



Deep-learning recognition and tracking of individual nanotubes in low-contrast microscopy videos

Vladimir Pimonov*, Said Tahir and Vincent Jourdain

Full Research Paper

Open Access

Address:
Laboratoire Charles Coulomb, Univ Montpellier, CNRS, Montpellier,
France

Email:
Vladimir Pimonov* - vladimir.v.pimonov@gmail.com

* Corresponding author

Keywords:
carbon nanotubes growth kinetics; deep learning; in situ microscopy;
object recognition and tracking; polarization microscopy

Beilstein J. Nanotechnol. **2025**, *16*, 1316–1324.
<https://doi.org/10.3762/bjnano.16.96>

Received: 26 May 2025
Accepted: 29 July 2025
Published: 13 August 2025

Associate Editor: S. Giordani



© 2025 Pimonov et al.; licensee Beilstein-Institut.
License and terms: see end of document.

Abstract

This study addresses the challenge of analyzing the growth kinetics of carbon nanotubes using in situ homodyne polarization microscopy (HPM) by developing an automated deep learning (DL) approach. A Mask-RCNN architecture, enhanced with a ResNet-FPN backbone, was employed to recognize and track individual nanotubes in microscopy videos, significantly improving the efficiency and reproducibility of kinetic data extraction. The method involves a series of video processing steps to enhance contrast and used differential treatment techniques to manage low signal and fast kinetics. The DL model demonstrates consistency with manual measurements and increased throughput, laying the foundation for statistical studies of nanotube growth. The approach can be adapted for other types of in situ microscopy studies, emphasizing the importance of automation in high-throughput data acquisition for research on individual nano-objects.

Introduction

Carbon nanotubes (CNTs), discovered over three decades ago, continue to present unresolved questions and challenges. Their exceptional properties, both theoretically [1,2] and experimentally demonstrated [3], make them desirable for electronic and optical devices. However, the widespread application of CNTs is hindered by the lack of control over their structure during growth. Therefore, developing highly selective synthesis methods is crucial for advancing CNT-based devices. This requires a deep understanding of the relationship between nano-

tube structure and selectivity, particularly kinetic selectivity. To address this, we developed a method based on in situ homodyne polarization microscopy (HPM), which is highly sensitive and can detect changes in optical absorption caused by a single carbon nanotube. The technique allows for imaging tens to hundreds of individual carbon nanotubes during growth at up to 40 frames per second [4]. However, the vast amount of information generated requires meticulous and time-consuming analysis to extract kinetic data.

This challenge, common in imaging-related fields, can be addressed through advances in artificial intelligence (AI), particularly in computer vision (CV). Early attempts to automate visual information processing began over three decades ago with one of the first convolutional neural networks recognizing handwritten zip codes on postage envelopes [5]. Since then, CV algorithms have significantly progressed, finding numerous applications in scientific research. Deep learning models have been used to identify two-dimensional materials in microscopic images [6], characterize mineral composition in scanning electron microscopy (SEM) samples [7], and determine nanotube chirality from transmission electron microscopy images [8]. Here, we report on using a deep learning algorithm to automatically recognize and track individual carbon nanotubes during growth, captured in polarization microscopy video sequences.

Materials and Methods

CNT growth

The experimental setup was previously described in [4,9]. In short, horizontally aligned carbon nanotubes (HA-CNTs) were synthesized inside a miniature chemical vapor deposition (CVD) cell with an optical window (Linkam TS1500). ST-cut quartz and iron nanoparticles served as substrate and catalyst, respectively. Ethanol and argon were, respectively, used as carbon precursor and carrier gas. Oxygen and water sensors monitored gas-phase contaminants at the outlet line.

In situ microscopy

Nanotube growth was imaged in situ using a custom-built optical setup for homodyne polarization microscopy. A super-continuum source (Fianium SC-400-4, 2 ps pulses, 40 MHz, spectral range 400–2000 nm) provided white light excitation across the visible spectrum. Two crossed polarizers were employed, with a polarizer and analyzer used to enhance the scattered field from the nanotubes relative to the stronger reflected field from the substrate. A low-pass optical filter with a cutoff wavelength of 700 nm was placed after the analyzer to filter out blackbody radiation generated by the heating crucible. A long-distance objective (Nikon Plan Fluor ELWD 20× 0.45 C L) was used for illumination and collection. Growth process videos were captured using a digital camera (Hamamatsu c11440 ORCA-Flash4.0 LT) with maximum acquisition rate up to 40 frames per second (fps).

Video processing

To extract kinetic data from each nanotube, the raw videos were first processed to enhance contrast. The detailed processing methodology is provided in Supporting Information File 1. In short, after alignment the initial frame rate, typically between 40 and 25 fps, was reduced by averaging frames to one frame per second. This operation boosts further analysis and improves

signal-to-noise ratio without compromising kinetic data quality, as the nanotube growth rate is typically of the order of $0.5 \mu\text{m}\cdot\text{s}^{-1}$ and the localization precision of the optical setup is $0.33 \mu\text{m}$. Frames were then aligned using a template matching algorithm from the OpenCV Python library. Shade correction was applied to compensate for uneven illumination caused by the optics. Residual noise was reduced using fast Fourier transform (FFT) band-pass filtering, and object edge contrast was enhanced using Gaussian difference filtering. Finally, image contrast was optimized through histogram equalization [10].

Deep learning model

The deep learning model and its training are detailed in Supporting Information File 1. In short, an image recognition system was implemented using the PyTorch library [11]. For recognizing nanotubes in videos, we employed the Mask-RCNN architecture [12] with the ResNet-FPN neural network combination as a backbone [13,14], pre-trained on the Microsoft COCO-2017 dataset [15,16]. Training was conducted on graphics processing units provided by Google Colab [17] over 150 epochs with a learning rate of 0.005, decreasing by 10% every 5 epochs. A weight decay of 0.0005 with momentum of 0.9 was used.

Results and Discussion

In situ videos of CNT growth obtained using HPM contain extensive kinetic information. However, the low contrast of raw videos (Figure 1a) is inadequate for kinetic assessment, necessitating raw video processing for statistical data acquisition on nanotube growth kinetics. A rolling-frame method (also called differential-treatment) was developed to enhance contrast (see details in Supporting Information File 1). This approach, which is a variation of the shade correction commonly used in image processing, involves subtracting a background snapshot containing the illumination profile from the image [18,19]. Typically, the first frame (Figure 1a, left image) of a sequence is used as the background (Figure 1b). However, using the frames with a fixed delay time (chosen between 5 and 30 s, see Supporting Information File 1, Figure S2) significantly increases contrast, by up to an order of magnitude (Figure 1c as well as Supporting Information File 1, Figure S3, and Supporting Information Files 2 and 5). The method is termed “differential” or “rolling frame” shade correction due to the rolling of the background and processed frames.

In such differential videos, the length of a nanotube segment is proportional to its instantaneous rate, as described in Equation S5 (Supporting Information File 1), which adds useful information to the video. Additionally, differential videos capture other processes causing local changes in optical absorption. For instance, if the nanotube structure (also called helicity or

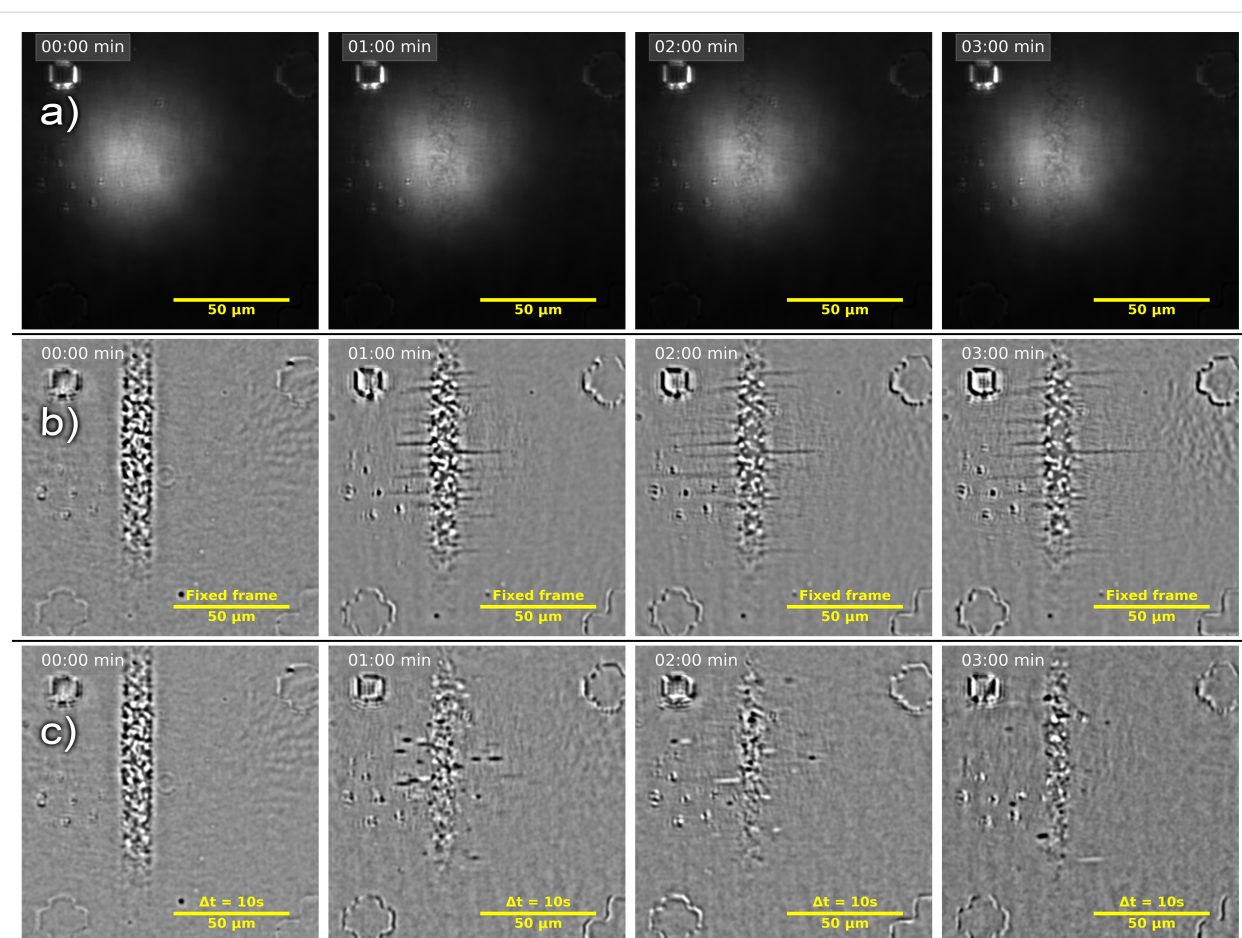


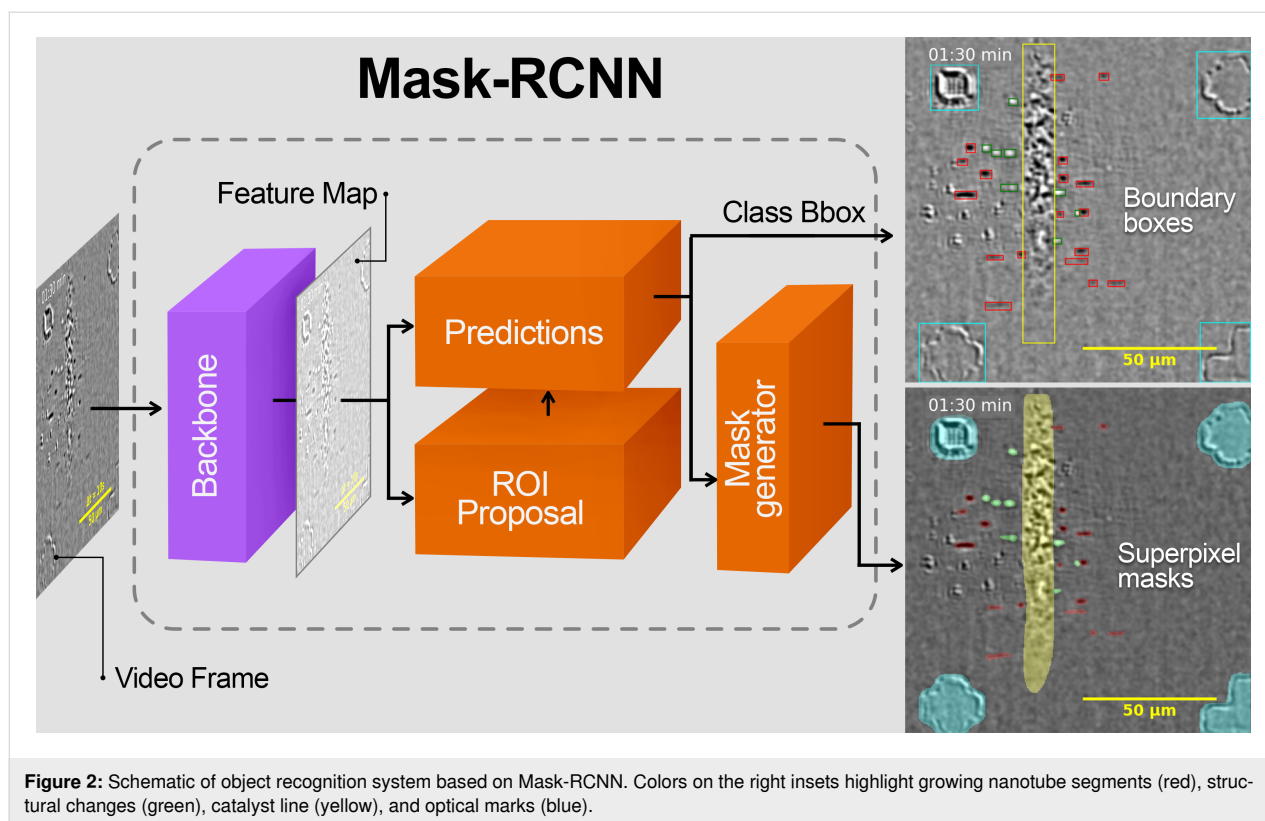
Figure 1: Snapshots from (a) the raw video of carbon nanotube (CNT) synthesis on a stable cut (ST) quartz substrate and from the fully processed video using (b) fixed-frame and (c) differential shading correction (with a 10 s delay). The thick vertical line in each image indicates the catalyst line. Optical markers (i.e., squares, crosses, and L-shapes) are visible in the corners. CNTs appear as thin horizontal lines in the fixed-frame images and as segmented features in the differential images. Grown CNTs show dark contrast in both treatments, while shrinking CNTs are only visible in the differential images and appear as bright segments.

chirality) changes during growth, this manifests as a second segment moving synchronously with the first one: The new chirality appears either as a bright segment if it has lower optical absorption, or as a dark segment otherwise. If the nanotube switches from growth to shrinkage, it appears as a single bright segment moving backward, corresponding to lower optical absorption [20].

We developed a deep learning model to recognize and track both growing nanotubes (dark segments) and structural changes (bright segments) in such differential videos. The model was also trained at recognizing optical marks and catalyst lines (Figure 2) [20]. Kinetic data extraction proceeded in the following steps: (1) object recognition, (2) tracking of recognized objects, (3) verification of tracking, and (4) kinetic curve extraction and analysis. The initial stage of the process utilizes the Mask-RCNN neural network implemented in PyTorch Python library [11]. This architecture integrates several neural

networks in a single pipeline (Figure 2). The image is first processed by the backbone network to produce a feature map. Region of interest (ROI) proposal and prediction networks then localize boundary boxes of the objects and classify them. Finally, a fully connected convolutional network generates a super pixel mask of the recognized objects [21].

The training dataset consisted of 580 manually labeled images from various videos. During training, the dataset was split into a training set of 550 images and a validation set of 30 images. In image segmentation, this data amount does not hinder model quality, and subsets as small as 30 images have been reported sufficient in some cases [22,23]. Nonetheless, we analyzed the dependence of the model prediction quality on the training set size (see Supporting Information File 1, Figure S6 and Table S1). Given the varying quality of our videos, augmentations were applied to account for possible variations in image brightness, contrast, and nanotube localization



(Supporting Information File 1, Figure S4). This approach not only expanded the dataset without additional manual labeling but also increased the robustness and stability of the trained model [24] (Supporting Information File 1, Table S1).

The fully trained model detected segments corresponding to nanotube growth and structural changes, as well as optical marks and catalyst lines (Supporting Information File 1, Figure S5, and Supporting Information File 3). This process was conducted frame by frame through the video. To extract the growth kinetics of each nanotube, the as-recognized objects were tracked using the Hungarian method [25] and Kalman filter (or linear quadratic estimation) [26], which are widely used for object tracking [27]. The Hungarian algorithm matched masks across successive frames. However, some video frames are unrecognizable due to imaging artifacts, illumination instability, or uncompensated vibrations. Following the initial tracking stage, all segments were grouped into clusters of varying sizes, corresponding to objects recognized across consecutive frames. The clusters were then subjected to a Kalman filter to merge segments corresponding to the same nanotube (Supporting Information File 4).

The information about tracked segments is entered into tables for final manual verification and labeling of events. This manual step remains essential due to the complexity of nano-

tube kinetics, which involves switches between growth, pauses, shrinkage, and structure change during growth [20]. Pauses (Figure 3c,d) cannot be efficiently traced by the Hungarian method or Kalman filter, necessitating manual verification to ensure correct assignment of newly grown segments to the same or another close nanotube (Figure 3). Despite this final manual check, automating recognition and tracking steps has accelerated kinetic extraction, significantly enhanced the time resolution of kinetic curves (Figure 3b–d), and increased the repeatability of kinetic measurements. A comparison of manually measured and DL model-extracted results is shown in Figure 3a,b.

Figure 4 compares the main kinetic parameters extracted manually and using the AI-assisted method from in situ videos of nanotube growth performed at the same growth temperature and ethanol partial pressure. In Figure 4a, the lifetime and segment growth duration represent the time during which the nanotube grew at a constant rate (see dots clusters of the same color in Figure 4). Linearly grown nanotubes exhibit constant growth rates from start to finish, observed in about half of the cases. The other half show one or more stochastic rate changes, or even growth–shrinkage switches separated by pauses [20]. In both cases, the close alignment of distribution centers, overlapping standard deviations, and similar overall profiles of the kinetic parameters extracted manually (MR) and via AI-assisted

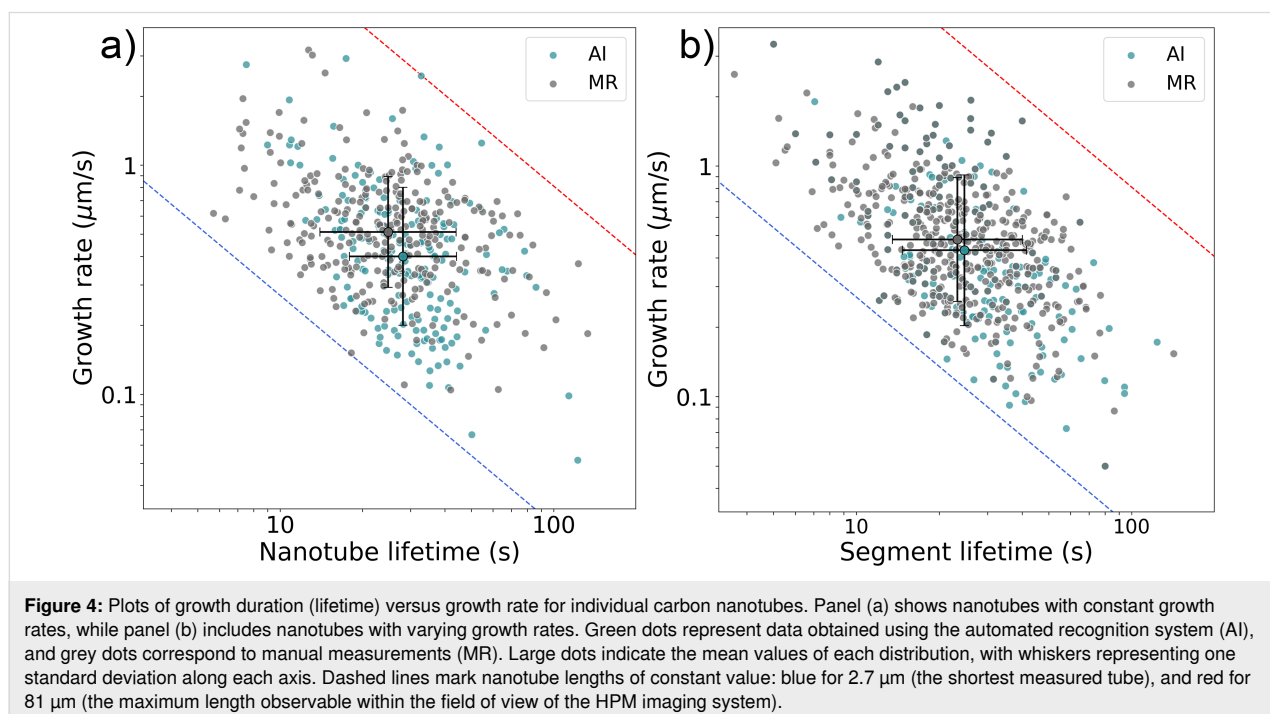
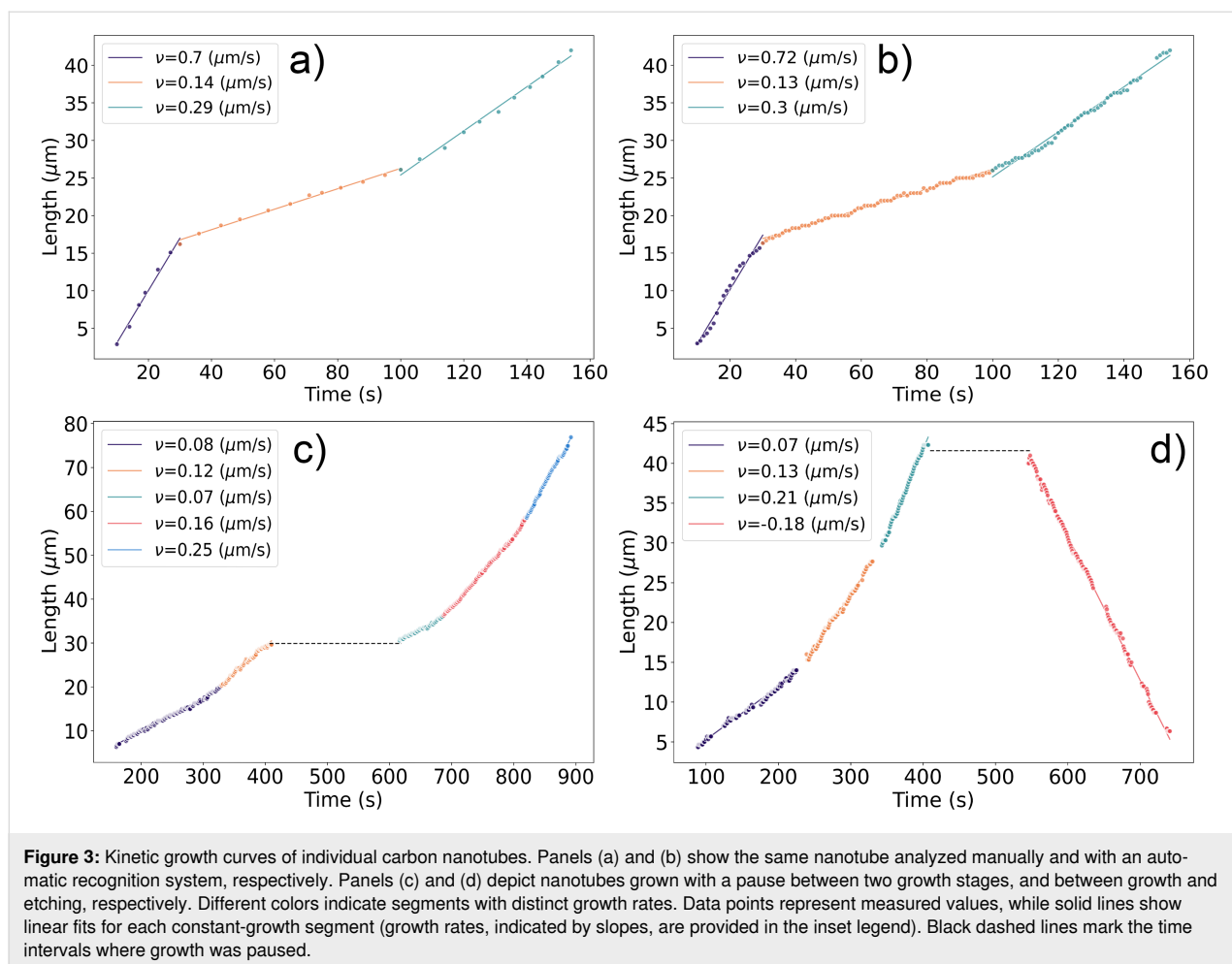


Table 1: Statistical metrics corresponding to the data shown in Figure 4, comparing kinetic parameters derived from manual and AI-assisted analyses.

Plot ref.	Kinetic parameter	Mean \pm std	p-value	Cohen's d	Interpretation
Figure 4a	growth rate AI	0.51 ± 0.43	0.026	0.213	difference significant, but minor
	growth rate MR	0.59 ± 0.39			
	lifetime AI	31.2 ± 16.1	0.226	0.112	insignificant difference
	lifetime MR	29.2 ± 18.6			
Figure 4b	growth rate AI	0.55 ± 0.48	0.66	0.03	insignificant difference
	growth rate MR	0.57 ± 0.41			
	segment growth durations AI	27.4 ± 16.4	0.37	0.069	insignificant difference
	segment growth durations MR	26.3 ± 15.2			

tracking confirm the consistency between the two approaches. Statistical testing (Table 1) supports this: While one instance of growth rate comparison (Figure 4a) yielded a statistically significant difference ($p = 0.026$), the associated effect size was small ($d = 0.213$), indicating only a minor practical difference. All other comparisons showed no statistically significant differences and negligible effect sizes, further underscoring the agreement between methods. This consistency validates the applicability of both manual and AI-assisted techniques for analyzing both linear and non-linear CNT growth kinetics. The difference in the number of data points shown in Figure 4 arises from the exclusion of videos manually analyzed during model training from the AI-based analysis, in order to avoid bias and ensure an independent performance assessment. The AI-extracted data in Figure 4 represents a small fraction of all data acquired using this method, forming the foundation for developing a statistically supported model of carbon nanotube growth kinetics. On average, extracting kinetic data from a single video takes approximately 6 h for a time resolution of 5 s using the manual method, compared to just 2 h for a time resolution of 1 s with the deep learning model, meaning that the throughput was increased by about a factor of 15 (Figure 3a,b).

Conclusion

We have identified several critical steps in developing a DL model for extracting kinetic data from in situ microscopy videos of individual nano-objects. A specific challenge is the low signal from nano-objects and fast kinetics on the video time scale. Applying differential treatment not only drastically improves contrast but also enhances a human operator's understanding of ongoing processes. This is crucial for manual recognition and tracking and for properly labeling the dataset used to train the DL model, allowing for comparison of data obtained by both methods to validate their coherence. Properly designed training datasets and augmentations improve model robustness and stability without increasing training data volume.

Despite significant advancements, there is room for improvement in processing algorithms. In particular, manual verification and labeling of tracked nanotubes, particularly for complex cases, remains essential and is currently the most time-consuming step in the process [20]. Currently, brightness and contrast adjustments rely on empirical hyperparameters, which could benefit from self-tuning or non-parametric algorithms. The same applies to FFT and Gaussian filtering parameters, which are presently constant. Additionally, the AI-based algorithm for kinetic extraction from differentially treated videos can be refined, regardless of the performance comparable to the current state-of-the-art models [28]. Expanding the dataset to include more diverse video quality and nanotube localization can lead to more stable models [29]. Moreover, replacing computationally demanding models like Mask-RCNN with smaller models could enhance overall performance, although potentially reducing recognition accuracy.

Automating recognition and tracking is essential for high-throughput video analysis, which is critical for understanding and modeling complex nanoscale phenomena. The differential video treatment approach, combined with the deep learning-enhanced kinetic extraction framework presented in this work, enabled the extraction of detailed kinetic parameters, specifically, growth rates, lifetimes, and final segment lengths of over 2000 individual carbon nanotubes observed across more than 50 in situ homodyne polarization microscopy videos acquired under a range of synthesis pressures and temperatures. This statistic is an order of magnitude larger than any other study on the growth kinetics of individual carbon nanotubes realized using other methods [30,31].

This large dataset provided unprecedented statistical insight into CNT growth kinetics, revealing complex behaviors such as intermittent switching between growth, pause, and etching modes, even under nominally constant synthesis conditions

[20]. These observations, further supported by complementary Raman spectroscopy, served as a foundation for the development of a new mechanistic model of CNT growth, in which the structure and dynamics of the CNT edge at the catalyst interface govern transitions between kinetic regimes [32].

The versatility of the proposed method extends beyond the specific imaging modality used in this study. Its robustness makes it applicable to other in situ imaging platforms for nanomaterials, such as environmental transmission electron microscopy (ETEM) of CNTs [31,33] and environmental scanning electron microscopy (ESEM) of graphene growth and etching processes [34].

Supporting Information

Supporting Information includes a PDF file with the expanded description of data processing and ten videos derived from two distinct experimental samples. The videos (Supporting Information Files 2–11) illustrate various stages of video processing, object recognition, and nanotube tracking as described below. **Note:** Recognition and tracking were performed only on the video in Supporting Information File 2 as it was excluded from model training to avoid bias. The image sequences in Supporting Information Files 6–11 can be viewed using the open source free software ImageJ [35,36].

Supporting Information File 1

Additional information regarding video processing. Expanded description of the differential video processing, comprehensive explanation of model training process, evaluation of different models, and description of tracking process.

[<https://www.beilstein-journals.org/bjnano/content/supplementary/2190-4286-16-96-S1.pdf>]

Supporting Information File 2

This video shows raw, fixed-frame, and rolling-frame (differentially processed) in situ sequences.

[<https://www.beilstein-journals.org/bjnano/content/supplementary/2190-4286-16-96-S2.tiff>]

Supporting Information File 3

This video presents the differentially processed sequence after object recognition. Masks and bounding boxes are superimposed on the left and right halves of the frame, respectively.

[<https://www.beilstein-journals.org/bjnano/content/supplementary/2190-4286-16-96-S3.tiff>]

Supporting Information File 4

This finalized differential video demonstrates the complete recognition and tracking pipeline, including application of the Hungarian algorithm, Kalman filtering, and manual verification. Masks are displayed on the left side and bounding boxes on the right.

[<https://www.beilstein-journals.org/bjnano/content/supplementary/2190-4286-16-96-S4.mp4>]

Supporting Information File 5

This video shows raw, fixed-frame, and rolling-frame (differentially processed) in situ sequences.

[<https://www.beilstein-journals.org/bjnano/content/supplementary/2190-4286-16-96-S5.mp4>]

Supporting Information File 6

This video shows an uncompressed .tiff image sequence of the raw sequence corresponding to the content shown in the video in Supporting Information File 2.

[<https://www.beilstein-journals.org/bjnano/content/supplementary/2190-4286-16-96-S6.mp4>]

Supporting Information File 7

This video shows an uncompressed .tiff image sequence of the fixed-frame processed sequences corresponding to the content shown in the video in Supporting Information File 2.

[<https://www.beilstein-journals.org/bjnano/content/supplementary/2190-4286-16-96-S7.mp4>]

Supporting Information File 8

This video shows an uncompressed .tiff image sequence of the rolling-frame (differentially processed) sequences corresponding to the content shown in the video in Supporting Information File 2.

[<https://www.beilstein-journals.org/bjnano/content/supplementary/2190-4286-16-96-S8.tiff>]

Supporting Information File 9

This video shows an uncompressed .tiff image sequence of the raw sequence corresponding to the content shown in the video in Supporting Information File 5.

[<https://www.beilstein-journals.org/bjnano/content/supplementary/2190-4286-16-96-S9.tiff>]

Supporting Information File 10

This video shows an uncompressed .tiff image sequence of the fixed-frame processed sequences corresponding to the content shown in the video Supporting Information File 5.

[<https://www.beilstein-journals.org/bjnano/content/supplementary/2190-4286-16-96-S10.tiff>]

Supporting Information File 11

This video shows an uncompressed .tiff image sequence of the rolling-frame (differentially processed) sequences corresponding to the content shown in the video in Supporting Information File 5.

[<https://www.beilstein-journals.org/bjnano/content/supplementary/2190-4286-16-96-S11.tiff>]

Acknowledgements

V.J. designed and coordinated the research. S.T. prepared the catalyst samples and performed SEM characterization. V.P. performed the in situ optical imaging, developed the protocols of data treatment and analyzed the experimental data.

This work is partially based on Vladimir Pimonov's doctoral thesis ("Growth kinetics of individual carbon nanotubes studied by in situ optical microscopy", Université Montpellier, 2021) [37].

Funding

The authors acknowledge the support of the Agence Nationale de la Recherche (grant ANR-20-CE09-0007-01).

Author Contributions

Vladimir Pimonov: conceptualization; data curation; investigation; software; validation; visualization; writing – original draft; writing – review & editing. Said Tahir: methodology; resources. Vincent Jourdain: conceptualization; formal analysis; funding acquisition; investigation; methodology; project administration; resources; supervision; validation; writing – review & editing.

ORCID® iDs

Vladimir Pimonov - <https://orcid.org/0000-0001-8067-4790>

Vincent Jourdain - <https://orcid.org/0000-0002-2695-078X>

Data Availability Statement

The code used for training the model, with the example of model and training data is available as a zenodo repository "Training Script for Mask R-CNN on In-Situ Homodyne Microscopy Videos of Carbon Nanotube Growth" <https://doi.org/10.5281/zenodo.15697833>.

Preprint

A non-peer-reviewed version of this article has been previously published as a preprint: <https://doi.org/10.48550/arXiv.2410.13594>

References

- Bernholc, J.; Brenner, D.; Buongiorno Nardelli, M.; Meunier, V.; Roland, C. *Annu. Rev. Mater. Res.* **2002**, *32*, 347–375. doi:10.1146/annurev.matsci.32.112601.134925
- Ajiki, H. *J. Phys.: Condens. Matter* **2012**, *24*, 483001. doi:10.1088/0953-8984/24/48/483001
- Franklin, A. D.; Luisier, M.; Han, S.-J.; Tulevski, G.; Breslin, C. M.; Gignac, L.; Lundstrom, M. S.; Haensch, W. *Nano Lett.* **2012**, *12*, 758–762. doi:10.1021/nl203701g
- Monniello, L.; Tran, H.-N.; Vialla, R.; Prévot, G.; Tahir, S.; Michel, T.; Jourdain, V. *Phys. Rev. B* **2019**, *99*, 115431. doi:10.1103/physrevb.99.115431
- Denker, J. S.; Gardner, W. R.; Graf, H. P.; Henderson, D.; Howard, R. E.; Hubbard, W.; Jackel, L. D.; Baird, H. S.; Guyon, I. Neural Network Recognizer for Hand-Written Zip Code Digits. In *Advances in Neural Information Processing Systems 1 (NIPS 1988)*, Touretzky, D., Ed.; 1988. https://proceedings.neurips.cc/paper_files/paper/1988/file/a97da629b098b75c294dfdc3e463904-Paper.pdf
- Han, B.; Lin, Y.; Yang, Y.; Mao, N.; Li, W.; Wang, H.; Yasuda, K.; Wang, X.; Fatemi, V.; Zhou, L.; Wang, J. I.-J.; Ma, Q.; Cao, Y.; Rodan-Legrain, D.; Bie, Y.-Q.; Navarro-Moratalla, E.; Klein, D.; MacNeill, D.; Wu, S.; Kitadai, H.; Ling, X.; Jarillo-Herrero, P.; Kong, J.; Yin, J.; Palacios, T. *Adv. Mater. (Weinheim, Ger.)* **2020**, *32*, 2000953. doi:10.1002/adma.202000953
- Chen, Z.; Liu, X.; Yang, J.; Little, E.; Zhou, Y. *Comput. Geosci.* **2020**, *138*, 104450. doi:10.1016/j.cageo.2020.104450
- Förster, G. D.; Castan, A.; Loiseau, A.; Nelayah, J.; Alloyeau, D.; Fossard, F.; Bichara, C.; Amara, H. *Carbon* **2020**, *169*, 465–474. doi:10.1016/j.carbon.2020.06.086
- Liu, K.; Hong, X.; Zhou, Q.; Jin, C.; Li, J.; Zhou, W.; Liu, J.; Wang, E.; Zettl, A.; Wang, F. *Nat. Nanotechnol.* **2013**, *8*, 917–922. doi:10.1038/nnano.2013.227
- Hum, Y. C.; Lai, K. W.; Mohamad Salim, M. I. *Complexity* **2014**, *20*, 22–36. doi:10.1002/cplx.21499
- Pytorch. <https://pytorch.org/> (accessed Nov 2, 2024).
- He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask R-CNN. In *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, Oct 22–29, 2017; IEEE, 2017; pp 2980–2988. doi:10.1109/iccv.2017.322
- He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, June 27–30, 2016; IEEE, 2016; pp 770–778. doi:10.1109/cvpr.2016.90
- Lin, T.-Y.; Dollár, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S. Feature Pyramid Networks for Object Detection. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, July 21–26, 2017; IEEE, 2017; pp 936–944. doi:10.1109/cvpr.2017.106
- Lin, T.-Y.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Dollár, P.; Zitnick, C. L. Microsoft COCO: Common Objects in Context. In *Lecture Notes in Computer Science, vol 8693*, Computer Vision – ECCV 2014; Fleet, D.; Pajdla, T.; Schiele, B.; Tuytelaars, T., Eds.; Springer: Cham, Switzerland, 2014; pp 740–755. doi:10.1007/978-3-319-10602-1_48

16. George Karimpanal, T.; Bouffanais, R. *Adapt. Behav.* **2019**, *27*, 111–126. doi:10.1177/1059712318818568
17. Google Colab. <https://colab.research.google.com/notebooks/intro.ipynb> (accessed Nov 2, 2024).
18. Peng, T.; Thorn, K.; Schroeder, T.; Wang, L.; Theis, F. J.; Marr, C.; Navab, N. *Nat. Commun.* **2017**, *8*, 14836. doi:10.1038/ncomms14836
19. Goldman, D. B. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 2276–2288. doi:10.1109/tpami.2010.55
20. Pimonov, V.; Tran, H.-N.; Monniello, L.; Tahir, S.; Michel, T.; Podor, R.; Odorico, M.; Bichara, C.; Jourdain, V. *Nano Lett.* **2021**, *21*, 8495–8502. doi:10.1021/acs.nanolett.1c03431
21. Schwing, A. G.; Urtasun, R. *arXiv* **2015**, 1503.02351. doi:10.48550/arxiv.1503.02351
22. Ronneberger, O.; Fischer, P.; Brox, T. Convolutional Networks for Biomedical Image Segmentation. In *Lecture Notes in Computer Science*, vol 9351, Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015; Navab, N.; Hornegger, J.; Wells, W.; Frangi, A., Eds.; Springer: Cham, Switzerland, 2015; pp 234–241. doi:10.1007/978-3-319-24574-4_28
23. Ghosh, P.; Mitchell, M. Segmentation of medical images using a genetic algorithm. In *GECCO '06: Proceedings of the 8th annual conference on Genetic and evolutionary computation*, Seattle, Washington, USA; Association for Computing Machinery: New York, NY, USA, 2006; pp 1171–1178. doi:10.1145/1143997.1144183
24. Shorten, C.; Khoshgoftaar, T. M. *J. Big Data* **2019**, *6*, 60. doi:10.1186/s40537-019-0197-0
25. Kuhn, H. W. *Nav. Res. Logist. Q.* **1955**, *2*, 83–97. doi:10.1002/nav.3800020109
26. Kalman, R. E. *J. Basic Eng.* **1960**, *82*, 35–45. doi:10.1115/1.3662552
27. Bewley, A.; Ge, Z.; Ott, L.; Ramos, F.; Upcroft, B. Simple online and realtime tracking. In *2016 IEEE International Conference on Image Processing (ICIP)*, Phoenix, AZ, USA, Sept 25–28, 2016; IEEE, 2016; pp 3464–3468. doi:10.1109/icip.2016.7533003
28. Papers With Code. Benchmark of neural network for object detection. <https://paperswithcode.com/sota/real-time-object-detection-on-coco> (accessed March 2, 2025).
29. Sharma, P.; Singh, A. Era of deep neural networks: A review. In *2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, Delhi, India, July 3–5, 2017; IEEE, 2017; pp 1–5. doi:10.1109/icccnt.2017.8203938
30. Koyano, B.; Inoue, T.; Yamamoto, S.; Otsuka, K.; Xiang, R.; Chiashi, S.; Maruyama, S. *Carbon* **2019**, *155*, 635–642. doi:10.1016/j.carbon.2019.09.031
31. Vincent, P.; Panciera, F.; Florea, I.; Blanchard, N.; Cojocaru, C. S.; Ezzedine, M.; Taoum, H.; Perisanu, S.; De Laharpe, P.; Ayari, A.; Chaste, J.; Saidov, K.; Mirsaidov, U.; Purcell, S. T.; Legagneux, P. *Carbon* **2023**, *213*, 118272. doi:10.1016/j.carbon.2023.118272
32. Förster, G. D.; Pimonov, V.; Tran, H.-N.; Tahir, S.; Jourdain, V.; Bichara, C. *ACS Nano* **2023**, *17*, 7135–7144. doi:10.1021/acsnano.2c07388
33. He, M.; Wang, X.; Zhang, S.; Jiang, H.; Cavalca, F.; Cui, H.; Wagner, J. B.; Hansen, T. W.; Kauppinen, E.; Zhang, J.; Ding, F. *Sci. Adv.* **2019**, *5*, eaav9668. doi:10.1126/sciadv.aav9668
34. Kidambi, P. R.; Bayer, B. C.; Blume, R.; Wang, Z.-J.; Baehtz, C.; Weatherup, R. S.; Willinger, M.-G.; Schloegl, R.; Hofmann, S. *Nano Lett.* **2013**, *13*, 4769–4778. doi:10.1021/nl4023572
35. Schindelin, J.; Arganda-Carreras, I.; Frise, E.; Kaynig, V.; Longair, M.; Pietzsch, T.; Preibisch, S.; Rueden, C.; Saalfeld, S.; Schmid, B.; Tinevez, J.-Y.; White, D. J.; Hartenstein, V.; Eliceiri, K.; Tomancak, P.; Cardona, A. *Nat. Methods* **2012**, *9*, 676–682. doi:10.1038/nmeth.2019
36. Schneider, C. A.; Rasband, W. S.; Eliceiri, K. W. *Nat. Methods* **2012**, *9*, 671–675. doi:10.1038/nmeth.2089
37. Pimonov, Y. *Growth kinetics of individual carbon nanotubes studied by in situ optical microscopy*. Ph.D. Thesis, Université Montpellier, Montpellier, France, 2021.

License and Terms

This is an open access article licensed under the terms of the Beilstein-Institut Open Access License Agreement (<https://www.beilstein-journals.org/bjnano/terms>), which is identical to the Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0>). The reuse of material under this license requires that the author(s), source and license are credited. Third-party material in this article could be subject to other licenses (typically indicated in the credit line), and in this case, users are required to obtain permission from the license holder to reuse the material.

The definitive version of this article is the electronic one which can be found at:
<https://doi.org/10.3762/bjnano.16.96>