* **2017**, *131* (6), 1529–1533. https://doi.org/10.12693/aphyspola.131.1529.

28. Heath, G. R.; Micklethwaite, E.; Storer, T. M. NanoLocz: Image Analysis Platform for AFM, High-Speed AFM, and Localization AFM. *Small methods* **2024**. https://doi.org/10.1002/smtd.202301766.

29. Miyashita, S.-I.; Sagane, Y.; Uchino, T.; Karatsu, S.; Hosoya, K.; Huang, I.; et al. Atomic Force Microscopic Image Data of Botulinum Neurotoxin Complexes with Different Molecular Sizes. *Data in Brief* **2019**, *25*, 104193. https://doi.org/10.1016/j.dib.2019.104193.

30. Bai, X.; Zhou, F.; Liu, Z.; Xue, B.; Jin, T. Analysis of Top-Hat Selection Transformation and Some Modified Top-Hat Transformations. *Optik* **2012**, *123* (10), 892–895. https://doi.org/10.1016/j.ijleo.2011.07.001.

31. Guerroudji, M. A.; Ameur, Z. A New Approach for the Detection of Mammary Calcifications by Using the White Top-Hat Transform and Thresholding of Otsu. *Optik* **2015**, *127* (3), 1251–1259. https://doi.org/10.1016/j.ijleo.2015.10.192.

32. Wu, B.; Zhang, J.; Zhao, Y. A Novel Method to Extract Narrow Water Using a Top-Hat White Transform Enhancement Technique. *Journal of the Indian Society of Remote Sensing* **2018**, *47* (3), 391–400. https://doi.org/10.1007/s12524-018-0910-z.

33. Li, Y.; Yong, B.; Wu, H.; An, R.; Xu, H.; Xu, J.; He, Q. Filtering Airborne Lidar Data by Modified White Top-Hat Transform with Directional Edge Constraints. *Photogrammetric Engineering & Remote Sensing* **2014**, *80* (2), 133–141. https://doi.org/10.14358/pers.80.2.133-141.

34. Golilarz, N. A.; Demirel, H.; Gao, H. Adaptive Generalized Gaussian Distribution Oriented Thresholding Function for Image De-Noising. *International Journal of Advanced Computer Science and Applications* **2019**, *10* (2). https://doi.org/10.14569/ijacsa.2019.0100202.

35. Amiri Golilarz, N.; Gao, H.; Demirel, H. Satellite Image De-Noising with Harris Hawks Meta Heuristic Optimization Algorithm and Improved Adaptive Generalized Gaussian

Distribution Threshold Function. *IEEE Access* **2019**, *7*, 57459–57468. https://doi.org/10.1109/access.2019.2914101.

36. Rehman, N. A.; Haroon, F. Adaptive Gaussian and Double Thresholding for Contour Detection and Character Recognition of Two-Dimensional Area Using Computer Vision. **2023**. https://doi.org/10.3390/engproc2023032023.

37. Lu, X.; Zhang, Y. An Improved Canny Detection Method for Detecting Human Flexibility; Academic Press, **2021**; pp. 207–234. https://doi.org/10.1016/B978-0-12-821187-8.00008-3.