UCLA
UCLA Electronic Theses and Dissertations

Title
Deep Learning Accelerated Studies of Electrochemical Systems – Applications in the Analysis of Wire Electrodes and Cyclic Voltammetry

Permalink
https://escholarship.org/uc/item/9vv1m0g3

Author
Hoar, Benjamin

Publication Date
2023

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA

Los Angeles

Deep Learning Accelerated Studies of Electrochemical Systems – Applications in the Analysis of Wire Electrodes and Cyclic Voltammetry

A dissertation submitted in partial satisfaction

of the requirements for the degree Doctor of Philosophy in Chemistry

by

Benjamin Hoar

2023

© Copyright by

Benjamin Hoar

2023

ABSTRACT OF THE DISSERTATION

Deep Learning Accelerated Studies of Electrochemical Systems –

Applications in the Analysis of Wire Electrodes and Cyclic Voltammetry

by

Benjamin Hoar

Doctor of Philosophy in Chemistry

University of California, Los Angeles, 2023

Professor Chong Liu, Chair

In chemistry, the procurement of data for the training of machine learning models is often an intractable task. The reliance of data generated via experimental routes is two-fold problematic. First, it requires a vast amount of menial labor that will result in large amounts of chemical waste and is thus wasteful of resources and human capital, and second, the establishment of a truly exhaustive dataset in practical terms is highly limited by the accessibility of affordable, diverse experimental systems. Even if data is gathered from the literature, the lack of reported negative results can result in a concerning level of bias in experimentally driven studies. While experimental data is highly valuable as it provides real-world behaviors that are difficult to capture in idealized simulations, these challenges remain and thus simulation and computational approaches to data generation become necessary for the practical evaluation of many important chemical systems.

Given these considerations, I established a workflow in my graduate career for the study of electrochemical systems that are difficult to investigate via standard experimental or computational approaches. Via the marriage of fundamental electrochemistry, automation software, COMSOL finite element simulations, and deep-learning, I have investigated data intensive questions related to wire electrodes and cyclic voltammetry.

In my first research project (**Chapter 2**), I investigated wire electrodes for nitrogen fixation by leveraging models to investigate the influence of wire arrays on the competition between the nitrogen reduction reaction (NRR) and the hydrogen evolution reaction (HER). The challenge with the evaluation of a morphology's influence on an electrode's outputs of interest (current density, A/m^2 and Faradaic efficiency, %) is that bulk analysis via fabrication is extremely time and resource intensive, while also requiring the investigation of large swaths of morphological parameter space (periodicity, P, length, L, and diameter, D) that result in extremely minor differences in electrode output. This concern is exacerbated by the issues of reproducibility and the introduction of multiple materials, where morphological parameters may have a unique influence on each material. While these issues are inherent of laboratory approaches, they can be mitigated by the establishment of a computationally generated dataset that can be used to cheaply explore large regions of the parameter space that constitutes both morphological and kinetic parameters of materials. My approach enabled the investigation of nitrogen fixation in acidic, ambient condition, wire electrodes in general, establishing a putative limit for wire electrode performance with Faradaic efficiencies of approximately 90% and current densities of approximately $-2mA/cm^2$ (under the assumptions of the model). Further, this data-intensive approach provided insights into electrode development – demonstrating the morphology is sometimes a weak predictor of output, while at other times a strong predictor. Further, an

importance analysis of the trained deep learning model indicated that the second step of the associative NRR pathway is likely limiting in many cases, providing fundamental insight towards the rational development of transport restrictive NRR electrodes going forward.

Following this approach to the study of wire electrodes, I investigated deep learning's applicability to the custom design of reactant gradients in oxygen reducing electrodes (**Chapter 3**). In nature, reactant concentration heterogeneity is ubiquitous (e.g., oxygen availability in biofilms). However, the study of systems with a desired diffusion gradient is challenging as reproducible direct control over reactant availability is non-trivial. One solution to this challenge is electrochemically generated concentration gradients using wire electrodes, where applied potentials result in depletion of redox active species along the z-axis (perpendicular to base) of the electrode. While the technology exists for the establishment of these gradients, the selection of a specific gradient was still an open problem. To address this need, I developed models to predict oxygen concentration gradients (assuming a platinum electrode) and hydrogen peroxide gradients (assuming a gold electrode) to permit the inverse design of biologically relevant reactant gradients for subsequent study of microbial systems.

In a pivot to an investigation of a fundamental problem in electrochemistry, I began the study of deep-learning approaches for mechanism assignment in cyclic voltammetry (**Chapter 4**). Mechanism assignment is fundamental to the understanding of an electrochemical system – where further analyses of the chemistry is dependent on correct deduction of the number, order, and type of electrochemical (*E* steps) and chemical (*C* steps). Despite the fundamental and foundational importance of this assignment, experimentalists still perform mechanism deduction using qualitative and semi-quantitative means. This is problematic due to diversity in experience level, biased approaches to assignment, and the inherent sluggishness of manual inspection. Because of

these concerns, I applied computer vision techniques to the elucidation of electrochemical mechanisms from cyclic voltammetry data. Beginning with the *E, EC, CE, DISP1*, and *ECE* mechanisms, I developed a framework to simulate these mechanisms while ensuring a diverse representation of the possible CV traces for each mechanism, with care taken to ensure their separability (i.e., a poorly constrained *EC* framework would be indistinguishable from *E*). Upon implementation of COMSOL models and MATLAB/Python scripts for each, a training set was generated and used to train a ResNet–18 neural network, which achieved accuracies of 98.5%. Further, the classification probabilities inherent of deep-learning classification tasks have proven valuable for the nascent development of automated, deep-learning driven artificial intelligence platforms that seek to reduce ambiguity in classification probabilities by tuning experimental parameters.

In a follow up to this research, I have investigated an object detection approach to enable the analysis of far more complex, multi-event CVs (**Chapter 5**). ResNet–18 is a powerful feature extraction tool – hence its success as an image/matrix data classifier – so it is amenable to object detection techniques that rely on the extraction of regions of data that have features correlating to several objects. In my case, those objects are the constituent electrochemical mechanisms that combine to give an overall complex CV trace. The implementation of a custom Faster R–CNN framework with one dimensional region proposals (as compared to the typical 2D) has achieved F_1 scores up to 0.932 when classifying and regressing CVs with contributions from up to eight possible mechanisms: *E, EC*$_a$*, EC*$_b$*, ECE, DISP1, T, SR,* and *EC'*.

In the final chapter (**Chapter 6**), I will discuss my work developing software for the purposes of improving course evaluation practices in the UCLA physical sciences. Therein I developed software that digests and recapitulates student course evaluations as quantitative,

summative reports that allow for quicker analysis of course strengths and weaknesses. This tool provided evidence of the power of sentiment analysis, a form of deep learning related to correlation of text data with implied positivity (e.g., how good/effective course practices were).

During my graduate career, I have shown deep learning's applications in studying problems in electrochemistry that were otherwise intractable to explore. This combination of chemistry with deep learning is not only valuable for the generation of useful tools and the designing of systems, but it can also provide fundamental insights and allow for technologies that will accelerate the advancement of chemistry in a world where automation and data analysis is driving advancements in all fields.

The dissertation of Benjamin Hoar is approved.

Anastassia N. Alexandrova

Alexander M. Spokoyny

Xiangfeng Duan

Chong Liu, Committee Chair

University of California, Los Angeles

2023

DEDICATION

I dedicate my thesis to my family and my cat. My parents Alyson and Bruce have always supported me in whatever path I chose – my mother especially ensuring my sanity away from home and my father when I return home. I also must thank my grandmother Brianne and grandfather David for their support, without them I could not have pursued the life I have. Finally, thank you Snow Pea, you bring immense comfort when I need it most. I love you all.

Table of Contents

Chapter 4. Applications of bulk generation of simulated electrochemical data for automation of mechanism assignment in cyclic voltammetry

Chapter 5. An object detection approach for the automated analysis of more complex cyclic voltammograms

Chapter 6. Sentiment analysis for student course reviews, an exploration of non-electrochemical data

List of Figures

Chapter 2. The accelerated study of simulated nitrogen fixing wire electrodes using deep learning

Chapter 3. The inverse design of oxygen and hydrogen peroxide concentration gradients in wire electrodes

Chapter 4. Applications of bulk generation of simulated electrochemical data for automated mechanism assignment in cyclic voltammetry

Chapter 5. An object detection approach for the automated analysis of more complex cyclic voltammograms

Chapter 6. Sentiment Analysis for student course reviews, an exploration of non-electrochemical data

List of Tables

Chapter 2. The accelerated study of simulated nitrogen fixing wire electrodes using deep learning

Chapter 4. Applications of bulk generation of simulated electrochemical data for automated mechanism assignment in cyclic voltammetry

List of Supplemental Texts

Acknowledgements

I would like to first acknowledge my advisor and mentor Prof. Chong Liu. He took me in as a relatively aimless graduate student (I didn't want to continue what I had worked on as an undergraduate) and provided me excellent and patient advisement as I established my own graduate career. Without Chong, I doubt I would have been able to explore the machine learning, simulation, and research software directions that I did. I likely would not have had the confidence to attempt this path alone, and Chong's trust in me allowed me build confidence and succeed in an area I had never even considered before.

In addition to Chong, the other group members in my research team made my graduate career possible. The general air of collaboration and welcoming attitude of lab members made for a graduate career I think many other students would envy. First, I have to thank Dr. Shengtao Lu for helping me to join his nitrogen fixation project and allowing me to learn from him while building the skills necessary for my first paper, resulting in my first publication in a new field with efficiency that I would have not achieved without him.

In a similar vein, I have to thank Yi Chen and Jingyu Wang for their work in our shared first-author work published in *PNAS*. Such multi-faceted research is not possible alone, and as a computational researcher, I am highly dependent on, and thankful for, those who are able to do experiments alongside me. Further, their support and professionalism made the process easy and headache free, I especially appreciate Yi leading the charge on writing as that part is often stressful.

Finally, for my research into cyclic voltammetry, I have a ton of people to thank. Thank you to Chong for his help establishing an electrochemical framework for this work, his expertise in the field of electrochemistry is amazing, and I am lucky to have worked with someone so capable. Similarly, thank you to Professor Cyrille Costentin for his advisement on the work –

again, Cyrille is intimidatingly well-versed in all aspects of cyclic voltammetry, and it is humbling to work with someone so revered as a graduate student. Following this, I can say the same things about Professor Quanquan Gu. Working with the director of UCLA's General Artificial Intelligence Lab as a chemistry graduate student is an honor and I am extremely fortunate to have worked under his advisement. Finally, I would like to thank Dr. Hongyuan Shen, Dr. Shuangning Xu, Weitong Zhang, Yuanzhou Chen, and Jingwen Sun for their contributions to this work on practical aspects. It would be overly verbose to enumerate their contributions, but I humbly thank all of them for enabling my research into the topic of automated mechanism assignment.

I would like to thank Dr. Erin Sparck, Dr. Roshini Ramchandran, Dr. Marc Levis, and Ke Wu for their help in developing software with the Center for the Advancement of Teaching. It was a treat getting a break from thinking about chemistry all the time, and our work on augmenting the course evaluation practices at UCLA was a great diversion within which I learned a ton about software development and education as a field of study.

While I am sure I am missing some people (a graduate career is long and impossible without help), I must now thank people in my personal life. Thank you to Alyson, my mom, and Bruce, my dad, and Josh, my brother. Without your support, I would not have been able to stay committed and dedicated, keeping things in perspective so as not to get too overwhelmed. Thank you to my grandparents David and Brianne for their financial support as an undergraduate and general support throughout my life, without you I would not be remotely the same person I am today. Thank you to Snow Pea, my cat, our furry friends can never be forgotten as there are times when they are the only source of comfort available. Thank you, Sarah Smyth, you have been a tremendous addition to my life my last couple years at UCLA, you have helped ensure that I live my life and don't get too lost in work or stress. Thank you, Roy Howie, you have been a truly loyal

friend, someone who is always there for me no matter what and sees right past my flaws no matter how apparent they are at times. Finally, I have to thank Ben Natinsky, Roselyn Rodrigues, Xun Guan, and Zach Hern. You were the “skwaad” for me early on, without you this whole experience would have been impossible, thank you for your consistent friendship during the hardest times of my career. I have to especially thank Ben, mostly because he just “gets me,” an informal, but necessary acknowledgment.

In addition to all of these people, I must thank one final group of non-UCLA people, my “gaming friends.” While I do know these “gaming friends” in “real life,” I don’t get to see them often. Fortunately with modern technology, I get to maintain great friendships with people thousands of miles away, enabling my reclusive nature without becoming fully isolated. Thank you to Paul Wesnofske, Gregory Blachly, and Henry Capants, you guys are the best.

I would like to thank my committee: Professors Anastassia Alexandrova, Alex Spokoyny, and Xianfeng Duan. As a student in the academic machine, it is always humbling to be surrounded by such intelligent people, and fortunately professors such as those on my committee and Professor Liu have made the experience both educational and welcoming. Thank you all for your time and help.

On funding, thank you to UCLA, Chong Liu, and the National Science Foundation for supporting me throughout my career. I have been lucky to have financial support in the form I most desired at all times throughout my career, and I am incredibly fortunate and thankful for that.

One more time, thank you to everyone I have mentioned above, you are the people that actually made my graduate career what it was: stimulating, fun, welcoming, and rewarding.

VITA

EDUCATION

University of California, Los Angeles — **Los Angeles, CA**
Ph.D. Track Chemistry – Inorganic Concentration — Expected June 2023
M.S. Chemistry – Inorganic Concentration — March 2020
Advanced to Candidacy — July 2020
University of Miami — **Coral Gables, FL**
Bachelor of Science: Chemistry, Minor in Computer Science — May 2017

WORK EXPERIENCE

University of California, Los Angeles — **Los Angeles, CA**
Graduate Researcher, Department of Chemistry and Biochemistry — August 2018 – Present
Advisor: Chong Liu

- Developed Faster R–CNN approach for reaction mechanism detection in cyclic voltammograms composed of up to four redox reactions and eight possible reaction types – up to 95% classification accuracy achieved
- Researched reaction condition optimization by leveraging residual neural network (ResNet) classification probabilities that indicate which reaction types are present in cyclic voltammetry experiments
    - Bayesian optimization can be leveraged to find reaction conditions that maximize desired reaction classes by minimizing classification probabilities of unwanted classes
- Wrote software to recycle expensive to obtain (two months) dataset of ~20,000 single-event cyclic voltammograms into merged multi-event voltammograms necessary for Faster R–CNN research
    - 20,000 merged voltammograms are generated in less than three hours using this software
- Trained ResNet using PyTorch framework to classify simulated electrochemical cyclic voltammetry data into five possible electrochemical mechanism classes: *E*, *EC*, *CE*, *DISP*, *ECE*
    - Data was simulated using MATLAB for simulation automation, COMSOL for mathematical modelling, and Python for experiment constraint implementation and data cleansing
- Developed JavaScript/React and Flask based website that allows public access to my mechanism classification tool (cv-ml.com, you can download samples to explore at https://github.com/benhoar/CVMLSAMPLES)
- Led interdepartmental collaboration with Prof. Quanquan Gu's computer science group, oversaw five students on yearlong project to build software tools for the automated analysis of electrochemistry data
- Implemented TensorFlow framework and Hyperopt Bayesian optimization library for the tuning of simulated CO_2 reduction electrode surface morphology
- Wrote python scripts to expedite research activities including the analysis of rotating ring disk electrode experiments
- Published two (N_2 fixation, O_2 reduction) papers with continuing work (CO_2) related to the inverse design of wire electrodes
- Experienced with simulation of electrochemistry phenomenon (electrolysis, reactant gradients, cyclic voltammetry, etc.) with MATLAB/COMSOL software in combination with custom python scripts for data sanitization and post-processing
- Acted as primary communicator of chemistry and machine learning/coding concepts between my research group and computer science collaborators with whom I develop research tools
- Mentored undergraduate researcher in the applications of machine learning in chemistry, guiding him to develop a script to automate the sampling of appropriate training data for tasks depending on user requirements

University of California, Los Angeles **Los Angeles, CA**
Graduate Teaching Assistant, Department of Chemistry and Biochemistry September 2018 – March 2020
Advisors: Prof. Chong Liu, Prof. Laurence Lavelle

- Taught General Chemistry for the Life Sciences I (14A) and General Chemistry for the Life Sciences II (14B)

University of California, Los Angeles **Los Angeles, CA**
Software Developer, UCLA Center for the Advancement of Teaching June 2019 – December 2020

- Provided regression model and feature importance-based visualizations to chemistry department in its efforts to correctly place students in extra-aid, regular, and honors general chemistry courses
- Led interdepartmental collaboration with the Center for the Advancement of Teaching (CAT), organizing two chemists, three education experts, and twenty undergraduate data labelers for the development of large-enrollment-course-feedback augmentation software
- Implemented sentiment analysis and text processing python scripts using the Natural Language Toolkit (NLTK) to enhance the UCLA student course-review process by converting bulk student course review data into automatically generated summative, quantitative, and qualitative LaTeX reports using PyLaTeX
- Worked with Google's Sentiment Analysis API and the VADER sentiment analysis tool to predict the relative positivity of student course review statements
- Iteratively gathered feedback and suggestions from CAT and UCLA educators to improve the format of and confidence in the student course review tool

University of Texas, Austin **Austin, TX**
Visiting Undergraduate Researcher, UT Austin Department of Chemistry June 2017–August 2018

- Supported research into the synthesis of molecular catalysts for hydrogen evolution
- Conducted inorganic syntheses and silicon wafer preparation for attachment of catalysts to silicon for better electronic coupling between photovoltaic and catalyst

PUBLICATIONS (†: equal contribution)

(1) Hoar, B., et al., Machine learning enabled exploration of morphology influence on wire-array electrodes for electrochemical nitrogen fixation. *The Journal of Physical Chemistry Letters* **2020**, *11*, 4625–4630
(2) Chen, Y., † Wang, J., † Hoar, B., † et al., Machine learning based inverse design for electrochemically controlled microscopic gradients of O_2 and H_2O_2. *PNAS* **2022**,*19(32),* e2206321119
(3) Hoar, B., et al., Electrochemical mechanistic analysis from cyclic voltammograms based on deep-learning. *ACS Measurement Science Au* **2022**, *2(6),* 595–604
(4) Hoar, B., et al. Summative Student Course Review Tool Based on Machine Learning Sentiment Analysis to Enhance Life Science Feedback Efficacy, *arXiv:2301.06173* **2023**
(5) Sharma, S.K., Micic, M., Li S., Hoar, B., Paudyal, S., Zahran, E.M., Leblanc, R.M., Conjugation of Carbon Dots with β-Galactosidase Enzyme: Surface Chemistry and Use in Biosensing. *Molecules* **2019,** *24(18)***,** 3275

HONORS AND AWARDS

(1) Phi Beta Kappa Honor Society Fellow *March 2018*
(2) Recipient of Dean's Scholar Award and Grant at UCLA *September 2018*
(3) NSF–Innovation at the Nexus of Food, Energy, and Water Systems Fellow *December 2018*
(4) Recipient of Excellence in Second Year Research and Academics Award UCLA. *May 2021*
(5) Inventor on filed provisional patent related to mechanism classification tool *December 2022*

Chapter 1. General Introduction

Motivation: Obtaining Deep Learning Training Sets is Challenging

In chemistry, there are inherent limitations on studies that may otherwise be of interest due to time and practicality.[1, 2] For example, while researchers may wish to know the ideal morphology of an electrode for a given reaction, exhaustively searching every electrode with any reasonable parametric resolution becomes unreasonable. This general issue of efficiency extends beyond optimization challenges too, where inverse design and automated data analyses are further examples where the development of protocols depends on large amounts of data that are not readily attainable. In addition to temporal challenges, experiments are limited by the accessibility of materials with known parameters (e.g., exchange current density for intrinsic parameters and morphological characteristics such as wire periodicity for extrinsic parameters), resulting in poor mapping of possible parameter spaces and the risk of missing combinations of materials and experimental conditions that result in optimal outputs. Finally, even if data is aggregated from the literature to preclude the need for in-house experimental data generation, the gathered data is likely to be biased due to the emphasis of "positive" results published in the literature.[3]

Considering the issues of gathering experimental data via lab-work or data mining, simulation as a means of generating training sets reveals itself as a valuable alternative. Simulating data allows for relatively rapid acquisition of diverse datasets; datasets of sufficient size and representational power for the successful training of deep learning models that must perform well given any reasonable input (not just inputs that lead to "good" outputs). In addition to providing large amounts of data, the direct control over the framework under which data is generated permits experts to ensure that the appropriate permutations and combinations of inputs are explored.

Through the combination of software engineering, modern simulation tools such as COMSOL Multiphysics, and open-source deep learning tools, chemists have the opportunity to explore systems that otherwise may have been bypassed due to practical considerations. In the research presented in this thesis, I will demonstrate how these tools can be combined to address issues in electrochemistry ranging from optimization of wire electrodes to inverse design of concentration gradients, to mechanism assignment in cyclic voltammetry.

Precedence: Deep Learning in Electrochemistry and Simulated Data in Deep Learning

Machine learning, and more specifically deep learning, have seen widespread adoption in chemistry in recent years. In the field of electrochemistry, deep learning has been applied to a number of problems such as the analysis of biosensor data,[4] discovery of electrocatalysts for the hydrogen evolution reaction,[5] the screening of viable battery materials,[6] as a guide tool for discovering electrochemical reactions of interest to organic chemists,[7] and an ever-expanding list of directions not mentioned here. While these publications relied on experimental data as guides to their insights, not all situations are amenable to such practices. In some cases, simulated data is the most practical avenue for the training of deep learning models. For example, the denoising[8] and segmentation[9] of tunneling electron microscope (TEM) images requires a high number and diversity of TEM measurements, measurements that would have to be gathered with intentionally added noise and edge conditions so as to make the task of gathering data infeasible. Further, computational tools as the basis for generalizations about real-world systems is common, as with the use of DFT for the establishment design rules for oxidation catalysts[10] or with the inverse design of nanoparticles with specific scattering behavior.[11] In fact, the simulation of cyclic voltammograms with the goal of mechanism assignment has been studied before, but the system studied was highly constrained.[12] Given the applications of deep learning to electrochemistry and

the use of simulated data as the basis for studying real-world systems, the work presented herein is following well-established routes of scientific exploration.

This Work: Accelerated Studies of Electrochemical Systems

Herein, a general workflow is discussed that enables the study of chemical systems that are resistant to purely lab-based approaches.[13, 14] This workflow can be discussed in the following general steps: establish a framework of variables, equations, and boundary conditions, implement frameworks using COMSOL, MATLAB, and python, generate datasets in bulk, train deep-learning models for the task at hand, and use those models to learn about systems.

In **chapter 2**, the development of a deep-learning model to predict the efficiency of nitrogen reduction reaction (NRR) electrodes will be discussed. Nitrogen conversion to ammonia is itself a valuable reaction to study due to its foundational role in the stability of the world's food supply.[15] Further, wire electrodes provide a proof-of-concept system that can be readily described, simulated, and investigated for their influence on electrode performance – revealing achievable limits on electrode performance (under the set of constraints applied) and fundamental insights into kinetic bottlenecks that seem to limit electrode performance. In a similar vein, but with the goal of inverse design, **chapter 3** will discuss the application of this research framework towards the rational selection of wire array morphologies for specific O_2[16] and H_2O_2[17] concentration gradients with investigation of microbes' interaction with gradients in mind.

For a further exploration of this frameworks' utility, I will present research into its ability to address the problem of mechanism assignment via cyclic voltammetry. Given the foundational role of mechanism assignment in studying electrochemical systems,[18, 19] the nascent research into automated mechanism assignment in cyclic voltammetry,[12] and the established use of deep

learning to analyze analytical data,[8, 20] this workflow was well suited to the establishment of a more powerful deep learning tool for mechanism assignment. In **chapter 4** I will discuss the use of residual neural networks (ResNet) for the assignment of mechanisms to data consisting of five possible electrochemical mechanism classes: *E, EC, CE, DISP1*, and *ECE*. This tool shows strong performance (98.5% overall accuracy) and enables the beginning of automated cyclic voltammetry experiments based on correlations between ResNet SoftMax output and implied contributions of competing mechanistic pathways. In a follow-up, in **chapter 5**, I present a more powerful tool using object detection, specifically Faster R-CNN, for the assignment of multiple mechanism classes to complex cyclic voltammograms constituted of multiple, distinct electrochemical events. This tool approaches more practical, general utility, adding mechanisms representing surface redox events (*SR*), homogenous electrocatalysis (*EC'*), and heterogeneous electrocatalysis (*T*). The object detection approach along with the image classification approach provide two tools that enable easier adoption of cyclic voltammetry by non-domain experts while also supporting high throughput experimentation that often relies on automated data analysis.

Finally, in **chapter 6**, I will present a discussion of deep learning as a means to enhance the utility of feedback provided to instructors in physical science courses. Typically, unstructured feedback is provided to instructors at UCLA, resulting in a difficult to parse dataset ripe for the application of modern data analysis tools.

References

1. Zahrt, A. F.; Henle, J. J.; Rose, B. T.; Wang, Y.; Darrow, W. T.; Denmark, S. E., Prediction of higher-selectivity catalysts by computer-driven workflow and machine learning. *Science* **2019,** *363* (6424), eaau5631.
2. Gómez-Bombarelli, R.; Wei, J. N.; Duvenaud, D.; Hernández-Lobato, J. M.; Sánchez-Lengeling, B.; Sheberla, D.; Aguilera-Iparraguirre, J.; Hirzel, T. D.; Adams, R. P.; Aspuru-Guzik, A., Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules. *ACS Central Science* **2018,** *4* (2), 268-276.
3. Beker, W.; Roszak, R.; Wołos, A.; Angello, N. H.; Rathore, V.; Burke, M. D.; Grzybowski, B. A., Machine Learning May Sometimes Simply Capture Literature Popularity Trends: A Case Study of Heterocyclic Suzuki–Miyaura Coupling. *Journal of the American Chemical Society* **2022,** *144* (11), 4819-4827.
4. Cui, F.; Yue, Y.; Zhang, Y.; Zhang, Z.; Zhou, H. S., Advancing Biosensors with Machine Learning. *ACS Sensors* **2020,** *5* (11), 3346-3364.
5. Wang, M.; Zhu, H., Machine Learning for Transition-Metal-Based Hydrogen Generation Electrocatalysts. *ACS Catalysis* **2021,** *11* (7), 3930-3937.
6. Jiang, B.; Gent, W. E.; Mohr, F.; Das, S.; Berliner, M. D.; Forsuelo, M.; Zhao, H.; Attia, P. M.; Grover, A.; Herring, P. K.; Bazant, M. Z.; Harris, S. J.; Ermon, S.; Chueh, W. C.; Braatz, R. D., Bayesian learning for rapid prediction of lithium-ion battery-cycling protocols. *Joule* **2021,** *5* (12), 3187-3203.
7. Zahrt, A. F.; Mo, Y.; Nandiwale, K. Y.; Shprints, R.; Heid, E.; Jensen, K. F., Machine-Learning-Guided Discovery of Electrochemical Reactions. *Journal of the American Chemical Society* **2022,** *144* (49), 22599-22610.

8. Lin, R.; Zhang, R.; Wang, C.; Yang, X.-Q.; Xin, H. L., TEMImageNet training library and AtomSegNet deep-learning models for high-precision atom segmentation, localization, denoising, and deblurring of atomic-resolution images. *Scientific Reports* **2021,** *11* (1), 5386.

9. Ziatdinov, M.; Dyck, O.; Maksov, A.; Li, X.; Sang, X.; Xiao, K.; Unocic, R. R.; Vasudevan, R.; Jesse, S.; Kalinin, S. V., Deep Learning of Atomically Resolved Scanning Transmission Electron Microscopy Images: Chemical Identification and Tracking Local Transformations. *ACS Nano* **2017,** *11* (12), 12742-12752.

10. Foppa, L.; Sutton, C.; Ghiringhelli, L. M.; De, S.; Löser, P.; Schunk, S. A.; Schäfer, A.; Scheffler, M., Learning Design Rules for Selective Oxidation Catalysts from High-Throughput Experimentation and Artificial Intelligence. *ACS Catalysis* **2022,** *12* (4), 2223-2232.

11. Peurifoy, J.; Shen, Y.; Jing, L.; Yang, Y.; Cano-Renteria, F.; DeLacy, B. G.; Joannopoulos, J. D.; Tegmark, M.; Soljačić, M., Nanophotonic particle simulation and inverse design using artificial neural networks. *Science Advances 4* (6), eaar4206.

12. Kennedy, G. F.; Zhang, J.; Bond, A. M., Automatically Identifying Electrode Reaction Mechanisms Using Deep Neural Networks. *Analytical Chemistry* **2019,** *91* (19), 12220-12227.

13. Chen, Y.; Huang, Y.; Cheng, T.; Goddard, W. A., III, Identifying Active Sites for CO2 Reduction on Dealloyed Gold Surfaces by Combining Machine Learning with Multiscale Simulations. *Journal of the American Chemical Society* **2019,** *141* (29), 11651-11657.

14. Chandrasekaran, A.; Kamal, D.; Batra, R.; Kim, C.; Chen, L.; Ramprasad, R., Solving the electronic structure problem with machine learning. *npj Computational Materials* **2019,** *5* (1), 22.

15. Suryanto, B. H. R.; Du, H.-L.; Wang, D.; Chen, J.; Simonov, A. N.; MacFarlane, D. R., Challenges and prospects in the catalysis of electroreduction of nitrogen to ammonia. *Nature Catalysis* **2019,** *2* (4), 290-296.

16. Lewis, D. M.; Blatchley, M. R.; Park, K. M.; Gerecht, S., O2-controllable hydrogels for studying cellular responses to hypoxic gradients in three dimensions in vitro and in vivo. *Nature Protocols* **2017,** *12* (8), 1620-1638.

17. Virgile, C.; Hauk, P.; Wu, H.-C.; Shang, W.; Tsao, C.-Y.; Payne, G. F.; Bentley, W. E., Engineering bacterial motility towards hydrogen-peroxide. *PLOS ONE* **2018,** *13* (5), e0196999.

18. Allen J. Bard, L. R. F., *Electrochemical Methods Fundamentals and Applications*. 2 ed.; John Wiley & Sons, INC: 2001.

19. Jean-Michel Savéant, C. C., Single-Electron Transfer at an Electrode. In *Elements of Molecular and Biomolecular Electrochemistry*, 2019; pp 1-80.

20. Lee, S.; Oh, J.; Lee, K.; Cho, M.; Paulson, B.; Kim, J. K., Diagnosis of Ischemic Renal Failure Using Surface-Enhanced Raman Spectroscopy and a Machine Learning Algorithm. *Analytical Chemistry* **2022,** *94* (50), 17477-17484.

Chapter 2. The accelerated study of simulated nitrogen fixing wire electrodes using deep learning

This chapter is a version of Hoar, B. B.; Lu, S.; Liu, C. "Machine–Learning Enabled Exploration of Morphology Influence on Wire-Array Electrodes for Electrochemical Nitrogen Fixation." *J. Phys. Chem. Lett.* **2020**, *11(12)*, 4625–4630.

Abstract

Neural networks, trained on data generated by a microkinetic model and finite-element simulations, expand explorable parameter space by significantly accelerating the predictions of electrocatalytic performance. In addition to modelling electrode reactivity, we use micro/nano-wire arrays as a well-defined, easily tuned, and experimentally relevant exemplary morphology for electrochemical nitrogen fixation. This model system provides the data necessary for training neural networks which are subsequently exploited for electrocatalytic material morphology optimizations and explorations into the influence of geometry on nitrogen fixation electrodes, feats untenable without large-scale simulations, on both a global and local basis.

Content

An efficient electrochemical process based on heterogenous catalysts demands fast electrocatalytic turnover with effective mass transport, which leads to a challenging tradeoff for high-surface area electrodes between reactant flux and active site density.[1-3] Experimental, multi-feature optimization of electrode morphology is relatively slow, may depend on the availability of developed synthetic methods, and hence hopes to receive suggestions from computational predictions.[4] Simulations of finite element methods (FEM), often based on relevant and

expeditious microkinetic models,[5-6] predict local concentrations of chemical species, electrode current density distributions, and electrocatalytic performance.[7-11] Yet despite their comparative speed, the repetitive calculations inherent of FEM renders them untenably slow for global optimization problems and other data-intensive questions, which often precludes their synchronization into common experimental practice.

Considering these challenges, we seek to use machine learning[12-13] to utilize FEM's effectiveness in correlating a catalyst's intrinsic kinetic and morphological features with electrocatalytic performance *in silico*[6, 14] – maintaining its utility in inverse design and optimization[7, 15-19]– while avoiding its expensive, repetitive calculations as much as possible. Furthermore, we aim to exploit the neural network's speed[20-21] to guide interrogations of electrode geometry for its relation to electrocatalytic performance to gain insights into design principles.

As an exemplary case, our machine learning method will be instructive in exploring electrodes for ambient, aqueous electrocatalytic nitrogen fixation – a proposed alternative to the prominent, yet environmentally deleterious Haber-Bosch process.[2, 3, 22, 23] Multiple catalytic materials have been reported recently;[24-30] however, their performance as measured by total current density (i_{total}) and Faradaic efficiency (*F.E.*) can be easily diminished by the low solubility of dinitrogen N_2 (0.74 mM in water, 1 bar, 25 ºC) and the local reactant concentration sensitive nature of proton-coupled electron transfer[31-32] during electrocatalysis.[2, 3, 8, 23] In recent years, it has been shown that the performance of electrodes depends highly on this local reactant concentration,[33-36] which is dependent on overpotential (η), electrode morphology, and kinetic predisposition of the material for a reaction of interest. Because of this, one may expect that a non-obvious morphology – one that optimizes reactant availability – exists for each system that maximizes i_{total} and *F.E.*, simultaneously.

In an effort to discover and study such morphologies, we trained two feed-forward[13] multilayer perceptron neural networks (MLPNNs) to accelerate the prediction of optimized wire-array electrodes. Aqueous electrocatalytic nitrogen reduction reaction (NRR) processes on nano/micro-wire array electrodes were selected as a proof-of-concept, due to their exploitation in (photo)electrochemistry[37-39] and accessibility for future experimental validation. To develop an MLPNN-based algorithm, we established a database of i_{total} and *F.E.* data for electrocatalytic NRR on wire arrays (Figure 1A) based on our previously reported microkinetic models of NRR and hydrogen evolution reaction (HER).[8] Therein, the mass transport of chemical species in the electrolyte was simulated under the assumptions of supporting electrolyte[5] and surface electrochemical reactions were simulated using concentration-dependent microkinetic models (Figure 1B, Figure S1).[5, 8, 10] $i_{1,0,HER}$ and $i_{2,0,HER}$, denoted as the exchange current densities for the one-electron Volmer and Heyrovský steps, respectively, were applied as microkinetic variables for HER process (Supplementary Note 1).[5, 8, 10] In order to account for the six-electron NRR process, the model includes another concentration-dependent kinetic equation assuming that either the first electron-transfer or proton-transfer step, whose equivalent exchange current densities are represented as $i_{1,0,NRR}$ and $i_{2,0,NRR}$, respectively, are the only possible turnover-limiting steps in the associative pathway (Supplementary Note 1 and Figure 1B, S1)[3, 8, 23]– an assumption consistent with experimentally reported Tafel slopes that imply an *EC* or *E* mechanism for NRR.[24-30, 40-42] Because these first two steps of the nitrogen fixation mechanisms account for all likely rate-limiting steps, discussion of downstream reaction steps is precluded; however it is worth mentioning that the associative mechanism is generally considered for heterogenous electrocatalysts under ambient conditions.[23, 43-44] Furthermore, our microkinetic model assumes mean-field approximation,[45-46] and does not directly account for surface coverage; however,

surface bound species and their reactivity are implicitly addressed in our approach by using exchange current densities as system-defining parameters. While the constructed model is generally applicable, predictions presented here were simulated under conditions frequently used in experimental investigations (Table S1).[42, 47-49] Microwire/nanowire array morphologies, defined by wire diameter (d), length (l) and array periodicity in a square lattice (p), were also incorporated in the simulation (Figure 1C). FEM simulations using COMSOL Multiphysics version 5.3 were conducted to yield values of i_{total} and *F.E.* as a function of feature variable vector $\boldsymbol{x} = (\eta, d, l, p, i_{1,0,NRR}, i_{2,0,NRR}, i_{1,0,HER}, i_{2,0,HER})$ in 0.1 M HCl under 1 bar N_2, 25 ºC,[24-30, 40-42] in which overpotential (η) is referenced to the reversible hydrogen electrode (RHE). About 200,000 FEM simulations with different $\boldsymbol{x}$ were conducted, establishing a dataset sparsely mapping the experimentally accessible parameter spaces (Table S2) due to the large degree-of-freedom in the feature variable $\boldsymbol{x}$ – highlighting the necessity of our alternative approach for optimal morphology discovery.

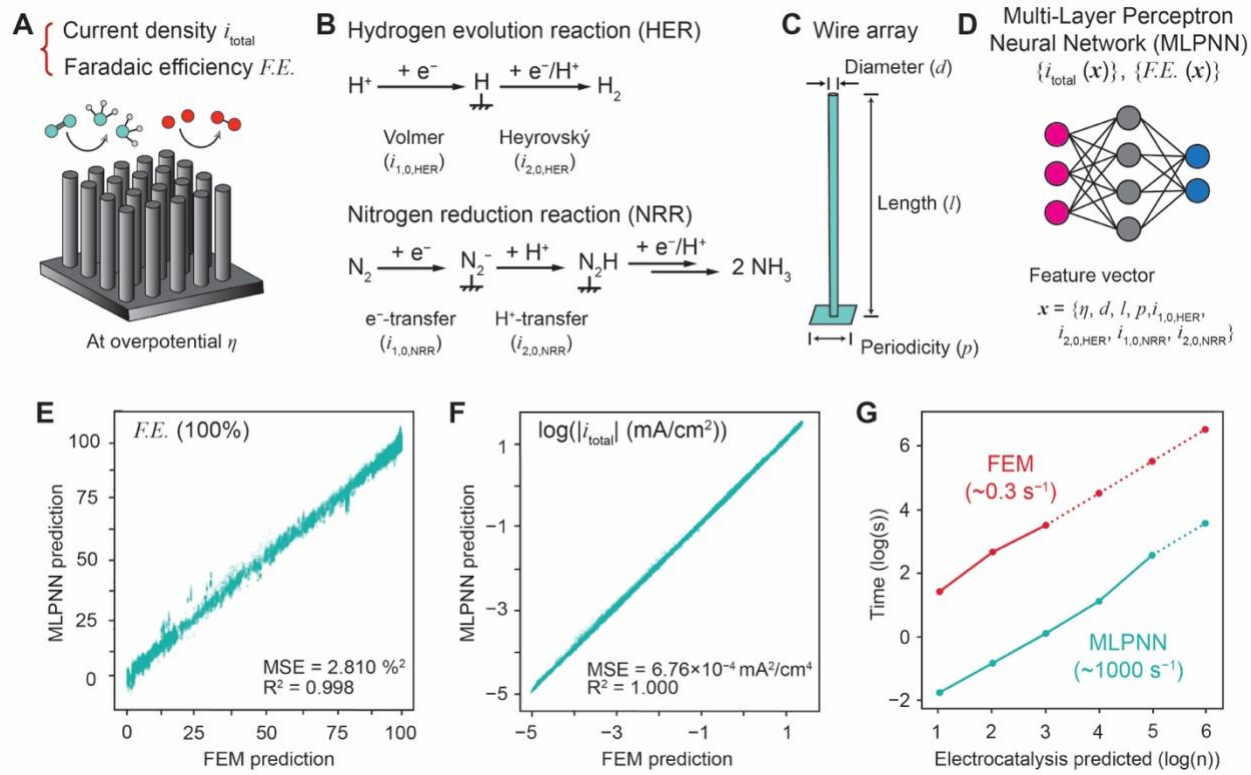


Figure 1. The microkinetic model includes **A** measures of efficacy, **B** microkinetic models of both HER and NRR, **C** and electrode geometry definitions. MLPNNs were trained and validated by FEM-based datasets $\{F.E.(\boldsymbol{x})\}$ and $\{i_{\text{total}}(\boldsymbol{x})\}$with $\boldsymbol{x}$ as the feature vector **D**. **E** and **F** show correlation between FEM and MLPNN approaches as measured by mean squared error (MSE) and linearity (R^2). **G**, Double-logarithmical plot between simulation time and number of predictions for the FEM and MLPNN approaches (Supplementary Note 2).

Using our generated FEM datasets for $\{F.E.(\boldsymbol{x})\}$ and $\{i_{\text{total}}(\boldsymbol{x})\}$, we trained two MLPNNs with $\boldsymbol{x}$ as the feature input (Figure 1D, model details in Table S3). The two trained neural networks yielded values of i_{total} and *F.E.* that were validated on a hold-out test set (Supplementary Information, Materials and Methods). The yielded MLPNNs displayed satisfying prediction consistency, both qualitatively shown in Figure 1E & 1F and quantitatively described by their mean square error (MSE) and coefficients of determination (R^2) as shown in the inset of Figures 1E and F.

As indicated by their small MSE's and near unity R^2 values, the developed neural networks provide accurate predictions and simultaneously alleviate the time burden of FEM calculations, amplifying the power of existing simulation method.[7, 17, 51] Figure 1G compares the computational time against the numbers of accomplished predictions of i_{total} and *F.E.* for two approaches: the one based on our MLPNNs on a 2014 Macbook Air laptop (green curve) and the other one based on FEM simulations on a powerful workstation (red curve) (hardware details Table S4). The neural network approach yielded about 1,000 outputs per second, while the one utilizing FEM software was at a rate of about 0.3 outputs per second (Supplementary Note 2). Thanks to the fast calculation of the neural network approach, within hours we can now generate large datasets, denoted as $\{F.E._{\text{NN}}(\boldsymbol{x})\}$ and $\{i_{\text{total,NN}}(\boldsymbol{x})\}$, predicting millions of electrocatalysis electrodes in short order.

Exploiting this speed boost provided by MLPNN, we sought to answer a more applied yet data-intensive question: what are the near-global optima of electrocatalytic NRR activities assuming control over both the intrinsic activities and geometry of wire array electrodes? To find an approximate "ultimate" value of NRR for our conditions, we generated a greatly expanded $\{F.E._{\mathrm{NN}}(\boldsymbol{x})\}$ and $\{i_{\mathrm{total,NN}}(\boldsymbol{x})\}$ of over 4,000,000 scenarios (Supplementary Data 1) within a few hours on a laptop (*vide supra*), which otherwise would cost an estimated thirty-four days via the FEM methods on a powerful workstation. For convenience, all of our machine-learning studies operate with a geometric resolution of 1 μm, which provided smooth changes in electrode output predictions and was considered satisfactory for the purposes of this investigation. Nonetheless, there is no apparent technical hurdle towards an increase of geometric resolution and the authors acknowledge the necessity of higher geometric resolution for other studies that involves finer or individual nanoparticles. We found over 200 scenarios whose *F.E.* > 90% and $|i_{\mathrm{total}}| > 2$ mA/cm^2 (Figure 2A, Supplementary Data 2). In Figure 2A, a tradeoff between i_{total} and *F.E.* was noted and a dashed line in red was added to illustrate a putative limit of NRR electrocatalysis under our simulated experimental condition. Conditions for four of the top scenarios, circled in Figure 2A, are shown in Table S5. It is interesting to note that all of these top scenarios somewhat converge to a similar range of parameter values: $\eta \approx -0.7$ V vs. RHE, $d \approx 8$ μm, $l \approx 40$ μm, $p \approx 10$ μm, $i_{1,0,\mathrm{NRR}} \approx 10^{-3}$, $i_{2,0,\mathrm{NRR}} \approx 10^{-3}$, $i_{1,0,\mathrm{HER}} \approx 10^{-8}$, $i_{2,0,\mathrm{HER}} \approx 10^{-8}$ mA/cm^2. This is reasonable for the catalyst's kinetic reactivity as HER is greatly suppressed as compared to NRR; yet, it is also impressive that there seems to be an non-obvious optimal geometry for wire array electrode, indicative to a tradeoff between reactant availability and the density of active sites.[1-3] Since we were unable to find a real NRR electrocatalyst that possess the above properties in both literature and our own experimental efforts,[2, 3, 22, 23] an experimental validation of such a claim was precluded. Nonetheless, we further

showcase the scope of answerable questions enabled by the employment of machine learning. We deduced from literature the values of catalysts' intrinsic kinetic information ($i_{1,0,NRR}$, $i_{2,0,NRR}$, $i_{1,0,HER}$, and $i_{2,0,HER}$) (Supplementary Information, Materials and Methods),[42, 47, 50, 52, 53] and predicted hypothetical wire array geometries of those reported catalysts that increased both *F.E.* and i_{total} (Figure 2B, catalyst details Table S6) or either *F.E.* and i_{total} , independently (Figure S2). Depending on whether the target of optimization is maximized *F.E.*, i_{total}, or both, the neural-network model suggests different wire array geometries catering to each reported catalyst, which are predicted to yield significantly higher figure-of-merits. In the case of Au_6/Ni catalyst that was claimed to possess a *F.E.* = 67.8% and i_{total} = 0.21 mA/cm^2,[52] optimized wire array geometry is predicted to simultaneously achieve *F.E.* ≈ 100% and i_{total} ≈ 0.37 mA/cm^2 (Figure 2B), much improved compared to the original values,[52] when $\eta = -0.3$ V vs. RHE, $d \approx 8.5$ μm, $l \approx 80$ μm, $p \approx 23$ μm (Table S6). We note that our results give no indication that slight deviations from these recommended parameters upon fabrication would drastically reduce performance as highly similar morphologies provide very similar outputs. Although the prediction remains to be experimentally validated, this example illustrates how inverse-design can be applied for electrocatalysis development, a tool to maximize the efficacy of well-defined morphologies before exploring other, more exotic options.

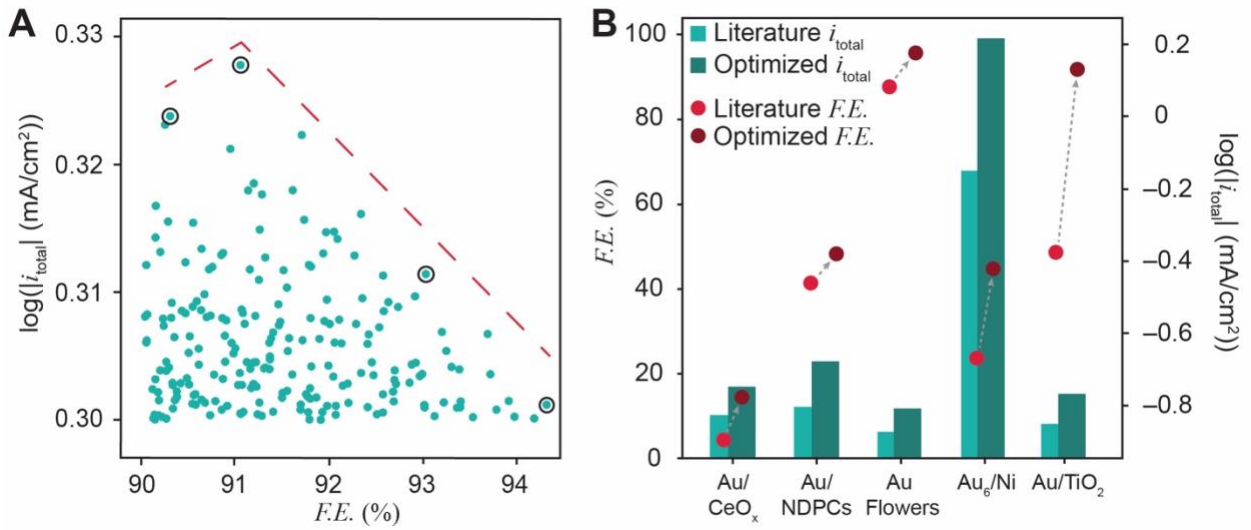


Figure 2. A, A selection of hypothetical NRR electrocatalysts ($i_{total} \geq 2$ mA/cm^2 and $F.E. \geq 90\%$) among over 4,000,000 scenarios generated by trained MLPNNs. Details of the circled points are shown in Table S5. Dotted line putatively denotes a subjective boundary of achievable performances. **B**, Predicted and reported performances of literature catalysts. Details in Table S6 for the cited NRR electrocatalysts.[42, 47, 50, 52, 53] Grey dotted line denotes a visual aid.

In addition to mitigating time-to-prediction challenges and allowing for such optimizations, our approach affords insights into feature influence on electrode efficacy – particularly for boosting *F.E.* for electrocatalytic NRR[2, 3, 23] – and subsequently directs further MLPNN and FEM investigation to elucidate how morphology impacts Faradaic efficiency (*F. E.*). As our first means of exploring model feature importance[20, 21, 54] in determining electrocatalytic performance, we iteratively randomized and permuted all values of a single feature, retrained the neural networks using the same hyperparameters as the original models, and calculated the change in the MSE with the data-shuffled model's predictions versus the initial one, thus deducing the relative influence of each feature on both *F.E.* (Figure 3A) and i_{total} (Figure S3). Unsurprisingly, η and kinetic descriptors for HER and NRR pose significant influence over the value of *F.E.*; however, some variables in the microkinetic model are significantly more important than other

variables. The value of $i_{2,0,NRR}$, indicating the kinetics of first proton transfer in NRR, is intriguingly the most prominent variable among all features. Given the weak Lewis basicity of N_2, our analysis suggests that the kinetics of proton transfer is the bottleneck towards a high *F.E.* value in electrocatalytic NRR, at least within our parameter space and simulated experimental conditions. We note that among the morphological features controlling reactant availability[7-11], the wire length, *l*, poses the smallest influence over *F.E.*. Such non-obvious co-influence on electrocatalysis by wire array's geometry highlight that all morphological parameters are consequential and ideally should be experimentally optimized simultaneously, not more intuitively one-by-one.

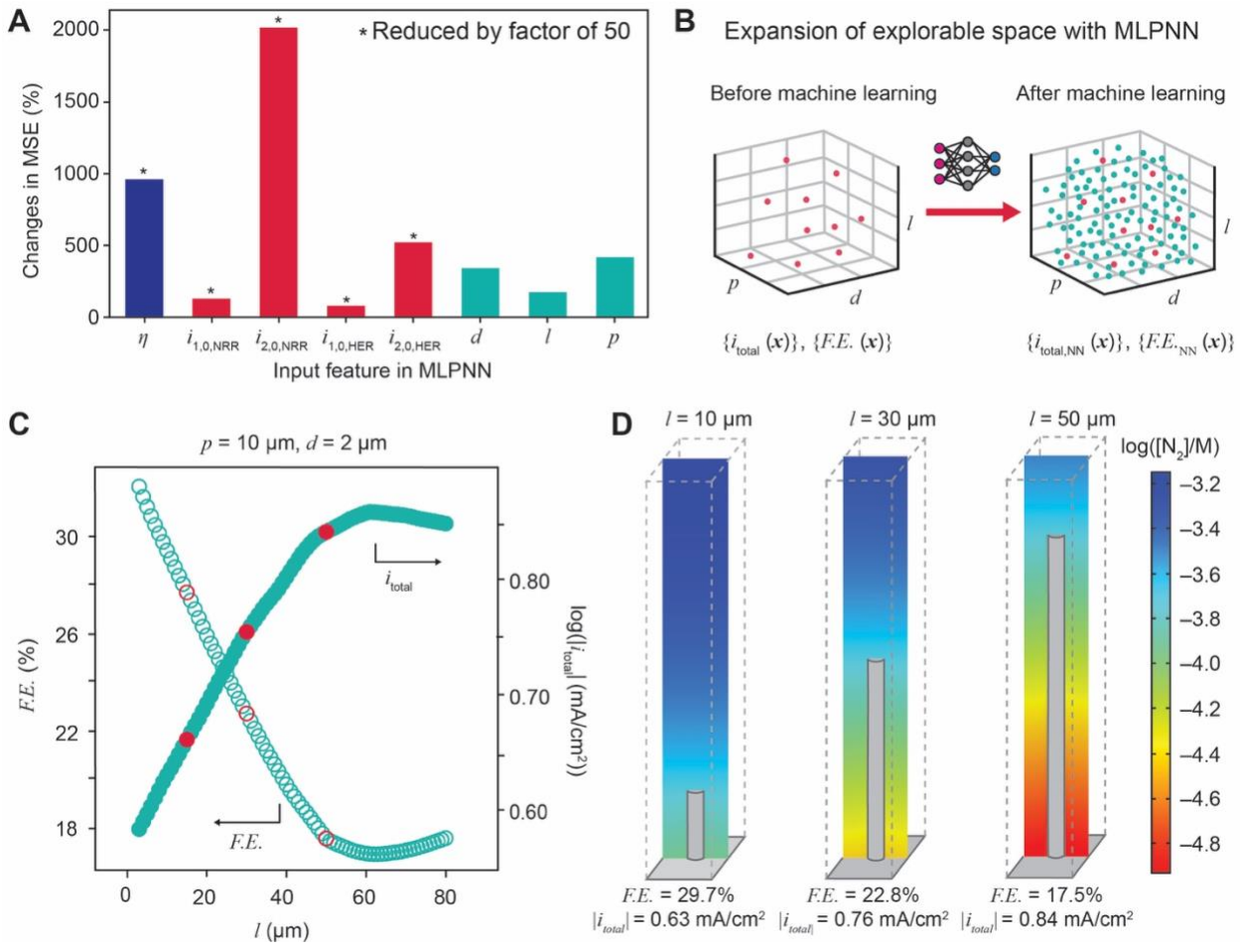


Figure 3. A, Data shuffling, retraining, and resultant change in MSE as a measure of feature importance for *F.E.* in NRR. **B**, MLPNN-based algorithm allows high-density sampling of parameter space. **C**, Values of i_{total} and *F.E.* versus length for a region sensitive to wire length *l*. Red circles indicate lengths of wires present in MLPNN training data. **D**, FEM-derived concentration profiles of N_2 for varying values of *l*. η = −0.5 V vs. RHE, $i_{1,0,NRR} = 10^{-6}$, $i_{2,0,NRR} = 10^{-3}$, $i_{1,0,HER} = 10^{-4}$, and $i_{2,0,HER} = 10^{-6}$ mA/cm^2 in **C** and **D**.

Expanding on these machine-learning enabled insights, we found that the influence of geometric parameters are sometimes exceptionally consequential despite their relatively low influence on average, a phenomenon elucidated only upon the generation of large amounts of data (Figure 3B, S4) and an occurrence elusive to sparse mapping of parameter space.[8] In one instance, we found that the geometric parameters (*d*, *l*, and *p*) have significant impacts on i_{total} and *F.E* (Figure S5) when η = −0.5 V vs. RHE, $i_{1,0,NRR} = 10^{-6}$, $i_{2,0,NRR} = 10^{-3}$, $i_{1,0,HER} = 10^{-4}$, and $i_{2,0,HER} = 10^{-6}$ mA/cm^2. As shown in Figures 3C, modifications of wire array length alone will alter the *F.E.* values between ~15% to ~35% (a 133% marginal increase), a range relevant to the reported performance in recent literature.[24-30] Noteworthy of this material is *l*'s exceptional influence on electrode performance despite its seemingly minimal importance implied by Figure 3A, a demonstration of how features should be considered holistically, instead of independently, when designing electrocatalysts' morphology. To probe this phenomenon for regions of high morphology dependence, we generated concentration profiles for both N_2 and H^+ using our FEM simulations. Shown in Figure 3D are FEM-derived concentration profiles (Supplementary Materials and Methods) for N_2 for the same catalytic material discussed in Figure 3C, with wire arrays of varying wire lengths (*p* = 10 μm, *d* = 2 μm). The profiles displayed in Figure 3D are in contrast to the much weaker variations of H^+ concentration profiles (Figure S9). As shown here and in other cases (Figures S6-8), wire geometry can shift the availability of N_2 by orders-of-

magnitude from the top of the wire to its base, and significantly affect the *F.E.*. This necessitates a careful selection of wire morphology unique to a given set of kinetic descriptors, potentially fulfilled by machine learning given the large degree-of-freedom and morphologies' convoluted impacts (*vide supra*), in order to balance a boost in *F.E.* with the often-coincident loss of i_{total} and active surface area.

Conclusion

In summary, our machine-learning approach not only provides the means to quickly optimize wire-array morphology for a given catalytic material, but also explores how electrode's morphology contributes to electrocatalytic performance based on global relative feature importance and local individual scenarios. The reported insights enabled by the fast prediction and the generation of large datasets will offer real-time guidance and reasoning of effective morphology optimizations on an affordable laptop, an important synchronization between experiments and simulations for the development of advanced electrocatalysis in general.

References

1. Birdja, Y. Y.; Pérez-Gallent, E.; Figueiredo, M. C.; Göttle, A. J.; Calle-Vallejo, F.; Koper, M. T. M. Advances and challenges in understanding the electrocatalytic conversion of carbon dioxide to fuels. *Nat. Energy* **2019,** *4* (9), 732−745.
2. Suryanto, B. H. R.; Du, H.-L.; Wang, D.; Chen, J.; Simonov, A. N.; MacFarlane, D. R. Challenges and prospects in the catalysis of electroreduction of nitrogen to ammonia. *Nat. Catal.* **2019,** *2* (4), 290−296.
3. Deng, J.; Iñiguez, J. A.; Liu, C. Electrocatalytic Nitrogen Reduction at Low Temperature. *Joule* **2018,** *2* (5), 846−856.
4. Collins, K. D.; Gensch, T.; Glorius, F. Contemporary screening approaches to reaction discovery and development. *Nat. Chem.* **2014,** *6* (10), 859−871.
5. Newman J, T.-A., K.E. *Electrochemical Systems*. 3rd ed.; Wiley: 2004.
6. Logan, D. L. *A First Course in the Finite Element Method*. 6th ed.; Cengage Learning: 2016.
7. Chen, Y.; Huang, Y.; Cheng, T.; Goddard, W. A. Identifying Active Sites for CO_2 Reduction on Dealloyed Gold Surfaces by Combining Machine Learning with Multiscale Simulations. *J. Am. Chem. Soc.* **2019,** *141* (29), 11651−11657.
8. Lu, S.; Lee, D. H.; Liu, C. Modeling of Electrocatalytic Dinitrogen Reduction on Microstructured Electrodes. *Small Methods* **2019,** *3* (6), 1800332.
9. Hashiba, H.; Weng, L.-C.; Chen, Y.; Sato, H. K.; Yotsuhashi, S.; Xiang, C.; Weber, A. Z. Effects of Electrolyte Buffer Capacity on Surface Reactant Species and the Reaction Rate of CO_2 in Electrochemical CO_2 Reduction. *J. Phys. Chem. C* **2018,** *122* (7), 3719−3726.

10. Singh, M. R.; Goodpaster, J. D.; Weber, A. Z.; Head-Gordon, M.; Bell, A. T. Mechanistic insights into electrochemical reduction of CO_2 over Ag using density functional theory and transport models. *Proc. Natl. Acad. Soc.* **2017,** *114* (42), E8812−E8821.
11. Singh, M. R.; Clark, E. L.; Bell, A. T. Effects of electrolyte, catalyst, and membrane composition and operating conditions on the performance of solar-driven electrochemical reduction of carbon dioxide. *Phys. Chem. Chem. Phys.* **2015,** *17* (29), 18924−18936.
12. Burkov, A. *The Hundred-Page Machine Learning Book*. 2019.
13. Ian Goodfellow, Y. B., Aaron Courville. *Deep Learning*. MIT Press: 2016.
14. Britz, D. S., J. *Digital Simulation in Electrochemistry*. 4th ed.; Springer: 2016.
15. Gómez-Bombarelli, R.; Wei, J. N.; Duvenaud, D.; Hernández-Lobato, J. M.; Sánchez-Lengeling, B.; Sheberla, D.; Aguilera-Iparraguirre, J.; Hirzel, T. D.; Adams, R. P.; Aspuru-Guzik, A. Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules. *ACS Cent. Sci.* **2018,** *4* (2), 268−276.
16. Palkovits, R.; Palkovits, S. Using Artificial Intelligence To Forecast Water Oxidation Catalysts. *ACS Catal.* **2019**, 8383−8387.
17. Mansouri Tehrani, A.; Oliynyk, A. O.; Parry, M.; Rizvi, Z.; Couper, S.; Lin, F.; Miyagi, L.; Sparks, T. D.; Brgoch, J. Machine Learning Directed Search for Ultraincompressible, Superhard Materials. *J. Am. Chem. Soc.* **2018,** *140* (31), 9844−9853.
18. Zahrt, A. F.; Henle, J. J.; Rose, B. T.; Wang, Y.; Darrow, W. T.; Denmark, S. E. Prediction of higher-selectivity catalysts by computer-driven workflow and machine learning. *Science* **2019,** *363* (6424), eaau5631.

19. Peurifoy, J.; Shen, Y.; Jing, L.; Yang, Y.; Cano-Renteria, F.; DeLacy, B. G.; Joannopoulos, J. D.; Tegmark, M.; Soljačić, M. Nanophotonic particle simulation and inverse design using artificial neural networks. *Sci. Adv.* **2018,** *4* (6), eaar4206.

20. Jerome H. Friedman, R. T., Trevor Hastie. *The Elements of Statistical Learning*. 2009.

21. Altmann, A.; Toloşi, L.; Sander, O.; Lengauer, T. Permutation importance: a corrected feature importance measure. *Bioinformatics* **2010,** *26* (10), 1340−1347.

22. Andersen, S. Z.; Čolić, V.; Yang, S.; Schwalbe, J. A.; Nielander, A. C.; McEnaney, J. M.; Enemark-Rasmussen, K.; Baker, J. G.; Singh, A. R.; Rohr, B. A.; Statt, M. J.; Blair, S. J.; Mezzavilla, S.; Kibsgaard, J.; Vesborg, P. C. K.; Cargnello, M.; Bent, S. F.; Jaramillo, T. F.; Stephens, I. E. L.; Nørskov, J. K.; Chorkendorff, I. A rigorous electrochemical ammonia synthesis protocol with quantitative isotope measurements. *Nature* **2019,** *570* (7762), 504−508.

23. Foster, S. L.; Bakovic, S. I. P.; Duda, R. D.; Maheshwari, S.; Milton, R. D.; Minteer, S. D.; Janik, M. J.; Renner, J. N.; Greenlee, L. F. Catalysts for nitrogen reduction to ammonia. *Nat. Catal.* **2018,** *1* (7), 490−500.

24. Geng, Z.; Liu, Y.; Kong, X.; Li, P.; Li, K.; Liu, Z.; Du, J.; Shu, M.; Si, R.; Zeng, J. Achieving a Record-High Yield Rate of 120.9 for N_2 Electrochemical Reduction over Ru Single-Atom Catalysts. *Adv. Mater.* **2018,** *30* (40), 1803498.

25. Chen, G.-F.; Cao, X.; Wu, S.; Zeng, X.; Ding, L.-X.; Zhu, M.; Wang, H. Ammonia Electrosynthesis with High Selectivity under Ambient Conditions via a Li^+ Incorporation Strategy. *J. Am. Chem. Soc.* **2017,** *139* (29), 9771−9774.

26. Li, L.; Tang, C.; Xia, B.; Jin, H.; Zheng, Y.; Qiao, S.-Z. Two-Dimensional Mosaic Bismuth Nanosheets for Highly Selective Ambient Electrocatalytic Nitrogen Reduction. *ACS Catal.* **2019,** *9* (4), 2902−2908.

27. Han, L.; Liu, X.; Chen, J.; Lin, R.; Liu, H.; Lü, F.; Bak, S.; Liang, Z.; Zhao, S.; Stavitski, E.; Luo, J.; Adzic, R. R.; Xin, H. L. Atomically Dispersed Molybdenum Catalysts for Efficient Ambient Nitrogen Fixation. *Angew. Chem. Int. Ed.* **2019,** *58* (8), 2321−2325.

28. Cao, N.; Chen, Z.; Zang, K.; Xu, J.; Zhong, J.; Luo, J.; Xu, X.; Zheng, G. Doping strain induced bi-Ti_3^+ pairs for efficient N_2 activation and electrocatalytic fixation. *Nature Commun.* **2019,** *10* (1), 2877.

29. Wang, Y.; Shi, M.-m.; Bao, D.; Meng, F.-l.; Zhang, Q.; Zhou, Y.-t.; Liu, K.-h.; Zhang, Y.; Wang, J.-z.; Chen, Z.-w.; Liu, D.-p.; Jiang, Z.; Luo, M.; Gu, L.; Zhang, Q.-h.; Cao, X.-z.; Yao, Y.; Shao, M.-h.; Zhang, Y.; Zhang, X.-B.; Chen, J. G.; Yan, J.-m.; Jiang, Q. Generating Defect-Rich Bismuth for Enhancing the Rate of Nitrogen Electroreduction to Ammonia. *Angew. Chem. Ed.* **2019,** *58* (28), 9464−9469.

30. Wang, J.; Yu, L.; Hu, L.; Chen, G.; Xin, H.; Feng, X. Ambient ammonia synthesis via palladium-catalyzed electrohydrogenation of dinitrogen at low overpotential. *Nat. Commun.* **2018,** *9* (1), 1795.

31. Costentin, C. Electrochemical Approach to the Mechanistic Study of Proton-Coupled Electron Transfer. *Chem. Rev.* **2008,** *108* (7), 2145−2179.

32. Huynh, M. H. V.; Meyer, T. J. Proton-Coupled Electron Transfer. *Chem. Rev.* **2007,** *107* (11), 5004−5064.

33. Hall, A. S.; Yoon, Y.; Wuttig, A.; Surendranath, Y. Mesostructure-Induced Selectivity in CO_2 Reduction Catalysis. *J. Am. Chem. Soc.* **2015,** *137* (47), 14834−14837.

34. Xie, C.; Niu, Z.; Kim, D.; Li, M.; Yang, P. Surface and Interface Control in Nanoparticle Catalysis. *Chem. Rev.* **2020,** *120* (2), 1184−1249.

35. Yoon, Y.; Hall, A. S.; Surendranath, Y. Tuning of Silver Catalyst Mesostructure Promotes Selective Carbon Dioxide Conversion into Fuels. *Angew. Chem. Int. Ed.* **2016,** *55* (49), 15282−15286.

36. Goyal, A.; Marcandalli, G.; Mints, V. A.; Koper, M. T. M. Competition between CO_2 Reduction and Hydrogen Evolution on a Gold Electrode under Well-Defined Mass Transport Conditions. *J. Am. Chem. Soc.* **2020,** *142* (9), 4154−4161.

37. Deng, J.; Su, Y.; Liu, D.; Yang, P.; Liu, B.; Liu, C. Nanowire Photoelectrochemistry. *Chem. Rev.* **2019,** *119* (15), 9221−9259.

38. Zhou, G.; Xu, L.; Hu, G.; Mai, L.; Cui, Y. Nanowires for Electrochemical Energy Storage. *Chem. Rev.* **2019,** *119* (20), 11042−11109.

39. Boettcher, S. W.; Spurgeon, J. M.; Putnam, M. C.; Warren, E. L.; Turner-Evans, D. B.; Kelzenberg, M. D.; Maiolo, J. R.; Atwater, H. A.; Lewis, N. S. Energy-Conversion Properties of Vapor-Liquid-Solid–Grown Silicon Wire-Array Photocathodes. *Science* **2010,** *327* (5962), 185−187.

40. Savéant, J.-M. *Elements of Molecular and Biomolecular Electrochemistry: An Electrochemical Approach to Electron Transfer Chemistry*. John Wiley & Sons, Inc.: 2006.

41. Bao, D.; Zhang, Q.; Meng, F.-L.; Zhong, H.-X.; Shi, M.-M.; Zhang, Y.; Yan, J.-M.; Jiang, Q.; Zhang, X.-B. Electrochemical Reduction of N_2 under Ambient Conditions for Artificial N2 Fixation and Renewable Energy Storage Using N_2/NH_3 Cycle. *Adv. Mater.* **2017,** *29* (3), 1604799.

42. Shi, M.-M.; Bao, D.; Wulan, B.-R.; Li, Y.-H.; Zhang, Y.-F.; Yan, J.-M.; Jiang, Q. Au Sub-Nanoclusters on TiO_2 toward Highly Efficient and Selective Electrocatalyst for N_2 Conversion to NH_3 at Ambient Conditions. *Adv. Mater.* **2017,** *29* (17), 1606550.

43. Chen, A.; Xia, B. Y. Ambient dinitrogen electrocatalytic reduction for ammonia synthesis. *J. Mater. Chem. A* **2019,** *7* (41), 23416−23431.

44. Cui, X.; Tang, C.; Zhang, Q. A Review of Electrocatalytic Reduction of Dinitrogen to Ammonia under Ambient Conditions. *Adv. Energy Mater.* **2018,** *8* (22), 1800369.

45. Neurock, M. The microkinetics of heterogeneous catalysis. By J. A. Dumesic, D. F. Rudd, L. M. Aparicio, J. E. Rekoske, and A. A. Treviño, ACS Professional Reference Book, American Chemical Society, Washington, DC, 1993, 315 pp. *AIChE Journal* **1994,** *40* (6), 1085−1087.

46. Chen, Z.; Wang, H.; Su, N. Q.; Duan, S.; Shen, T.; Xu, X. Beyond Mean-Field Microkinetics: Toward Accurate and Efficient Theoretical Modeling in Heterogeneous Catalysis. *ACS Catal.* **2018,** *8* (7), 5816−5826.

47. Li, S.-J.; Bao, D.; Shi, M.-M.; Wulan, B.-R.; Yan, J.-M.; Jiang, Q. Amorphizing of Au Nanoparticles by CeO_x–RGO Hybrid Support towards Highly Efficient Electrocatalyst for N_2 Reduction under Ambient Conditions. *Adv. Mater.* **2017,** *29* (33), 1700001.

48. Qiu, W.; Xie, X.-Y.; Qiu, J.; Fang, W.-H.; Liang, R.; Ren, X.; Ji, X.; Cui, G.; Asiri, A. M.; Cui, G.; Tang, B.; Sun, X. High-performance artificial nitrogen fixation at ambient conditions using a metal-free electrocatalyst. *Nat. Commun.* **2018,** *9* (1), 3485.

49. Lv, C.; Yan, C.; Chen, G.; Ding, Y.; Sun, J.; Zhou, Y.; Yu, G. An Amorphous Noble-Metal-Free Electrocatalyst that Enables Nitrogen Fixation under Ambient Conditions. *Angew. Chem. Int. Ed.* **2018,** *57* (21), 6073−6076.

50. Wang, Z.; Li, Y.; Yu, H.; Xu, Y.; Xue, H.; Li, X.; Wang, H.; Wang, L. Ambient Electrochemical Synthesis of Ammonia from Nitrogen and Water Catalyzed by Flower-Like Gold Microstructures. *ChemSusChem* **2018,** *11* (19), 3480−3485.

51. Chandrasekaran, A.; Kamal, D.; Batra, R.; Kim, C.; Chen, L.; Ramprasad, R. Solving the electronic structure problem with machine learning. *npj Comput. Mater.* **2019,** *5* (1), 22.

52. Xue, Z.-H.; Zhang, S.-N.; Lin, Y.-X.; Su, H.; Zhai, G.-Y.; Han, J.-T.; Yu, Q.-Y.; Li, X.-H.; Antonietti, M.; Chen, J.-S. Electrochemical Reduction of N_2 into NH_3 by Donor–Acceptor Couples of Ni and Au Nanoparticles with a 67.8% Faradaic Efficiency. *J. Am. Chem. Soc.* **2019,** *141* (38), 14976−14980.

53. Qin, Q.; Heil, T.; Antonietti, M.; Oschatz, M. Single-Site Gold Catalysts on Hierarchical N-Doped Porous Noble Carbon for Enhanced Electrochemical Reduction of Nitrogen. *Small Methods* **2018,** *2* (12), 1800202.

54. Ahneman, D. T.; Estrada, J. G.; Lin, S.; Dreher, S. D.; Doyle, A. G. Predicting reaction performance in C–N cross-coupling using machine learning. *Science* **2018,** *360* (6385), 186−190.

Supplementary Information

Materials and Methods

Finite Element Methods

Finite element methods (FEM) for simulating nitrogen reduction reaction (NRR) and hydrogen evolution reaction (HER) activity of electrodes were conducted using COMSOL Multiphysics (Version 5.3) (Figure 1A to 1C). The basis of the FEM model used here was reported previously as a study of micro-wire NRR electrode efficacy and its concomitant mass transport complications.[1] Electrode activity was calculated assuming aqueous non-oxidizing strong acid electrolyte with a concentration of 0.1 M at 293 K with other invariant FEM simulation parameters shown in Table S1.[2-4] Other parameters of the FEM were variable and consisted of morphology descriptors (wire geometry: periodicity *p*, length *l*, diameter *d*), applied potential (η vs. Reversible Hydrogen Electrode, RHE), and catalytic activity descriptors (exchange current densities of mechanisms shown in Figure 1B: $i_{1,0,NRR}$, $i_{2,0,NRR}$, $i_{1,0,HER}$, $i_{2,0,HER}$). Combinations of these descriptors were implemented using parametric sweep (Table S2).

The 3D model constructed for the above descriptors also accounted for the diffusion layer relevant to the electrochemical NRR system, the layer being defined by the distance between the highest point of the micro-wire and the diffusion layer boundary. This FEM model consisted of a micro-wire whose unit cell's base was defined by the periodicity of the electrode wires and whose height was defined by the length of the microwire plus the diffusion layer thickness. Of this model, one surface represented the electrode and the one opposite to this surface represented the diffusion layer boundary, beyond where reagents existed in their bulk electrolyte concentrations. For all

other surfaces, periodic boundary conditions were applied. Migration effects were assumed to be minimal due to supporting electrolyte effects and only diffusion was considered.

The nitrogen fixation reaction in acidic solution is given by:

$$N_2(aq) + 6e^- + 8H^+ \rightarrow 2NH_4^+(aq) \qquad \textbf{(S1)}$$

The kinetic rate description of the reaction is given by equation S8 provided in Supplementary Note 1. The parameters explored unique to this reaction were $i_{1,0,NRR}$ and $i_{2,0,NRR}$. Applied potential and morphological parameters were also explored in combination with these kinetic descriptors and could also influence the kinetics of the reaction.

The hydrogen evolution reaction in acidic solution is given by:

$$2H^+(aq) + 2e^- \rightarrow H_2(aq) \qquad \textbf{(S2)}$$

The kinetic rate description of the reaction is given by equation S12 provided in Supplementary Note 1. The parameters explored unique to this reaction were $i_{1,0,HER}$ and $i_{2,0,HER}$. Other swept parameters could influence HER kinetics as described for NRR.

The ranges of exchange current densities for NRR and HER activity used were 10^{-8} to 10^{-3} mA/cm^2 based on previously reported experimental data.[5-6]

The total current density (i_{total}) is calculated as:

$$i_{total} = \frac{I_{total}}{A} = \frac{I_{NRR} + I_{HER}}{A} = \frac{\int_A i_{loc,NRR}\, dA + \int_A i_{loc,HER}\, dA}{A} \qquad \textbf{(S3)}$$

here, A denotes the electrode surface area and I_{total}, I_{NRR}, and I_{HER} denote total current passing through the electrode, the current used for NRR and the current used for HER, respectively.

Faradaic efficiency (*F.E.)*, which is the proportion of i_{total} accounted for by $i_{loc,\,NRR}$, is given by:

$$F.E. = \frac{I_{NRR}}{I_{total}} = \frac{\int_A i_{loc,NRR}\, dA}{\int_A i_{loc,NRR}\, dA + \int_A i_{loc,HER}\, dA} \qquad \textbf{(S4)}$$

Model selection and training

The implementation of all machine-learning code was done on Google Cloud Platform's (GCP) AI Platform, with code specifically deployed with the beta version of their JupyterLab Notebook implementation. This allowed for virtual machines with adjustable power depending on the demands of the contemporary goal. All machine-learning methods were implemented using the scikit-learn machine-learning package (version 0.21.2) with data management largely provided by the pandas software package (version 0.24.2) as well as graphing provided by the matplotlib software package (version 3.1.0). For convenience all of our machine-learning studies arbitrarily operate with a geometric resolution of 1 μm, which can be improved in future investigations.

Data was imported from FEM results in the form of CSV files and combined into library of data. Before model development began, a random 30% of the data was split from the overall dataset and set aside for later model validation. The rest of the data was normalized (scikit-learn StandardScaler) and scaled (scikit-learn MinMaxScaler) in preparation for implementation of machine learning algorithms. Model selection proceeded via a script combining a list of candidate models with coarse grid search methods to provide options for final model deployment. Model success was evaluated using a mean squared error (MSE) loss function and built-in three-fold cross-validation.

Upon completion of model selection procedures, scikit-learn's multi-layer perceptron (MLPNN) neural network provided the most promising results. The hyperparameters of the MLPNN is displayed in Table S3. A finer grid search of MLPNN hyper-parameters was conducted to establish the final model with the early-stopping hyper-parameter set to "true" as our primary means to avoid model over-fitting.

Evaluation of model performance

The quality of the models' fit was further validated using the 30% of data separated pre-model selection. Model validity was measured using the MSE of our neural network predicted values versus FEM predictions. Testing data and training data covered the same parameter space and provided model evaluation as it related to parameters previously seen by the model (Figures 1E-F).

A sensitivity analysis was performed to determine the individual importance of each feature by one-at-a-time randomization of features, training the model with dataset containing the randomized feature, and re-evaluation of model predictive power (Figure 3A). Importance was determined by comparing the MSE of the data-altered model and the model trained on the unaltered data.

Evaluation of model's ability to predict electrode activity in regions with high morphology influence was explored by selecting a catalyst (defined by values: $i_{1,0,NRR} = 10^{-6}$ mA/cm^2, $i_{2,0,NRR} = 10^{-3}$ mA/cm^2, $i_{1,0,HER} = 10^{-4}$ mA/cm^2, $i_{2,0,HER} = 10^{-6}$ mA/cm^2) and potential value ($\eta = -0.5$ V vs. RHE) where a wide range of Faradaic efficiency values were present as a function of morphology change alone (Figure 3C, S5). Data was generated by FEM and MLPNN and compared for correlation in trend (Figure S4). Ranges of morphological parameters explored were determined by literature and simulation results and the accuracy of predictions was measured on a global (Figure S4) and parameter-by-parameter basis (Figures S4-5).

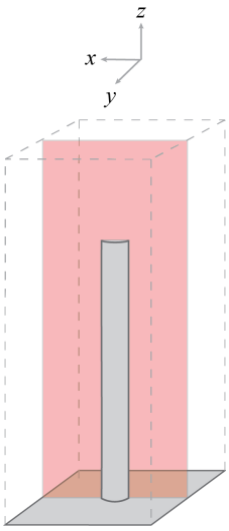
The procedure of comparing simulation duration

The comparison of simulation durations between the FEM and MLPNN approaches (Figure 1G) were conducted on an Intel® Core™ i7-7800 CPU @ 3.50 GHz with 32 GB of random access memory (RAM) and 6 central processing unit (CPU) cores with an NVIDIA GeForce GTX 1060 3GB graphics processing unit (GPU) and a 1.4 GHz Dual-Core i5 MacBook Air with 8 GB of RAM (MD760LL/B) for the FEM and MLPNN simulations, respectively (Table S4). Up to the order of 10^3 predictions could be taken using the FEM procedure before exhausting the RAM, and up to the order of 10^5 predictions could be made using the MLPNN before the same issue occurred.

Local and global electrode optimization

Global optimization (Figure 2A) was achieved by generating over 4,000,000 possible catalyst and morphology combinations via the neural-network approach and presenting the combinations predicted to have greater than 2 mA/cm^2 of i_{total} and 90% of *F.E.* Literature-based catalyst information was derived using equations S8 and S12 (Supplementary note 1) in combination with the WebPlotDigitizer tool provided by Ankit Rohatgi at *https://apps.automeris.io/wpd/* to extract data from published literature plots. The automeris.io application allowed the conversion of published graphical data into raw data. This data was then used to calculate NRR, HER, and *F.E.* data from published information. This data was sufficient to estimate the kinetic activity descriptors ($i_{1,0,NRR}$, $i_{2,0,NRR}$, $i_{1,0,HER}$, and $i_{2,0,HER}$) necessary for application in the neural network models. Morphologies to maximize *F.E.*, i_{total}, or both simultaneously were predicted for each literature-derived catalyst (Figure 2B, S2).[7-11]

Generation of reactant concentration profiles



Reactant profiles shown in Figure 3D and Figures S6-9 were generated using COMSOL Multiphysics 5.3. Upon solving for the current densities of the electrode materials, two-dimensional concentration gradients bisecting the wire of the cell were plotted for both proton and nitrogen (as shown right). For a given grouping of wire concentration profiles, an absolute scale was presented for the reactant of interest, this can give the appearance of no reactant concentration profile where diffusion in not highly obstructed (i.e. Figure S6 far left panels); for these cases, the change in concentration of species from the top of the cell to the bottom was minimal, so a mostly monochromatic concentration profile results from the simulation.

Supplementary Notes

Note 1: Summary of FEM equations

In our FEM model for electrocatalytic NRR, the kinetics of electrochemical nitrogen reduction consider the following steps (Figure 2B):

Electron transfer (ET)	$N_2(aq) + e^- \leftrightharpoons N_2^{\bullet -}(ad)$	$k_{f1}(\eta)$, $k_{b1}(\eta)$	(**S5**)
Proton transfer (PT)	$N_2^{\bullet -}(ad) + H^+(aq) \rightarrow HN_2(ad)$	k_{f2}	(**S6**)

Here, k_{f1} and k_{b1} are forward and backward rate constants – functions of overpotential – for the ET step, respectively. k_{f2} is the rate constant of PT step. States are given by *aq*, aqueous; *ad*, adsorbed to catalyst surface. At steady state, the concentration of $N_2^{\bullet -}(ad)$ remains constant. Using this information, the local current density of nitrogen fixation $i_{loc,\mathrm{NRR}}$ can be written as:

$$i_{loc,NRR} = -nF \frac{k_{f1,0}[N_2(aq)] exp\,(\frac{\alpha_c F\eta}{RT})}{\frac{k_{b1,0} exp(\frac{\alpha_c F\eta}{RT})}{k_{f2,0}[H^+(aq)]} + 1} \qquad \textbf{(S7)}$$

Where $k_{f1,0}$ and $k_{b1,0}$ are rate constants at 0 over potential. α_c is the cathodic charge exchange coefficient, F is Faraday's constant, T is temperature in kelvin, and R is the gas constant. Further modification can be made to the equation resulting in our generalized NRR equation:

$$i_{loc,NRR} = -\frac{i_{1,0,NRR}\frac{[N_2(aq)]}{c_{N_2(aq)}} exp\,(\frac{-\alpha_c F\eta}{RT})}{1 + \frac{i_{1,0,NRR} C_{H^+(aq)}\, exp(\frac{\alpha_c F\eta}{RT})}{i_{2,0,NRR}[H^+(aq)]}} \qquad \textbf{(S8)}$$

Where $i_{1,0,\mathrm{NRR}}$ and $i_{2,0,\mathrm{NRR}}$ are the exchange current densities when the rate determining step (RDS) is ET and PT respectively (Figure S1A). $c_{N_2(aq)}$ and $c_{H^+(aq)}$ are the bulk concentrations of N_2 and H^+ and $[N_2(aq)]$ and $[H^+(aq)]$ are the concentrations of species at the electrode surface.

The FEM equation can be discerned in a similar way, where we consider a Volmer-Heyrovský mechanism:[5,6]

Volmer: $H^+(aq) + e^- \leftrightharpoons H(ad)$ $k_{f3}(\eta), k_{b3}(\eta)$ (**S9**)

Heyrovksý: $H(ad) + H^+(aq) + e^- \rightarrow H_2(aq)$ (**S10**)

Here, k_{f3} and k_{b3} are forward and backward rate constants – functions of overpotential – for the ET step, respectively. k_{f4} is the rate constant of the Heyrovský step. H (*ad*) represents surface-bound hydrogen atom.

Applying steady-state conditions and potential dependency for the rate constants k_{f3}, k_{b3}, and k_{f4} we derive:

$$i_{loc,HER} = -nF\frac{k_{f3,0}[H^+(aq)]exp\left(\frac{-\alpha_c F\eta}{RT}\right)}{1+\frac{k_{b3,0}exp(\frac{2\alpha_c F\eta}{RT})}{k_{f4,0}[H^+(aq)]}} \quad \textbf{(S11)}$$

Here, $i_{loc,\mathrm{HER}}$ is the local hydrogen evolution current density. Similarly, this equation can be rewritten using the exchange current density notation relevant to our simulation:

$$i_{loc,HER} = -\frac{i_{1,0,HER}\frac{[H^+(aq)]}{c_{H^+(aq)}}exp\left(\frac{-\alpha_c F\eta}{RT}\right)}{1+\frac{i_{1,0,HER}C_{H^+(aq)}exp(\frac{2\alpha_c F\eta}{RT})}{i_{2,0,HER}[H^+(aq)]}} \quad \textbf{(S12)}$$

In this equation, $i_{1,0,\mathrm{HER}}$ and $i_{2,0,\mathrm{HER}}$ represent exchange current densities when the RDS is the Volmer and Heyrovský steps respectively (Figure S1B).

Note 2: Hardware Limitations

In Figure 1G, the plotted curves showing the time to prediction comparison between FEM and MLPNN methods transition from solid to dashed lines. The transition indicates the number of electrodes predicted before human intervention is needed to further predict new electrodes for their respective methods. Because of RAM depletion (computer memory becomes exhausted), the data must be exported and predictions manually restarted, meaning 10^3 and 10^5 electrodes are easily predictable for the FEM and MLPNN, respectively.

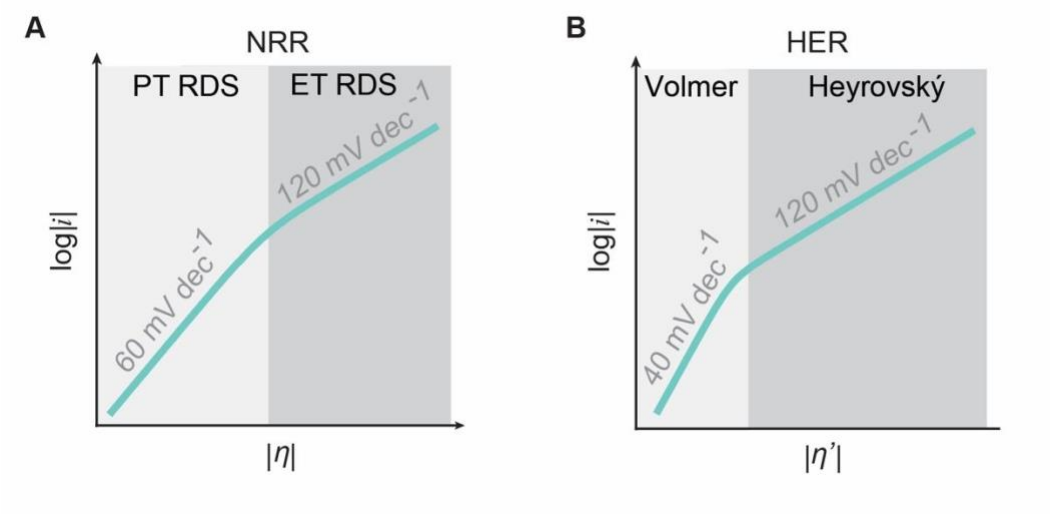


Figure S1. Representative log(i_{total}) versus η (overpotential) plots of our FEM simulations that show the mechanistic dependence on overpotential of the two competing processes. For the NRR, the proton transfer step is rate limiting at low overpotentials as the slope of 60 mV/decade implies an EC mechanism. At higher overpotentials, the electron transfer step becomes rate limiting as the slope of 120 mV/decade implies an E mechanism. For the HER, a similar trend is shown except at low overpotentials the 40 mV/decade slope implies an EE mechanism transitioning into an E mechanism at higher overpotentials.

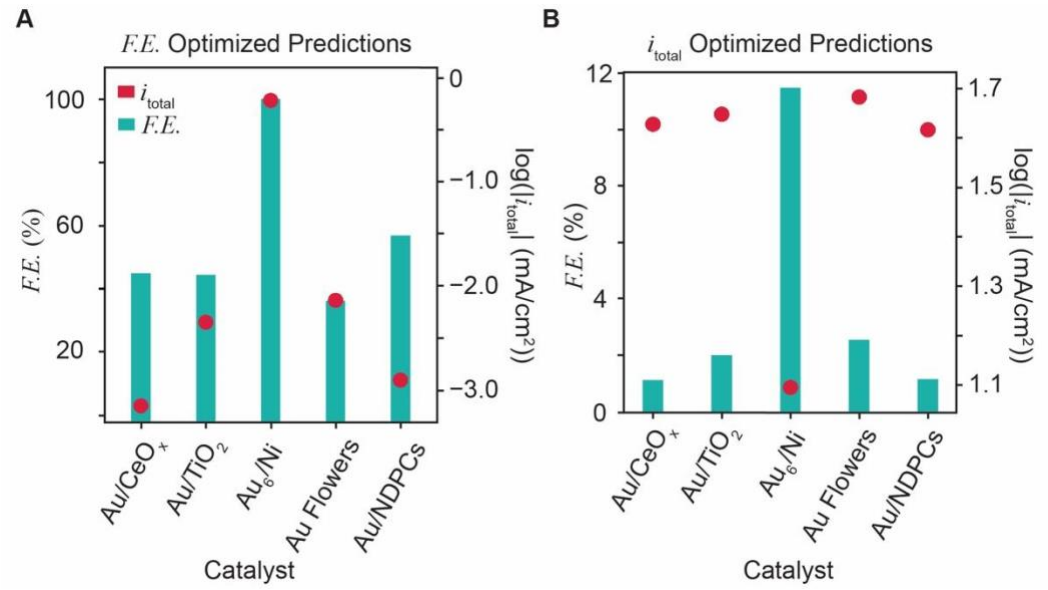


Figure S2. i_{total} and *F.E.* versus catalyst outputs for literature-based materials. Electrodes are maximized for either *F.E.* (**A**) or i_{total} (**B**).

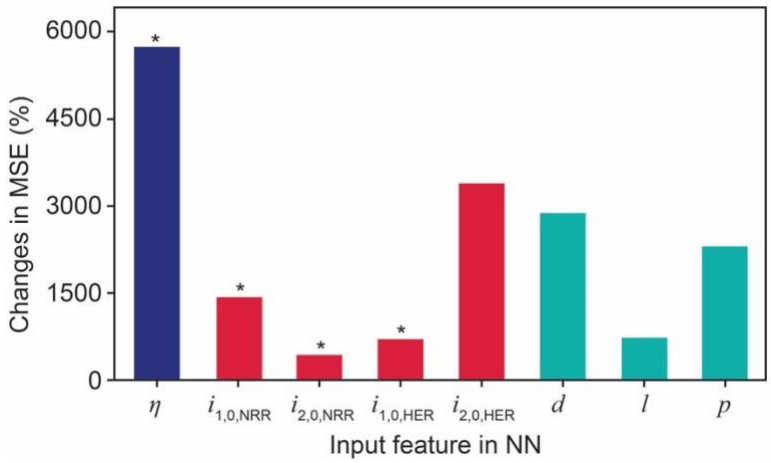


Figure S3. Change in MSE as a percentage versus input feature indicates relative feature importance values for the MLPNN tasked with predicting i_{total}. Bars marked with a "*" have been divided by a factor of 50. The procedure is the same as that presented for the feature importance analysis for *F.E.* in Figure 3A.

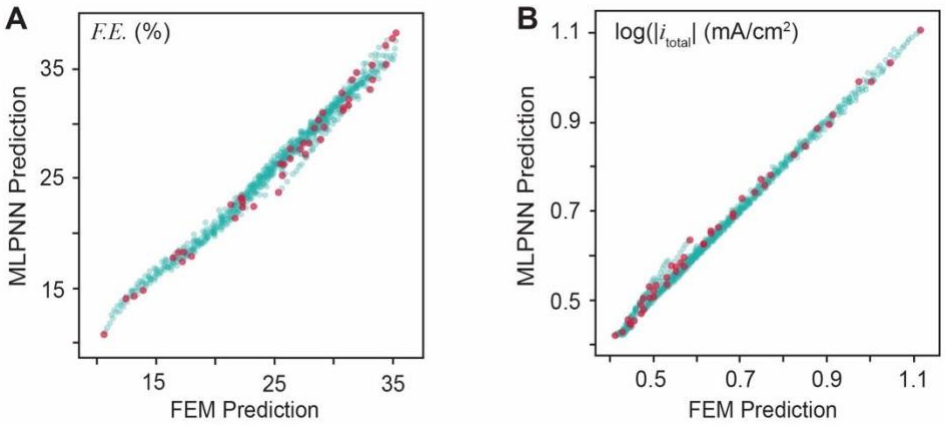


Figure S4. Correlation of FEM and MLPNN predictions for the expanded morphology set for the catalyst described by $i_{1,0,NRR} = 10^{-6}$, $i_{2,0,NRR} = 10^{-3}$, $i_{1,0,HER} = 10^{-4}$, and $i_{2,0,HER} = 10^{-6}$ mA/cm^2 (same catalyst discussed in Figure 3). Red dots show morphologies seen by the original training data and green dots indicate never-before-seen morphologies.

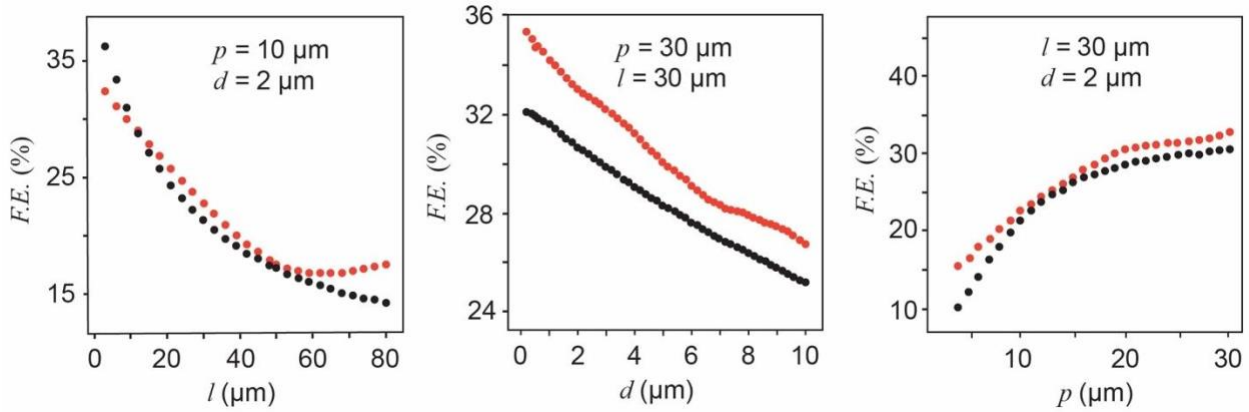


Figure S5. *F.E.* versus morphology-feature plots predicted by the MLPNN (red) and FEM (black) in a highly morphology-sensitive region.

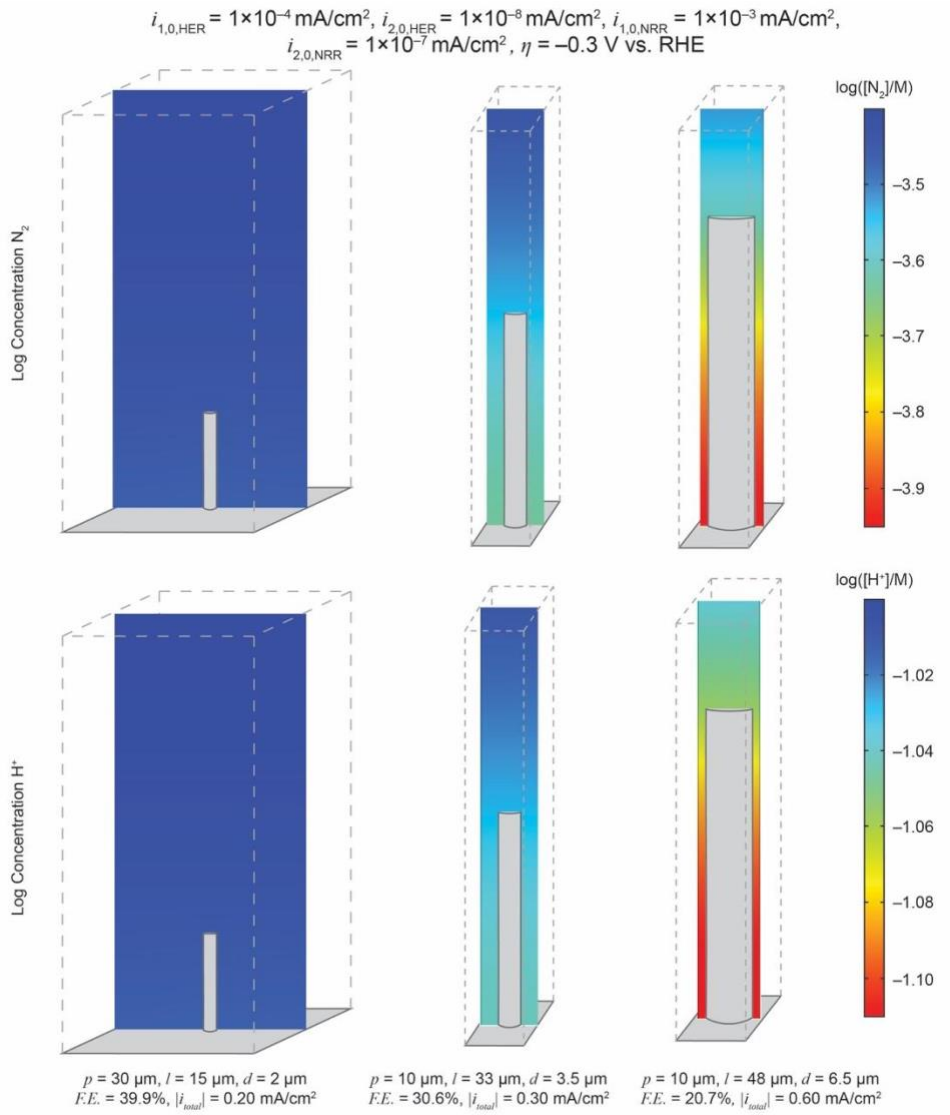


Figure S6. Reactant concentration profiles for a material with high morphology dependent Faradaic efficiency at an overpotential of –0.3 V. Nitrogen concentration profile above, proton concentration below.

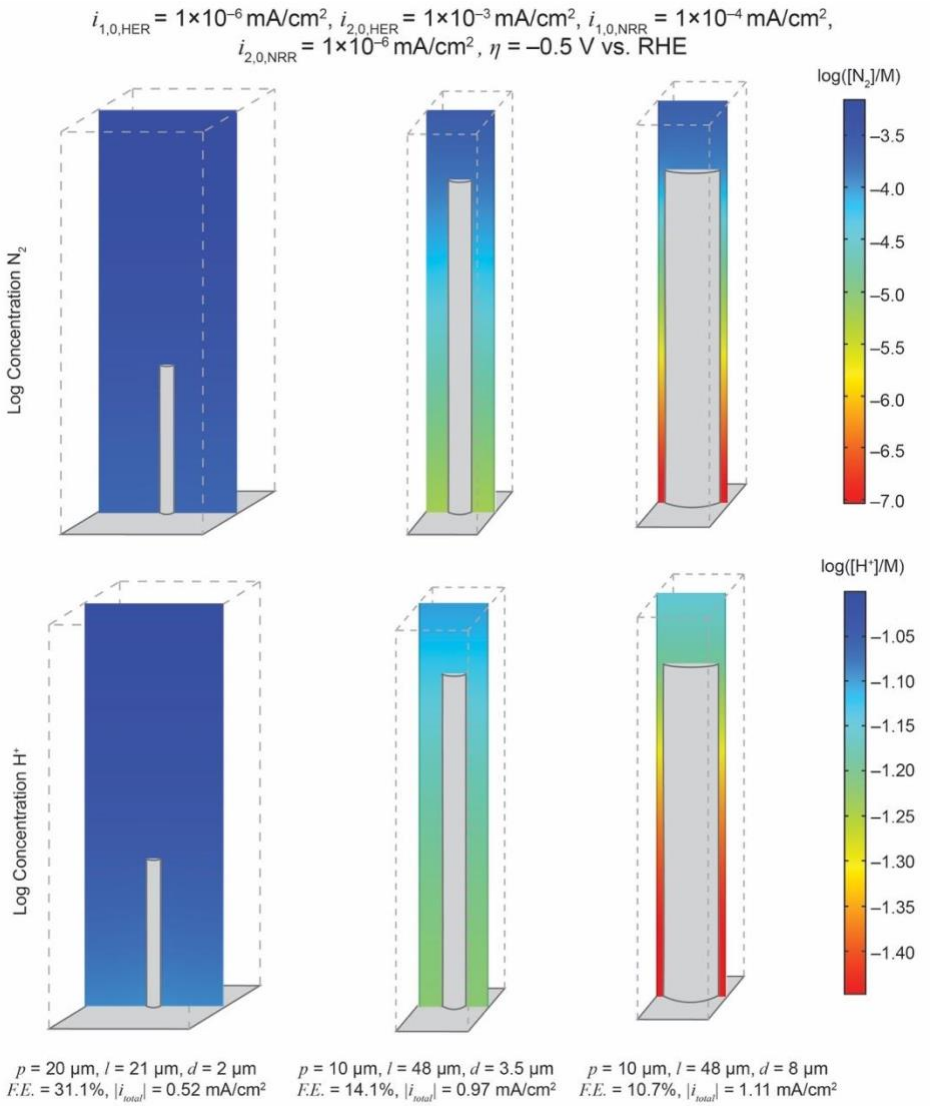


Figure S7. Reactant concentration profiles for a material with high morphology dependent Faradaic efficiency at an overpotential of –0.5 V. Nitrogen concentration profile above, proton concentration below.

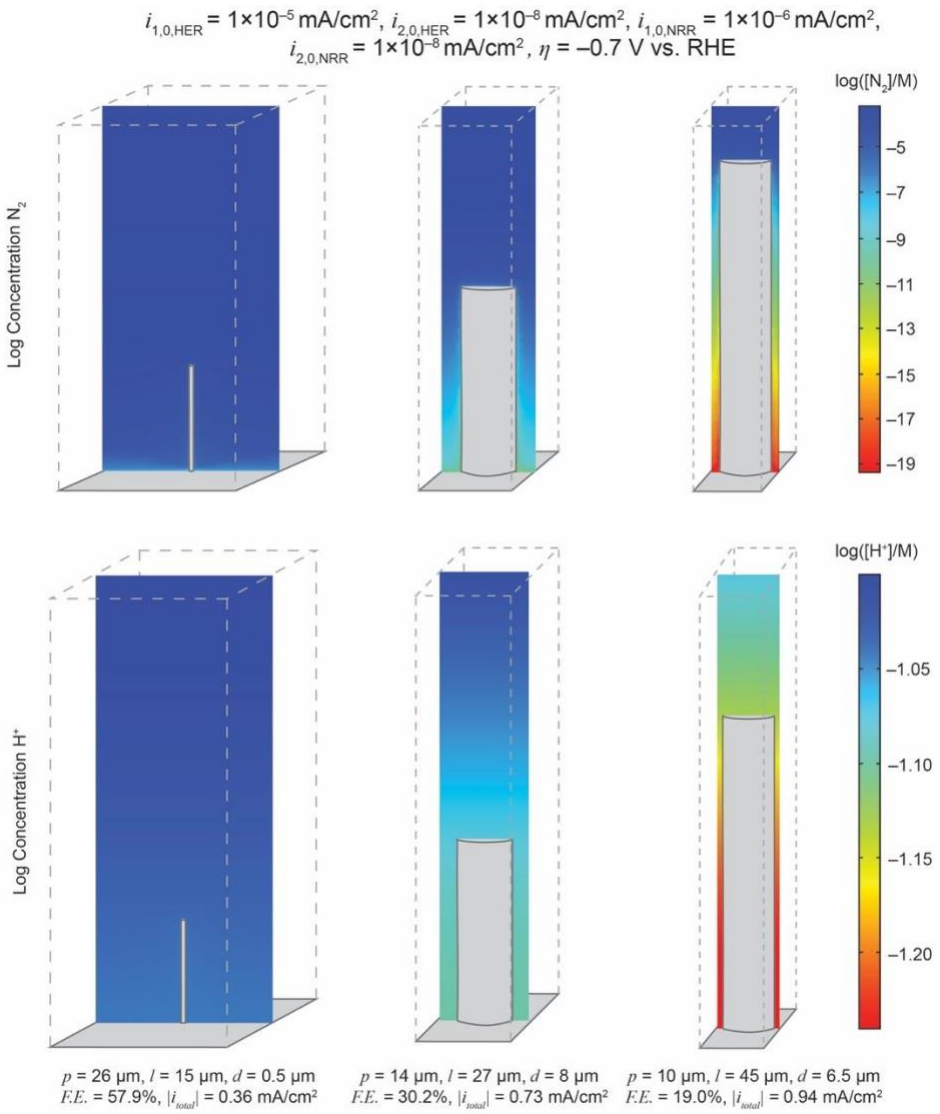


Figure S8. Reactant concentration profiles for a material with high morphology dependent Faradaic efficiency at an overpotential of –0.7 V. Nitrogen concentration profile above, proton concentration below.

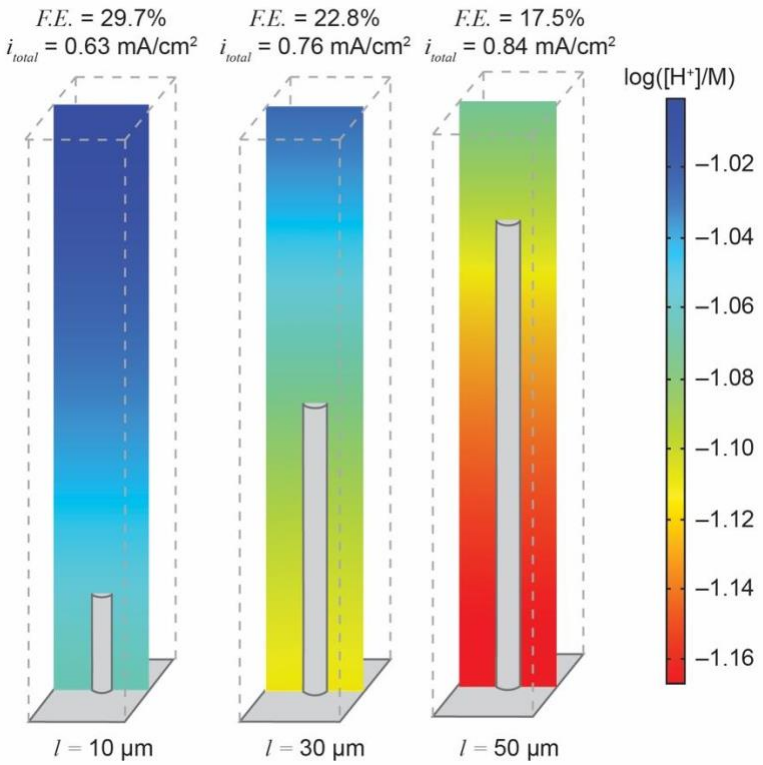


Figure S9. Proton concentration profiles for a catalytic material showing high Faradaic efficiency dependence on wire-array length. Nitrogen concentration gradient shown in Figure 3D. Kinetic descriptors of this system: $i_{1,0,NRR} = 10^{-6}$, $i_{2,0,NRR} = 10^{-3}$, $i_{1,0,HER} = 10^{-4}$, and $i_{2,0,HER} = 10^{-6}$ mA/cm^2 and η = –0.5 V vs. RHE.

Table S1. Invariant parameters of FEM simulation

Parameter	Value
Diffusion layer thickness	53.5 μm
Diffusion coefficient of N_2, D_{N_2}	2.0×10^{-9} m^2/s
Diffusion coefficient of H^+, D_{H^+}	1.9×10^{-9} m^2/s
Bulk concentration of N_2, c_{N_2}	7.14×10^{-4} mol/L
Bulk concentration of H^+, c_{H^+}	0.1 mol/L
Standard reduction potential of N_2/NH_4^+, $E^0_{N_2/NH_4^+}$	0.27 V vs. RHE
Standard reduction potential of H^+/H_2, $E^0_{H^+/H_2}$	0 V vs. RHE
Cathodic charge transfer coefficient, α_c	0.5

Parameters for the FEM model that were invariant in the simulations made relevant to this research.

Table S2. Ranges of parameters used in FEM simulations

Feature	Range
Periodicity (p)	10 ~ 30 μm
Length (l)	15 ~ 50 μm
Diameter (d)	0.5 ~ 8 μm
η (vs. RHE)	–0.7~–0.1 V
$i_{1,0,NRR}$	10^{-8} ~ 10^{-3} mA/cm^2
$i_{2,0,NRR}$	10^{-8} ~ 10^{-3} mA/cm^2
$i_{1,0,HER}$	10^{-8} ~ 10^{-3} mA/cm^2
$i_{2,0,HER}$	10^{-8} ~ 10^{-3} mA/cm^2

Ranges of parameters used in the establishment of the FEM-based database. Range values are inclusive of the stated boundary values. For convenience all of our machine-learning studies arbitrarily operate with a geometric resolution of 1 μm, which can be improved in future investigations

Table S3. Hyperparameters of trained MLPNNs

Feature	Default	Chosen
Hidden Layer Sizes	100	2000
Activation Function	relu	relu
Solver	adam	adam
Validation Fraction	0.1	0.3
Early Stopping	false	true

Hyperparameters of MLPNNs used for predictions. Those shown are either changed from default values or are fundamental to the machine learning model's architecture. The *F.E.* MLPNN was trained and optimized first (activation function, solver, hidden layer sizes, and epsilon focus for grid search). Once established, the parameters of the FE model were used as a basis for the i_{total} MLPNN. Upon initial utilization, it was found that these hyperparameters were satisfactory for prediction of i_{total} as shown in Figure 1F, and thus both models had the same hyperparameters.

Table S4. Computer hardware

Device	Specifications	Uses
PC (Laptop)	Processor: 1.4 GHz Dual-Core Intel Core i5 Memory: 8 GB	Predictions for electrodes using trained neural networks, development of figures and code development.
Workstation	Processor: Intel® Core™ i7-7800X CPU @ 3.50 GHz Memory: 32 GB GPU: NVIDIA GeForce GTX 1060 3GB	Used for FEM simulations.
Virtual Machine	Variable Specifications	A virtual machine run through Google Cloud Platform's Jupyter Notebooks was used for selecting and training machine learning algorithms.

Details of hardware used throughout the course of this research.

Table S5. Circled electrodes from figure 2A

entry	p (μm)	l (μm)	d (μm)	η (V)	$i_{1,0,NRR}$ (mA/cm^2)	$i_{2,0,NRR}$ (mA/cm^2)	$i_{1,0,HER}$ (mA/cm^2)	$i_{2,0,HER}$ (mA/cm^2)	*F.E.* (%)	i_{total} (mA/cm^2)
1	9	42	5	−0.5	10^{-4}	10^{-8}	10^{-6}	10^{-3}	90.3	−2.1
2	10	48	8	−0.7	10^{-3}	10^{-3}	10^{-8}	10^{-8}	91.0	−2.1
3	11	39	8	−0.7	10^{-3}	10^{-3}	10^{-8}	10^{-8}	93.0	−2.0
4	10	33	8	−0.7	10^{-3}	10^{-3}	10^{-8}	10^{-8}	94.3	−2.0

Morphological and kinetic information for circled catalysts in Figure 2A. All exchange current density and total current values have units of mA/cm^2 and potential is vs. RHE. The entry values correspond to the circled catalysts from left to right (lowest to highest *F.E.*). Note for convenience all of our search of optimized geometry arbitrarily operate with a geometric resolution of 1 μm, which can be improved in future investigations.

Table S6. Comparison of MLPNN selected electrodes and their literature counterparts

	Literature reported values		MLPNN optimized	
Catalyst	*F.E.* (%)	$\lvert i_{total} \rvert$ (mA/cm^2)	*F.E.* (%)	$\lvert i_{total} \rvert$ (mA/cm^2)
Au/CeO_x [7]	10.10	0.13	17.0	0.16
			η:–0.3 V, *p*:8 µm, *l*:80 µm, *d*:0.4 µm	
Au/TiO_2 [8]	8.11	0.56	15.1	1.35
			η:–0.3 V, *p*:30 µm, *l*:80 µm, *d*:10 µm	
Au_6/Ni [9]	67.8	0.21	99.0	0.37
			η:–0.3 V, *p*:23 µm, *l*:16 µm, *d*:8.5 µm	
Au Flowers [10]	6.05	1.2	11.7	1.53
			η:–0.3 V, *p*:23 µm, *l*:80 µm, *d*:2.5 µm	
Au/NDPCs [11]	12.3	0.34	22.7	0.42
			η:–0.3 V, *p*:30 µm, *l*:80 µm, *d*:10 µm	

Comparison of FE and i_{total} values for reported gold catalysts and those predicted by the MLPNNs with one prediction to maximize FE and one to maximize i_{total}. All literature values are FE maximized and reported at −0.2 V vs. RHE with the exception of Au_6/Ni which was reported at −0.14V vs. RHE. Under the FE and i_{total} values are summaries of relevant parameters: E, η (V) vs. RHE; *p*, periodicity (µm); *l*, length (µm); *d*, diameter (µm).

References

1. Li, Y.; Zheng, J.; Lyu, Y.; Qiao, M.; Veder, J. P.; Marco, R. D.; Bradley, J.; Wang, R.; Huang, A.; Jiang, S. P.; Wang, S. Electron localization of gold in control of nitrogen-to-ammonia fixation. *Angew. Chem. Int. Ed.*, **2019**, *58* (51), 18604–18609.
2. Ferrell, R. T.; Himmelblau, D. M. Diffusion coefficients of nitrogen and oxygen in water. *J. Chem. Eng. Data* **1967,** *12* (1), 111–115.
3. Chen, H.; Voth, G. A.; Agmon, N. Kinetics of Proton Migration in Liquid Water. *J. Phys. Chem. B* **2010,** *114* (1), 333–339.
4. Kolev, N. *Multiphase Flow Dynamics 4: Turbulence, Gas Adsorption and Release, Diesel Fuel Properties*. Springer, Berlin, Heidelberg, Germany: 2012.
5. Nørskov, J. K.; Bligaard, T.; Logadottir, A.; Kitchin, J. R.; Chen, J. G.; Pandelov, S.; Stimming, U. Trends in the Exchange Current for Hydrogen Evolution. *J. Electrochem. Soc.* **2005,** *152* (3), J23–J26.
6. Seh, Z. W.; Kibsgaard, J.; Dickens, C. F.; Chorkendorff, I.; Nørskov, J. K.; Jaramillo, T. F. Combining theory and experiment in electrocatalysis: Insights into materials design. *Science* **2017,** *355* (6321), eaad4998.
7. Li, S.-J.; Bao, D.; Shi, M.-M.; Wulan, B.-R.; Yan, J.-M.; Jiang, Q. Amorphizing of Au Nanoparticles by CeO_x–RGO Hybrid Support towards Highly Efficient Electrocatalyst for N_2 Reduction under Ambient Conditions. *Adv. Mater.* **2017,** *29* (33), 1700001.
8. Shi, M.-M.; Bao, D.; Wulan, B.-R.; Li, Y.-H.; Zhang, Y.-F.; Yan, J.-M.; Jiang, Q. Au Sub-Nanoclusters on TiO_2 toward Highly Efficient and Selective Electrocatalyst for N_2 Conversion to NH_3 at Ambient Conditions. *Adv. Mater.* **2017,** *29* (17), 1606550.

9. Xue, Z.-H.; Zhang, S.-N.; Lin, Y.-X.; Su, H.; Zhai, G.-Y.; Han, J.-T.; Yu, Q.-Y.; Li, X.-H.; Antonietti, M.; Chen, J.-S. Electrochemical Reduction of N_2 into NH_3 by Donor–Acceptor Couples of Ni and Au Nanoparticles with a 67.8% Faradaic Efficiency. *J. Am. Chem. Soc.* **2019,** *141* (38), 14976−14980.

10. Wang, Z.; Li, Y.; Yu, H.; Xu, Y.; Xue, H.; Li, X.; Wang, H.; Wang, L. Ambient Electrochemical Synthesis of Ammonia from Nitrogen and Water Catalyzed by Flower-Like Gold Microstructures. *ChemSusChem* **2018,** *11* (19), 3480−3485.

11. Qin, Q.; Heil, T.; Antonietti, M.; Oschatz, M. Single-Site Gold Catalysts on Hierarchical N-Doped Porous Noble Carbon for Enhanced Electrochemical Reduction of Nitrogen. *Small Methods* **2018,** *2* (12), 1800202.

Chapter 3. The inverse design of oxygen and hydrogen peroxide concentration gradients in wire electrodes

This chapter is a version of Chen, Y.†; Wang, J.†; Hoar, B. B. †; Lu, S.; Liu, C. "Mach-learning-based inverses design of electrochemically controlled microscopic gradients of O_2 and H_2O_2." *PNAS* **2022**, *119(32)*, e2206321119 (†, equal contribution)

Abstract

A fundamental understanding of extracellular microenvironments of O_2 and reactive oxygen species (ROS) such as H_2O_2, ubiquitous in microbiology, demands high-throughput methods of mimicking, controlling, and perturbing gradients of O_2 and H_2O_2 at microscopic scale with high spatiotemporal precision. However, there is a paucity for a high-throughput strategy of microenvironment design, and it remains challenging to achieve O_2 and H_2O_2 heterogeneities with microbiologically desirable spatiotemporal resolutions. Here we report the inverse design, based on machine learning (ML), of electrochemically generated microscopic O_2 and H_2O_2 profiles relevant for microbiology. Microwire arrays with suitably designed electrochemical catalysts enable independent control of O_2 and H_2O_2 profiles with spatial resolution of ~101 μm and temporal resolution of ~100 sec. Neural networks aided by data augmentation inversely design the experimental conditions needed for targeted O_2 and H_2O_2 microenvironments while being orders-of-magnitude faster (more than 250 times faster than experimental explorations). Interfacing ML-based inverse design with electrochemically controlled concentration heterogeneity creates a viable fast-response platform towards.

Introduction

Ubiquitous spatiotemporal heterogeneity of natural environments fosters the diverse and fascinating biology that our world embraces and motivates researchers to mimic natural environments with high spatiotemporal resolution.[1-5] Given their close relevance in biochemical metabolisms, dioxygen (O_2) and hydrogen peroxide (H_2O_2) as a surrogate of reactive oxygen species (ROS) are two ubiquitous biologically relevant species in extracellular medium.[1, 6] Their extracellular spatial and temporal distributions, particularly at microscopic scale ranging from 1 μm to 100 μm,[7-11] are critical for signal transduction, protein expression, biochemical redox balance, and regulation for cellular metabolism with extensive ecological, environmental, and biomedical implications (Fig. 1A).[1, 3, 8-13] A programmable creation of the spatiotemporal concentration profiles of $O_2$and H2O2 offers the freedom to mimic, control and perturb the microenvironments of $O_2$and H_2O_2 and hence advance our understanding in microbiology.

Despite the recent progresses,[14-18] there remain major technical challenges, particularly in the achievable spatiotemporal resolution and high-throughput design of concentration profiles to suit a plethora of scenarios in microbiology. Approaches based on microfluidics and hydrogels have been able to achieve concentration gradients of O_2 and H_2O_2 through the provision of either O_2/H2O2 source,[14, 19-21] O_2/H_2O_2 scavenging agents,[15, 22, 23] or a combination of both[24] across liquid-impermeable barriers such as agar layers or polymeric thin films.[25, 26] Yet such approaches dependent on passive mass transport and diffusion across more than 102 μm are inherently incapable of achieving spatial features of less than 100 μm and temporal resolution smaller than ~101 seconds, the prerequisites to investigate microbiology at cluster or single-cell levels.[10-12] Moreover, the large variations of extracellular O_2 and H_2O_2 gradients in different microbial systems demand an inverse-design strategy, which with minimal expenditure quickly programs a desired

concentration profile catering to a specific biological scenario.[2-5] The current lack of inverse-design protocol impedes the adoption of controllable extracellular heterogeneity to mimic and investigate microbial systems that are of environmental, biomedical, and sustainability-related significances.

We envision that the integration of electrochemically generated concentration gradients with inverse design based on machine learning (ML) will address the aforementioned challenges (Fig. 1B). Electrochemistry offers a venue of transducing electric signals into microscopic concentration profiles within 100 ~ 102 μm away from electrodes' surface, following the specific electrode reaction kinetics and the mass-transport governing equations in the liquid phase.[27] Proper designs of electrodes' microscopic spatial arrangement and electrochemical kinetics lead to concentration gradients that are spatiotemporally programmable by time-dependent electric signals of varying voltages.[28] Such benefits of electrochemically generated concentration gradients lead us to employ electrochemistry as a tool to spatiotemporally control the concentration profiles in the extracellular medium. In one example, we found that wire arrays electrochemically active towards O_2 reduction create anoxic microenvironment about 20 μm away from the aerobic external bulk environments, modulate the size and extent of O_2 depletion in the anoxic microenvironment by wire array's morphology and applied electrochemical potential (E_{appl}), and hence enable O_2-sensitive rhizobial N_2 fixation in ambient air powered by renewable electricity.[29] Moreover, while not reported before as far as we know, electrochemically generated concentration heterogeneity is commensurate with ML-based inverse design,[30, 31] thanks to the mathematically well-defined electrochemical processes that can be numerically simulated.[32, 33] We recently reported neural networks, trained by numerically simulated data, that explore the influence of electrode geometry on electrochemical N_2 fixation and achieve optimized morphologies of wire array electrodes

untenable without such ML-based strategy.[34] An inverse design for the electrochemically generated gradients will quickly program desirable microenvironments of O_2 and ROS with high spatiotemporal resolutions, thanks to the well reported electrochemical transformation related to O_2 and H_2O_2 with high electrochemical selectivity.[35, 36]

In this work, we report an inverse design based on neural networks for independent electrochemical creation of O_2 and ROS microscopic gradients that are relevant and mimic their extracellular heterogeneities in microbial systems. We hypothesize that careful design of electrocatalysis of O_2 reduction reaction (ORR) can either facilitate four-electron ORR on Pt electrocatalyst for a controllable O_2 spatiotemporal profile, or promote two-electron ORR on Au electrocatalyst for a programmable generation of H_2O_2 gradient without significantly perturbing the O2 one, thanks to their concentration differences in biological mediums (10−1~101 μM for H_2O_2 and 101~102 μM for O_2).2, 7-11 Electrochemically active microwire array electrodes as exemplary model systems (Fig. 1C) are experimentally shown to achieve tunable heterogeneities of O_2 and H_2O_2 independently, with spatial resolution of ~101 μm and temporal resolution of ~100 sec, and are suitable as a platform of independently perturbing biologically relevant O_2 and H_2O_2 profiles in microbial systems. We further established and experimentally validated two neural networks that inversely design the wire array electrodes' morphologies towards targeted microenvironments of O_2 and H_2O_2, respectively, which is at least one order of magnitude faster than trial-and-error numerical simulation and two orders of magnitudes faster than experimental explorations. The demonstrated inverse design of electrochemically generated controlled gradients not only demonstrates a full electrochemical control of concentration profiles in an electrode's proximity but also establishes an approach of spatiotemporally mimicking and perturbing extracellular space guided by artificial intelligence.

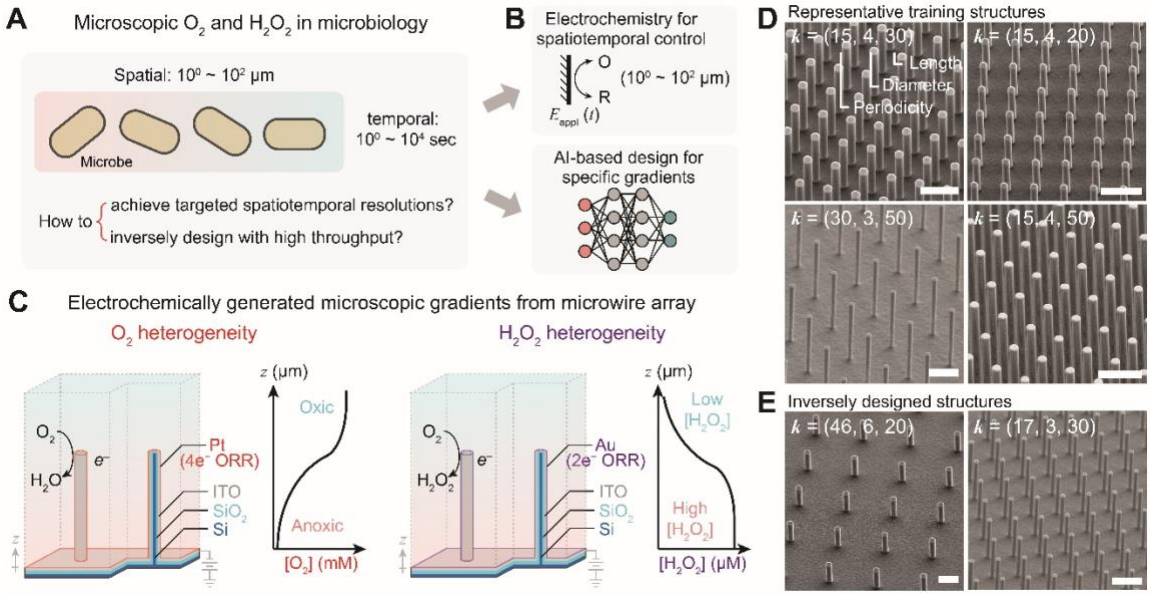


Figure 1. AI-based inverse design of electrochemically generated O_2 and H_2O_2 heterogeneities. **A** The ubiquitous spatiotemporal heterogeneities of O_2 and H_2O_2 in microbiology and the challenges posed in this research topic. **B** The combination of electrochemistry and machine-learning-based inverse design offers a viable approach of mimicking and control the heterogeneities of O_2 and H_2O_2 in microbiology. O, oxidant; R, reductant; E_{appl} (t), the time-dependent electrochemical voltages applied on electrodes. **C** The design of the electrochemically active microwire array electrodes for the generation of O_2 and H_2O_2 gradients. $4e^-$ ORR & $2e^-$ ORR, four-electron and two-electron oxygen reduction reaction into H_2O and H_2O_2, respectively; ITO, indium tin oxide. **D** and **E**, 45-degree tilting images of scanning electron microscopy for the representative microwire arrays used for the training of machine-learning model **D** and the ones inversely designed for targeted O_2 and H_2O_2 gradients (*E*). $\boldsymbol{k}$ = (P, D, L), the morphological vector that includes the periodicity (P), diameter (D) and length (L) of the synthesized wire arrays in the unit of μm. Scale bars are 20 μm.

Results and Discussion

Wire array electrodes for electrochemical generation of O_2 and H_2O_2 gradients

We applied microwire electrode array loaded with selective ORR electrocatalysts to establish customizable O_2 or H_2O_2 gradients (Fig. 1C). Si-based microwire arrays in a square lattice were constructed through photolithography and reactive ion-etching in a five-step fabrication process (see "Materials and Methods" section). After thermal annealing to generate an electrically insulating oxide layer, indium-doped tin oxide (ITO) of 500 nm was deposited via sputtering near conformally onto the wire array, followed by the deposition of about 7-nm Pt and Au for the generation of O_2 and H_2O_2 gradients via selective ORR, respectively. Here the deposition of electrically conducted ITO layer ensures a uniform distribution of applied electrochemical potential (E_{appl}). We employed Pt as the selectively electrocatalysts of four-electron ORR[35, 37, 38] in the generation of O_2 gradients, while Au for two-electron ORR[35, 39, 40] in the generation of H2O2 gradients. The morphologies (Fig. 1D, 1E, and S1) and compositions (Fig. S2 and Fig. S3) of the established wire array were characterized and confirmed by scanning electron microscopy (SEM) equipped with energy dispersive X-ray spectroscopy (EDS), with a vector k = (P, D, L) presenting the wire arrays' periodicity (P), diameter (D), and length (L), respectively, in the unit of μm.

The prepared wire array electrodes coated with Pt and Au exhibit desirable electrochemical properties for creating O_2 and H_2O_2 heterogeneities, respectively. In phosphate buffer saline (PBS) solution, linear scan voltammograms (20 mV/s) of the deposited Pt electrocatalysts on the wire array (Fig. S4) exhibited an onset potential of ORR at around 0.8 V vs. reversible hydrogen electrode (RHE). Linear scan voltammograms of the deposited Au electrocatalysts showed a similar onset potential of ORR at around 0.6 V vs. RHE (Fig. S5). Experiments of rotating ring-disk electrode for the Au electrocatalysts (Fig. S6) displayed a selectivity of H_2O_2 generation from O_2 reduction up to 50% at 0.5 V vs. RHE. Thanks to the reaction-diffusion model in the electrolyte and the electrochemical boundary conditions imposed by the microwire morphology,[29, 41, 42] the Pt

and Au-loaded wire array electrode transduces electric voltages E_{appl} into concentration gradients of O_2 and H_2O_2, respectively, at microscopic length scales.

Electrochemical generation and control of O_2 concentration profiles

The Pt-deposited microwire array electrode is capable of spatiotemporally controlling the electrochemically generated O_2 gradient. Thanks to its triplet-triplet quenching with $3O_2$,[43] the phosphorescence emission of tris(1,10-phenanthroline)ruthenium(II), $Ru(phen)_3^{2+}$, from the inter-system crossing after optical excitation between 350~500 nm,[44] was utilized to spatiotemporally probe the local concentration of O_2 ($[O_2]$) (Fig. 2A). The constructed microwire array electrode was housed in a home-made fluidic device (Fig. S7) under a confocal microscope with a 470-nm optical excitation (see "Materials and Methods" section). Under a constant flow of aerated PBS solution with 0.15 mM $Ru(phen)_3^{2+}$ ($[O_2]$ = 0.246 mM saturated with ambient air[45]), the phosphorescence emission intensity (Ip, 610~650 nm), inversely proportional to the value of $[O_2]$, was collected and a calibration curve was established for the quantification of local $[O_2]$ values (see "Materials and Methods" section, Fig. S8). When k = (15, 4, 50) for the wire array (Fig. 2B and 2C), the three-dimensional Ip mapping were recorded in a time sequence when the Pt-coated wire array was initially under an open-circuit condition (t < 15 sec), subject to an electrochemical potential (E_{appl} = 0.5 V vs. RHE) from t = 15 and 45 sec, and reverted back to the open-circuit condition when t > 45 sec (see "Materials and Methods" section). The side-views of the three-dimensional Ip mapping were displayed when E_{appl} was initially absent (t = 0 sec), E_{appl} = 0.5 V vs. RHE (t = 16 sec), and E_{appl} was absent again at t = 48 sec ("t = 0", "t = 16 sec", and "t = 48 sec" in Fig. 2B, respectively). The intensity of Ip was noticeably stronger within the wire array when E_{appl} = 0.5 V vs. RHE in comparison to the Ip values under the open-circuit conditions before and after the presence of E_{appl}. As Ip is inversely proportional to the local values of $[O_2]$, this

observation qualitatively suggested a decrease of $[O_2]$ hence a O_2 gradient covering the wire array region with microscopic resolution under a reductive electrochemical potential.

The averaged $[O_2]$ values ($[O_2]_{avg}$) at different distances above the bottom of the wire array z = 5, 50, 100 μm were quantified and displayed as a function of time t in Fig. 2C. While negligible change of $[O_2]_{avg}$ was recorded at z = 100 μm (black in Fig. 2C), which was quite far away from the wire array, at z = 5 and 50 μm (blue and red in Fig. 2C, respectively) significant changes of $[O_2]_{avg}$ up to complete anoxic condition were observed, concurrent with the temporal presence of E_{appl}. Such data suggest that the established O_2 gradient can be temporally switched by electrochemical potentials faster than the temporal resolution of the confocal microscopy under the tested conditions (~2.7 sec). The electrochemically established O_2 gradient for wire array k = (30, 3, 50) were similarly quantified at E_{appl} = 0.4, 0.5, and 0.6 V vs. RHE as a function of the distance above the bottom of wire array (z) (black, red, and blue in Fig. 2D, respectively). The steepness of the generated O_2 gradient increased at lower E_{appl} values under which the electrochemical activities of O_2 consumption on Pt was more pronounced thanks to the increased magnitude of reductive overpotential (Fig. S4). Such an E_{appl} -dependent O_2 gradient showed that electrochemical input was capable of spatially controlling and yielding a desirable O_2 gradient for potential biological applications given biologically relevant $[O_2]$ values and the spatial resolutions detected here.[2-4, 46, 47] The electrochemically driven O_2 gradients were also quantified at E_{appl} = 0.5 V vs. RHE for k = (15, 4, 50), (30, 3, 50), and (30, 3, 30) (black, red, and blue in Fig. 2E, respectively). Noticeably different O_2 gradients were observed, suggesting the capability of wire array morphology to yield a specific O_2 gradient. Particularly, a strictly O_2-free local environment in aerated medium was established for k = (15, 4, 50). Such a customizable O_2 gradient will be of

interests for the study of communal interactions for the study of communal interactions among microorganisms of varying degrees of O_2 demands that are prevalent in nature.2

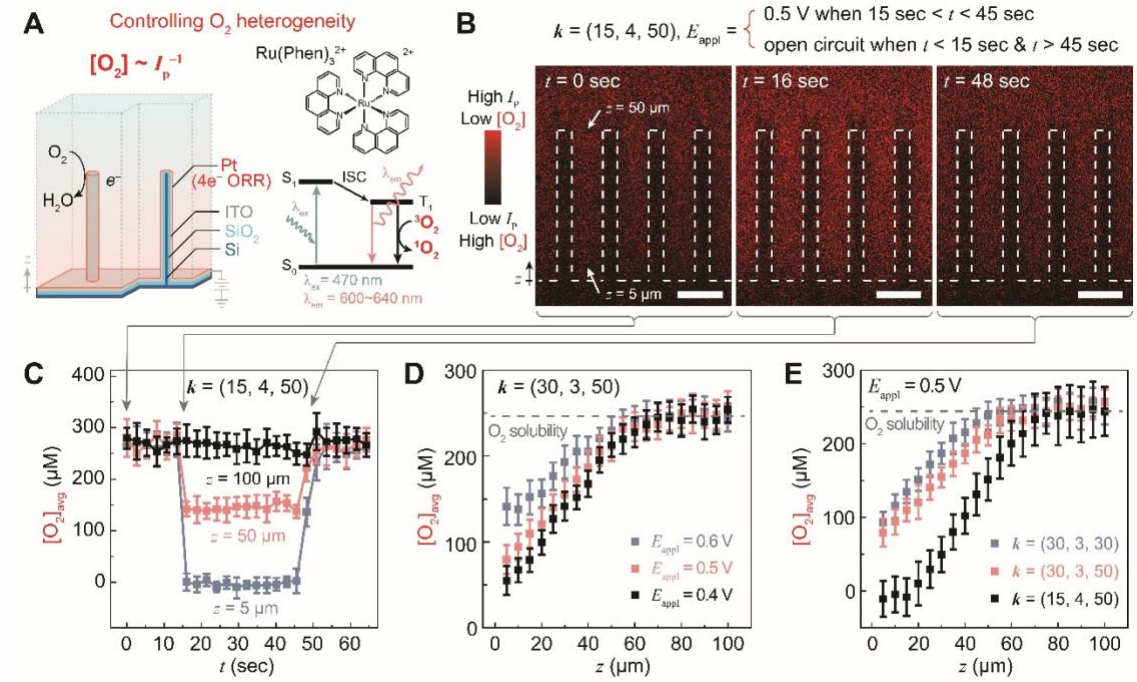


Figure 2. Spatiotemporal control of O_2 gradient on Pt-loaded microwire array. **A** Design of Pt-loaded microwire array and the fundamentals of spatiotemporal mapping local O_2 concentrations ($[O_2]$) based on the intensity of phosphorescence emission (I_p) from tris(1,10-phenanthroline)ruthenium(II) ($Ru(phen)_3^{2+}$). $^3O_2$ and $^1O_2$, the triplet and singlet dioxygen molecule, respectively; S_0 and S_1, the ground state and the first excited singlet state, respectively; T_1, the first excited triplet state; ISC, intersystem crossing; λ_{ex} and λ_{em}, the wavelengths of optical excitation and emission, respectively. **B** and **C** Cross-sectional I_p profiles on wire array $\boldsymbol{k}$ = (15, 4, 50) at t = 0, 16, and 48 secs **B** and the subsequent temporal evolution of averaged local O_2 concentrations ($[O_2]_{avg}$) at different distances z from the base of wire array **C**. The values of E_{appl} are reported against reversible hydrogen electrode (RHE). The microwires are depicted in dashed lines in **B**. Scale bars are 15 μm. **D** Plots of $[O_2]_{avg}$ versus z under different values of E_{appl} for wire array $\boldsymbol{k}$ = (30, 3, 50). (*E*) Plots of $[O_2]_{avg}$ versus z in wire arrays of different $\boldsymbol{k}$ when E_{appl} = 0.5 V. Error bars represent standard deviations across multiple separate measurements in the device ($n \geq 3$).

Electrochemical generation and control of H_2O_2 concentration profiles

We can similarly establish the gradients of H_2O_2, a potent reactive oxygen species (ROS) relevant to biology,[48] with the use of electrochemically active wire array electrodes and H_2O_2 -yielding Au ORR electrocatalysts. The local concentration of generated H_2O_2 ($[H_2O_2]$) was quantified based on the fluorogenic rection that converts non-fluorescent 10-acetyl-3,7-dihydroxyphenoxazine (Amplex Red) to fluorophore resorufin (λ_{ex} = 550 nm; λ_{em} = 590 ~ 650 nm) catalyzed by horseradish peroxidase (HRP)[49]. Under the similar setup mentioned above (Fig. S7), the emission intensities of resorufin (If), hence the local $[H_2O_2]$ value, were determined for Au-coated wire array electrode under the confocal microscopy, assisted by the corresponding calibration curves (Fig. S9 to Fig. S12, see "Materials and Methods" section). When k = (15, 4, 50) for the wire array (Fig. 3B and 3C), the three-dimensional If mapping was similarly recorded in a time sequence when the Au-coated wire array was initially under an open-circuit condition (t < 20 sec), subject to an electrochemical potential (E_{appl} = 0.5 V vs. RHE) from t = 20 and 50 sec, and reverted back to the open-circuit condition when t > 50 sec (see "Materials and Methods" section). The side-views of the three-dimensional If mapping was also displayed when E_{appl} was initially absent (t = 0 sec), E_{appl} = 0.5 V vs. RHE (t = 22 sec), and E_{appl} was absent again at t = 52 s ("t = 0 sec", "t = 22 sec", and "t = 52 sec" in Fig. 3B, respectively). While the absence of E_{appl} correlates with the absence of fluorescent emission from resorufin ("t = 0 sec" and "t = 52 sec" in Fig. 3B), the presence of E_{appl} = 0.5 V vs. RHE ("t = 22 sec" in Fig. 3B) yielded significant fluorescent emission near the wire array that suggested electrochemical generation of H_2O_2. Meanwhile, concurring monitoring of $[O_2]$ suggests that the local $[O_2]$ are not significantly perturbed (Fig. S13A) due to the relatively lower current density of ORR on the Au-based electrocatalysts under similar E_{appl} values (Fig. S4 and S5). This suggests that the electrochemically controlled H_2O_2 gradient is nearly

independent to aeration of the liquid medium, thanks to the catalytically selective generation of H_2O_2 and the low value of observed [H_2O_2] (at most up to 30 μM) relevant for biological studies[8, 48, 50] in comparison to the air-saturated O_2 solubility in water (246 μM).[45]

The averaged change of [H_2O_2] values ([H_2O_2]$_{avg}$) at different distances above the bottom of the wire array z = 5, 50, 90 μm were quantified and displayed as a function of time t in Fig. 3C. At all z values, the time-dependent generation of [H_2O_2]$_{avg}$ measured from If were well correlated with the presence of E_{appl}. A more gradual yet still relatively fast transition of the measured If (~10 sec) hence [H_2O_2]$_{avg}$ were observed, which was presumably due to limited temporal resolution of the fluorogenic reaction that was needed to track local [H_2O_2].[51] Nonetheless, such data indicated the capability of temporally controlling the formation of H_2O_2 electrochemically, which can be handy as a perturbation to study the microbial response towards H_2O_2 -based ROS.[52] We also determined the electrochemically induced H2O2 gradient for wire array k = (15, 4, 50) at E_{appl} = 0.45, 0.5, and 0.55 V vs. RHE (black, red, and blue in Fig. 3D, respectively). Significantly, different local accumulation of H_2O_2 up to 30 μM for different E_{appl} values were observed despite the 50 mV change of E_{appl}. Such observation suggests that the generated H_2O_2 gradient is highly sensitive and subsequently tunable by electrochemical driving forces. In addition, the morphology of wire array electrodes impacts the generated H_2O_2 gradient. The characterized H_2O_2 gradients for k = (15, 4, 50), (30, 3, 50), and (15, 4, 20) (black, red, and blue in Fig. 3E, respectively) were noticeably different at the same E_{appl} = 0.5 V vs. RHE. The achievable range of H_2O_2 gradients at the microscopic level is commensurate with biologically observed ROS microenvironments,8 heralding the utility of the electrochemically generated H_2O_2 gradients in microbial studies.

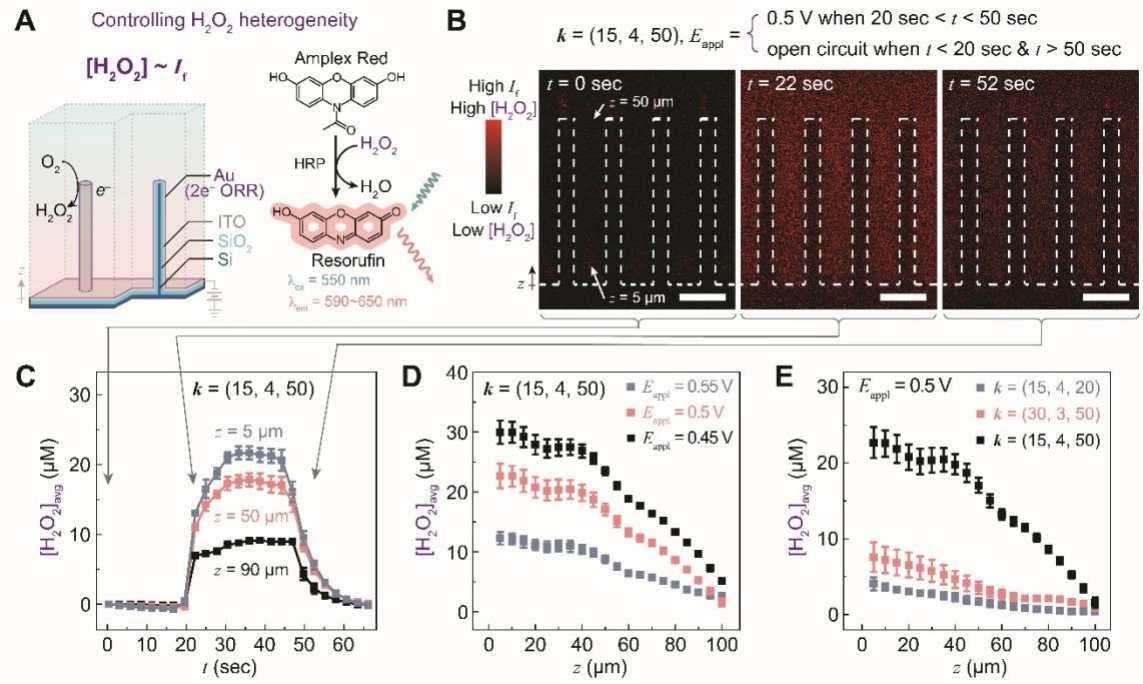


Figure 3. Spatiotemporal control of H_2O_2 gradient on Au-loaded microwire array. **A** Design of Au-loaded microwire array and the fundamentals of spatiotemporal mapping local H_2O_2 concentrations ($[H_2O_2]$) based on the intensity of fluorescence emission (I_f) in the fluorogenic reaction from Amplex Red to resorufin catalyzed by horseradish peroxidase (HRP). **B** and **C** Cross-sectional I_f profiles on wire array $\boldsymbol{k}$ = (15, 4, 50) at t = 0, 22, and 52 secs **B** and the subsequent temporal evolution of the averaged local H_2O_2 concentrations ($[H_2O_2]_{avg}$) at different distances z from the base of wire array **C**. The values of E_{appl} are reported against RHE. The microwires are depicted in dashed lines in **B**. Scale bars are 15 μm. $[H_2O_2]_{avg}$ = 0 when t = 0. **D** Plots of averaged local H_2O_2 concentrations ($[H_2O_2]_{avg}$) versus z under different values of E_{appl} for wire array $\boldsymbol{k}$ = (15, 4, 50). (*E*) Plots of $[H_2O_2]_{avg}$ versus z in wire arrays of different $\boldsymbol{k}$ when E_{appl} = 0.5 V. Error bars represent standard deviations across multiple separate measurements in the device ($n \geq 3$).

Establishing neural networks for an inverse design strategy

We seek to establish computational models that can inversely predict the values of E_{appl} and k = (P, D, L) of the Pt- and Au-loaded wire array electrodes for targeted corresponding O_2 and H_2O_2

gradients ([O_2](z) and [H_2O_2](z)), respectively. Such an inverse design strategy of O_2 and H_2O_2 microenvironments is proposed to be much more time-efficient in comparison with the classical trial-and-error approach (Fig. 4A), and will find plentiful applications given the high variabilities of biological applications in both spatial and temporal domains.[2, 3, 8] Critical inside such computational models are neural networks, trained with sufficient amounts of data, which correlate {E_{appl}, k = (P, D, L)} with the [O_2](z) and [H_2O_2](z) distributions. In such a regard, we employed numerical simulations based on finite-element methods (FEM)[32] to augment the available data (Fig. 4B). FEM-based electrochemical simulations have been widely used in the understanding and design of electrochemical applications with satisfactory accuracies.[32,33, 53-55] We established electrochemical microkinetic models that include the mass transport of redox species and the electrochemical reduction of O_2 for the Pt and Au electrocatalysts (see "Materials and Methods" section). FEM-based numerical simulations were conducted with COMSOL Multiphysics (version 5.3) for the O_2 and H_2O_2 gradients near the Pt- and Au-loaded wire array electrodes, respectively. Experimental [O_2]avg and [H_2O_2]avg values were compared with simulation results at different heights above the base of wire array (z), as shown in the exemplary case when consistent results of O_2 and H_2O_2 gradients were observed for k = (30, 3, 50) and E_{appl} = 0.5 V vs. RHE (Fig. S13B and Fig. S13C). The mean squared errors (MSEs) between FEM-based simulations and experimental results are 9.81×10^{-4} mM^2 and 4.84×10^{-6} mM^2 for O_2 and H_2O_2 gradients, respectively (see "Materials and Methods" section). Such consistence of results between experimental characterization and FEM-based simulations motivates us to use the augmented data from FEM-based simulations to establish neural networks to inversely predict O_2 and H_2O_2 gradients.

The established neuron networks display good accuracies for the $[O_2](z)$ and $[H_2O_2](z)$ distributions near the wire array electrodes loaded with Pt and Au electrocatalysts. We previously constructed multi-layer perception neuron networks (MLPNNs) that expand the explorable parameter space of wire array electrode morphologies in electrocatalytic reduction of N2.34 Here, we constructed MLPNNs that predicts $[O_2](z)$ and $[H_2O_2](z)$ based on inputs of $\{E_{appl}, k = (P, D, L)\}$, which were trained based on 10,000 data points augmented from the FEM-based simulations (see "Materials and Methods" section). As model-training process proceeds with an increasing number of epochs, monotonic decreases of the average MSE between the training and predicted data points (AMSE) for both the data sets of validation and training (red and black dots, respectively) were observed in Fig. 4C and 4D for the O_2 and H_2O_2 gradients near Pt- and Au-loaded wire array electrodes, respectively. The fact that the values of AMSEs against the validation date sets were similar to the ones from the training date sets in Fig. 4C and 4D indicates that there was no overfitting in the machine-learning process.[56] In the end, near-unity coefficients of determination ($R2 \rightarrow 1$) were observed for both MLPNNs (MLPNN 1 that predict O_2 gradients and MLPNN 2 that predict H_2O_2 gradients) (Fig. 4E and 4F, respectively). The values of AMSEs from the MLPNNs reach 1.74×10^{-4} mM^2 and 1.81×10^{-6} mM^2 for the prediction of $[O_2](z)$ and $[H_2O_2](z)$ based on inputs of $\{E_{appl}, k = (P, D, L)\}$, respectively. Such small values of AMSEs suggest good accuracies of the developed neural networks for the inverse design of desirable O_2 and H_2O_2 microenvironments.

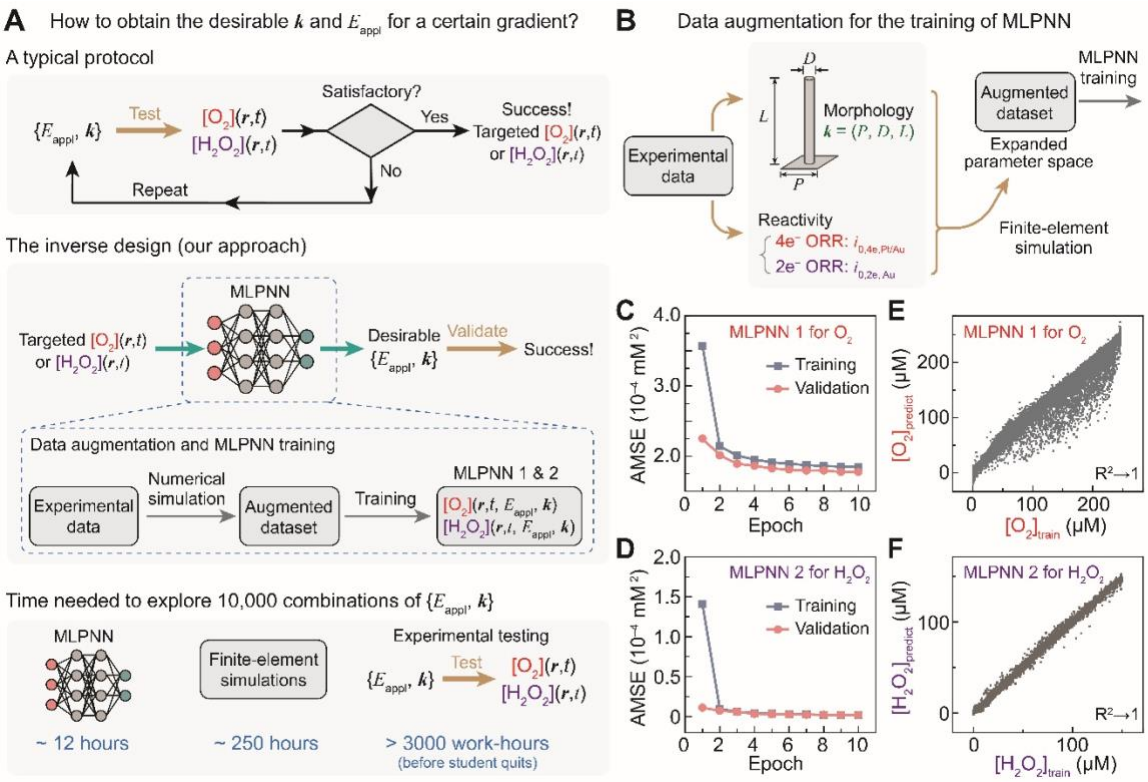


Figure 4. The development of inverse design for electrochemically generated O_2 and H_2O_2 gradients. **A** Comparison between the conventional protocol and our inverse design approach for the development of suitable experimental conditions, represented as $\{E_{appl}, \boldsymbol{k}\}$ in order to achieve desirable spatiotemporal distributions of O_2 and H_2O_2 concentrations ($[O_2](\boldsymbol{r}, t)$ and $[H_2O_2](\boldsymbol{r}, t)$, respectively). MLPNN 1 & 2, multiple-layer perceptron neural networks for O_2 and H_2O_2 gradients, respectively. **B** Protocols of data augmentation for the establishment of MLPNN. $i_{0,4e,Pt/Au}$ and $i_{0,2e,Au}$, the exchange current densities of four-electron and two-electron ORRs on Pt and/or Au electrocatalysts, respectively. **C** and **D** The averaged mean squared error (AMSE) in the training (blue) and validation (pink) datasets at different epochs for the gradients of O_2 (MLPNN 1 in **C**) and H_2O_2 (MLPNN 2 in **D**). **E** and **F** Comparisons between the MLPNN-predicted values ($[O_2]_{predict}$ and $[H_2O_2]_{predict}$) and training values ($[O_2]_{train}$ and $[H_2O_2]_{train}$) for the local average concentrations of O_2 **E** and H_2O_2 **F**, respectively. R^2, coefficient of determination.

Exemplary inverse design of O_2 and H_2O_2 microenvironments near wire array electrodes

Exemplary inverse design processes with direct biological relevance were experimentally tested with good predictabilities for the establishment of desirable O_2 and H_2O_2 microenvironments. In microbiology and microbial ecology, it is desirable to establish well-defined microenvironments whose sizes are 20~100 μm in order to mimic natural heterogenous distribution of biologically relevant extracellular species such as nutrients and other microbial resources.[57] Within such length scales, establishing microoxic niche (i.e. $[O_2] \sim 100$ μM) in the midst of an oxic external environment (Fig. 5A), prevalent in aquatic, terrestrial, and host-associated environments, is challenging yet desirable to understand the physiology and ecology of microaerophiles and advance our understanding of microbiomes;[2] extracellular H_2O_2 whose concentration can achieve 15 μM[58] (Fig. 5B) is also of particular interests in order to study microbial sensing, communal signaling, metabolic regulation, and genetic expression towards ROS.[48, 59] In such biological contexts, we aim to inversely design one O_2 gradient ($\Delta[O_2] \sim 100$ μM and $\Delta z \sim 40$ μm) and one H_2O_2 gradient ($\Delta[H_2O_2] \sim 15$ μM and $\Delta z \sim 100$ μm) based on our developed MLPNNs (Fig. 5A and 5B). We utilized the established MLPNNs that predict $[O_2](z)$ and $[H_2O_2](z)$ under different inputs of $\{E_{appl}, k = (P, D, L)\}$ and we scored the similarity percentages between the MLPNNs-predicted O_2/H_2O_2 gradients and the desirable ones. Fig. 5C and 5D exemplarily display the sliced mappings of similarity scores for the aforementioned O_2 and H_2O_2 gradients on Pt and Au-loaded wire arrays, respectively, as a function of $k = (P, D, L)$ at $E_{appl} = 0.5$ V vs. RHE. Such a multi-dimensional mapping, composed of 10,000 data points each in Fig. 5C and 5D, showcases the parameter spaces that are predicted to yield the desirable O_2 and H_2O_2 microenvironments within a certain relative uncertainty threshold (red region) (see "Materials and Methods" section). It is intriguing to note that there existed multiple different wire array morphologies to yield the same desirable O_2 and H_2O_2 gradients, which may not be straightforward intuitively. We estimated that

the determination of O_2/H_2O_2 gradients for one parameter coordinate in the space of {E_{appl}, k = (P, D, L)} will take ~4 sec for the MLPNNs-based method, ~90 sec from FEM-based simulations, and ~20 mins for the confocal characterization alone at one specific E_{appl} for a single wire-array morphology notwithstanding any time spent in any preceding protocols (see "Materials and Methods" section). Therefore, a comprehensive exploration of the parameter space {E_{appl}, k = (P, D, L)} with more than 104 trials as shown above is only possible with the use of MLPNN-based inverse design, because only the MLPNN is capable of screening 10,000 parameter combinations within a reasonable amount of time in practice (~12 hours) in comparison to the ones based on FEM (~250 hours, i.e. ~10 days) and experimental characterization (at least 3000 work-hours without considering any practical concerns) (Fig. 4A).

We also conducted a spot check for the predicted O_2 and H_2O_2 microenvironments by experimentally validations. Pt-based wire array electrode with k = (46, 6, 20) was picked from Fig. 5C as a desirable geometry, experimentally prepared (Fig. 1E), and experimented tested for the established O2 gradient at E_{appl} = 0.5 V vs. RHE. Satisfactory consistency with MSE = 5.63×10^{-4} mM^2 was achieved between the experimental and targeted values of $[O_2](z)$ (dots and line in Fig. 5E, respectively). Similarly, Au-based wire array electrode with k = (17, 3, 30) was picked from Fig. 5D, experimentally prepared (Fig. 1E), and experimented tested for the H_2O_2 gradient at E_{appl} = 0.45 V vs. RHE. We also observed satisfactory consistency with MSE = 7.22×10^{-6} mM^2 between the experimental and targeted values of $[H_2O_2](z)$ (dots and line in Fig. 5F, respectively). While we were unable to experimentally exhaust all of the predicted parameter space for the desirable microenvironments of O_2 and H_2O_2, our experimental validations offer convincing evidence for the validity of the developed MLPNN-based inverse design for future microbiology-related research.

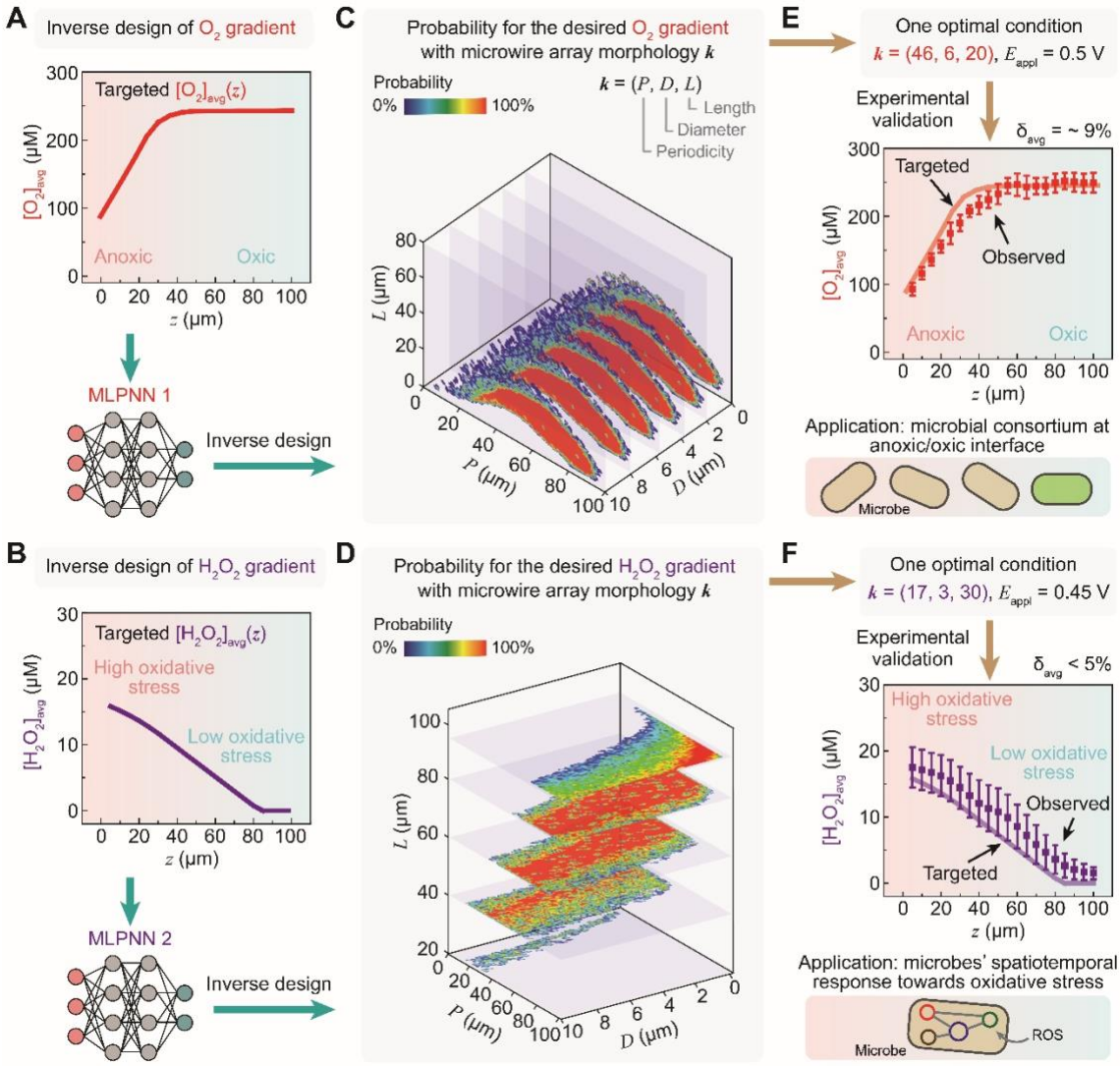


Figure 5. Experimental validations of MLPNN-assisted inverse design. **A** and **B** Targeted O_2 and H_2O_2 gradients ($[O_2]_{avg}(z)$ in **A** and $[H_2O_2]_{avg}(z)$ in **B**, respectively) for exemplary inverse design assisted by the developed MLPNNs. **C** and **D** the three-dimensional contour plots of the probabilities that MLPNN-predicted $[O_2]_{avg}(z)$ **C** and $[H_2O_2]_{avg}(z)$ **D** match the targeted ones as a function of P, D, L when E_{appl} = 0.5 V *vs.* RHE. **E** and **F**, Experimental characterizations of $[O_2]_{avg}(z)$ and $[H_2O_2]_{avg}(z)$ (scattered points) in comparison with the targeted ones (lines) when $\{E_{appl}, \boldsymbol{k}\}$ = (0.5, 46, 6, 20) in **E** and (0.45, 17, 3, 30) in **F** on Pt- and Au-loaded microwire arrays, respectively. The potential applications of those yielded gradients in microbiology are noted. Error bars represent standard deviations across multiple separate measurements in the device ($n \geq 3$).

Conclusion

In summary, we presented an ML-based inverse design strategy of O_2 and H_2O_2 concentration profiles with the use of electrochemical catalysis of ORR. We demonstrated a proof-of-concept closed-loop protocol of inversely designing O_2 and H_2O_2 gradients with properly designed microwire electrodes in PBS solution, the go-to culturing medium in microbiology. By achieving concentration differences and spatial resolutions that are relevant to microbial microenvironments, the demonstrated O_2 and H_2O_2 gradients are applicable for studies in microbiology. While the reported research focuses on one specific form of electrochemical boundary conditions, namely microwire array, the reported inverse-design procedures are generally applicable for any electrochemical systems that can be parameterized and analyzed by neural networks. As mass transport and concentration profiles in electrode's proximity are important components in electrochemistry, this work demonstrates the power of machine-learning-based inverse design in electrochemistry. Moreover, our results will lead to a general platform that inversely design suitable electrochemical systems for any targeted environments of O_2 and H_2O_2 in microbiology. Future research will focus on the platform's application of in vivo microbial communities and fundamental insights that can be fetched thanks to our system's unique capabilities. Moreover, since electrochemistry is capable of modulating any redox active species such as extracellular Fe(II)/(III) species,60 as well as other extracellular metabolites sensitive towards oxidative stress such as pyocyanin,61-63 our inverse design approach based on electrochemistry is capable of controlling the other microenvironments beyond O_2 and ROS and is generally applicable in the study of ubiquitous microenvironments in extracellular medium.

Materials and Methods

Chemical and materials

Indium-doped tin oxide (ITO)-coated glass slide (06499-AB) was purchased from SPI Supplies (30 ~ 60 Ω, 22 × 40 mm). Silver (Ag) paste (16040-30) was purchased from Ted Pella. Si wafers (*p*-type, boron-doped, <100>, 1~10 Ω•cm^{-1}) were purchased from University Wafer, Inc. Platinum (Pt) wires (CHI 115) and glassy carbon electrodes (CHI 104, diameter = 3 mm) were purchased from CH Instruments. Unless specially mentioned, all chemicals and materials were purchased from Sigma-Aldrich or VWR.

Experimental establishment of testing platform

Modified from a previously published protocol,[29] the experimental testing platform as shown in Fig. S7 consists of a fluidic cell with a three-electrode system, in order to electrochemically generate desirable gradients of O_2 and H_2O_2. As the working electrode in the setup, microwire array electrodes were fabricated similarly as previously described.[29] The microwire arrays were fabricated by photolithography with the use of deep reactive-ion etching process (DRIE). After treated in hexamethyldisilazane (HMDS) vapor for 10 minutes, pre-cleaned Si wafers were coated by photoresist (MicroChemicals, AZ5214E; 3000 rpm spin-coating) with a soft baking (100 °C for 75 sec), exposed in the hard contact mode by a contact aligner (Carl Suss MA6), hard-baked (120 °C, 5 min), and developed to generate the periodic patterns for the array (mixture of MicroChemicals AZ400K water, 1:4 volume ratio). After created by DRIE (Unaxis Versaline FDSE III), microwire arrays of desirable lengths were annealed under ambient air at 1050 °C for 9.5 hours to yield the insulating oxide layer, coated by 500-nm ITO through reactive sputtering (Denton Discovery 550 sputtering System), and finally deposited with a 7-nm layer of Pt or Au by a Anatech Hummer 6.2 sputtering system. The structure of wire array was examined by a SEM (ZEISS Supra 40VP SEM) and the element distribution was examined by a SEM (JEOL JSM-6700F) equipped with energy dispersive X-ray spectroscopy (EDS, Ametek). Patterned ITO-

coated glass slides were used as the counter electrodes in the established testing platform, after the selective removal of ITO by 6 M HCl in undesirable areas on the glass slides. Ag paint as the pseudo-reference electrode was applied on select area of the ITO-coated slides as to cover a 5 mm × 5 mm square with a layer of silver, serving as the reference electrode, while Pt was deposited on other ITO-coated areas for the creation of counter electrode in the setup. An optically transparent fluidic cell of 250-μm height was constructed by assembling the microwire array electrodes with the pre-fabricated ITO glass slides, while a Gamry Interface 1010B potentiostat was used to enforce electrochemical driving force. The setup was mounted on an inverted laser confocal microscope (Leica SP8 SMD) with sufficient working distance (680 μm) and syringe pump was used to maintain a fixed liquid flow rate. As we particularly ensure the accuracy of applied electrochemical potentials, cyclic voltammetry in a standard ZoBell's solution (3.3 mM $K_3Fe(CN)_6$, 3.3 mM $K_4Fe(CN)_6$, and 0.1 M KCl, 0.43 V *vs.* Standard Hydrogen Electrode (SHE)) was employed to calibrate the electrochemical potential of Ag-based pseudo reference electrode as shown in Fig. S14.[64, 65] The Ag-based pseudo-reference electrode was calibrated as 0.75 V *vs.* RHE in PBS solution at pH = 7.4.

Electrochemical characterization of the deposited electrocatalysts

While the deposited Pt electrocatalysts have been characterized in our prior report,[29] experiments were conducted to analyze the electrocatalytic activities of ORR for the coated Au electrocatalysts. An experiment of rotating ring-disk electrode (Pine Research, Inc., AFE6R1PT) was conducted in PBS solution using a setup with a Pt-wire counter electrode, Ag/AgCl (1M KCl) reference electrode and a modulated speed rotator (Pine Research, Inc., AFMSRCE). While Pt ring electrode was kept at 1.9 V *vs.* RHE, linear scan voltammograms (20 mV/s) were recorded between 0.1 V and 1.1 V *vs.* RHE with different rotating speeds (100 rpm, 225 rpm, 400 rpm, 625 rpm, 900 rpm,

1225 rpm, 1600 rpm, and 2025 rpm) in electrolytes saturated with O_2 and N_2, respectively. The measurements in N_2-saturated solution were used as signal background.

Quantification and calibration of O_2 concentration profiles

Aerated PBS solution consisting of 150 μM $Ru(phen)_3Cl_2$ solution was prepared in the dark, and fed into the assembled testing platform at a flow rate of 0.8 mL/min. The phosphorescence intensity mapping under the confocal microscopy was measured in a one-minute time sequence. During the one-minute confocal microscopy measurement, a programmed 30-s electrolysis was performed, with a particular potential on the working electrode from t = 15 sec to t = 45 sec. The excitation wavelength was set as 470 nm and we gathered emission intensity from 600 nm to 640 nm as phosphorescence emission intensity I_p. We defined the phosphorescence emission intensity with no potential applied as I_0. Normalized phosphorescence emission intensity I_{pn} was defined as $I_{pn} = I_p/I_0$. The phosphorescence emission intensity distribution was further translated into the concentration profiles based the linear relationship between O_2 concentration and the inverse of I_{pn} (noted as I_{pn}^{-1}) that was experimentally determined. $Ru(phen)_3Cl_2$-containing PBS solutions of different O_2 concentrations, ranging from $[O_2]$ = 25 μM to 375 μM, was prepared by bubbling N_2/O_2 mixture of tunable ratio through the solution, and pumped into the assembled fluidic device (0.8 mL/min) for calibration.

Quantification and calibration of H_2O_2 concentration profiles

Aerated PBS solution consisting of 0.2 U/mL HRP and 120 μM Amplex Red (1× working solution) was prepared in the dark, and fed into the assembled testing platform at a flow rate of 0.8 mL/min. The fluorescent intensity mapping under the confocal microscopy was conducted in a one-minute time sequence. During the one-minute confocal microscopy measurement, a programmed 30-sec electrolysis with a particular potential on the working electrode from t = 20 sec to t = 50 sec. The

excitation wavelength was set as 550 nm and we gathered emission intensity from 590 nm to 650 nm as fluorescence emission intensity I_f. The fluorescence emission intensity distribution was further translated into the concentration profiles based on the corresponding calibration curves. In the experiments of calibrating H_2O_2 concentrations, darkly prepared PBS solution consisting of 0.4 U/mL HRP and 240 μM Amplex Red was combined with PBS solution of different H_2O_2 concentrations, ranging from 5 μM to 60 μM, and was pumped into the assembled devices for the measurement of confocal microscopy. We found that the H_2O_2-induced fluorescence intensity I_f is also dependent on the specific morphologies of wire array (***k***) and the distance away from the bottom of wire array (z) (Fig. S15), owing to the scattering and, possibly, optical absorption of the wire array electrodes.[66-68] Therefore, individual calibration curves were determined for every z location in wire arrays with all possible ***k*** values (Fig. S9 to Fig. S12). Specific I_f correction was made to compensate the difference between calibration experiment and gradient optical detection due to practical restriction (Supplementary text and Fig. S16).

FEM-based numerical simulations for O_2 gradient and H_2O_2 gradients

FEM simulation of both O_2 gradient and H_2O_2 gradient was achieved in "electroanalysis" module in COMSOL Multiphysics (Version 5.3).

Geometry description: The shape of wire was represented by a column with diameter of D and length of L. We located each wire in the center a cuboid of $P \times P \times 200$ μm and the difference between the cuboid and column was geometrically defined for the electrolyte. For each point in the electrolyte, if its distance from the top of the wire not smaller than the diffusion distance, d_D, we considered it belonged to the bulk electrolyte in which $[O_2] = 246$ μM independent of time. The boundary surface was defined as the area of which the distance to the top of the wire is d_D. On the boundary surface, the $[O_2]$ was the same as that in the bulk. Periodic boundary condition

was applied to describe wire array. The value of d_D was set as 20 μm for O_2 gradient simulation on Pt-loaded wire array electrodes.[29] For O_2 gradient and H_2O_2 gradient simulation on Au-loaded wire array electrodes, d_D was measured as 50 μm (Fig. S17).

Transport properties: Diffusion of oxygen and hydrogen peroxide was simulated based on the following equations (1) and (2). D_{O_2} and $D_{H_2O_2}$ are the diffusion coefficients of oxygen and hydrogen peroxide in aqueous solution, which were 2.2×10^{-9} m^2/s and 1.5×10^{-9} m^2/s respectively.

$$\frac{\partial[O_2]}{\partial t} = D_{O_2}\nabla^2[O_2] \quad (1)$$

$$\frac{\partial[H_2O_2]}{\partial t} = D_{H_2O_2}\nabla^2[H_2O_2] \quad (2)$$

Electroanalysis: The potential window of E_{appl} is from 0.6 V *vs.* RHE to 0.2 V *vs.* RHE.

On the surface of Pt-loaded wire array electrodes, 4e-ORR took place within the potential window.

$$O_2 + 4H^+ + 4e^- \rightarrow 2H_2O$$

On the electrode surface of Au-loaded wire array electrodes, two-electron and four-electron ORR reactions (2e-ORR and 4e-ORR, respectively) took place at the same time as potential-dependent selectivity.

$$\text{2e-ORR: } O_2 + 2H^+ + 2e^- \rightarrow H_2O_2$$

$$\text{4e-ORR: } O_2 + 4H^+ + 4e^- \rightarrow 2H_2O$$

On the surface of Pt-loaded wire array electrodes, the supply-consumption equilibrium was simulated as the equations (3) and (4) in the formalism of concentration-dependent Tafel kinetics,[27]

$$i_{loc} = i_{4e} = -i_{0,4e,Pt}\frac{[O_2]}{C_{O_2}}exp\left(\frac{-\alpha_c F\eta_{Pt}n_{4e}}{RT}\right) \quad (3)$$

$$J_{O_2} = \frac{i_{loc}}{4F} \quad (4)$$

Here i_{4e} denotes the current density of 4e-ORR, $i_{0,4e,Pt}$ the exchange current density of 4e-ORR, i_{loc} the local current density of 4-electron reduction of O_2 on the electrode surface, J_{O_2} the local flux of O_2 consumption from electrolyte, and η_{Pt} is the over potential that is defined as the difference between E_{appl} and the standard redox potential of O_2/H_2O, $E^0_{O_2/H_2O}$ (1.23 V *vs.* RHE). α_c equals 0.5 as the transfer coefficient, *F* denotes the Faraday constant, *R* is the gas constant, *T* is the temperature, and n_{4e} is electron transfer number before the rate determining step of 4e-ORR, of which the value is 1 (Fig. S4). The local oxygen concentration is denoted as $[O_2]$. C_{O_2} is oxygen concentration in air-saturated water at *T*. The above equation follows the textbook equations that accounting for the mass transport and chemical stoichiometry at electrode interfaces.[27] On Pt-loaded wire array electrodes, $i_{0,4e,Pt}$ is found to be 3.0×10^{-6} A/m^2 (Fig. S4).

On the electrode surface of Au-loaded wire array electrodes, the supply-consumption equilibrium was simulated as the equations (5) to (7).

$$i_{loc} = i_{4e} + i_{2e} = -i_{0,4e,Au}\frac{[O_2]}{C_{O_2}}exp\left(\frac{-\alpha_c F\eta_{Au}n_{4e}}{RT}\right) - i_{0,2e,Au}\frac{[O_2]}{C_{O_2}}exp\left(\frac{-\alpha_c F\eta_{Au}n_{2e}}{RT}\right) \quad (5)$$

$$J_{O_2} = \frac{i_{4e}}{4F} + \frac{i_{2e}}{2F} \quad (6)$$

$$J_{H_2O_2} = -\frac{i_{2e}}{2F} \quad (7)$$

While most of the definitions of variables in the case of Au-loaded wire array electrodes with equation (5), (6), and (7) are the same as the Pt-loaded case in equation (3) and (4), we noted that i_{loc} instead stands for the local current density of both 2-electron and 4-electron reduction of O_2 on the electrode surface, i_{2e} The current density of 2e-ORR, and n_{2e} is electron transfer number before the rate determining step of 2e-ORR, of which the value is 0.7 (Fig. S5). On Au-loaded wire array electrodes, the catalysis selectivity of Au towards 4e-ORR and 2e-ORR are dictated by

the exchange current densities $i_{0,4e,Au}$ and $i_{0,2e,Au}$ respectively. $i_{0,4e,Au}$ is 2.0×10^{-8} A/m^2 and $i_{0,2e,Au}$ is 8.0×10^{-7} A/m^2 based on literature and measurement (Fig. S5).[69, 70]

Based on comparison with experimental gradients, the average mean squared error (AMSE) of O_2 gradient simulation and H_2O_2 gradient simulation on Pt-loaded wire array electrodes Au-loaded wire array electrodes is 9.81×10^{-4} mM^2 and 4.84×10^{-6} mM^2 respectively. The range of E_{appl} was set within the ORR potential windows, from 0.6 V *vs.* RHE to 0.2 V *vs.* RHE. Besides, a three-dimensional block was defined in the space of (*P*, *D*, *L*) as the range of morphology, from *P* = 1 to 100 μm, *L* = 1 to 150 μm, and *D* = 0.2 to 10 μm. By using Simulink in Matlab, we were able to generate random ***k*** value within the morphology space and calculate the corresponding gradient curves. To fulfill the machine learning functions, gradient profiles under 10000 experimental conditions were included in the dataset for each developed multiple-layer perceptron neural networks respectively.

Model selection and training

The implementation of all machine-learning code was done on a MacBook Pro with a 1.4 GHz quad-core Intel Core i5 processor and 8GB of RAM, with code specifically deployed using the JupyterLab Notebook, a Python-based programming platform widely used in data science and machine learning.[34] In this paper, we selected multi-layer perception neuron networks (MLPNNs) as the machine learning model for inverse design. We use FEM-simulated gradients to develop the MLPNNs. Gradient data was imported from FEM results in the form of .csv files and combined into library of data. Prior to the model development, a random 20% of the data was split from the whole dataset for later model validation. The rest of the data was split as training data and validation data of which the percentage were 65% and 15% respectively. Multiple cycles of model training, each cycle termed as one epoch, were conducted in order to develop the targeted

MLPNNs. In each epoch, MLPNNs will be trained from the training data, followed by a validation process in the validation data set used to provide estimates of final model accuracy after each round of training. The total machine learning process will include ten epochs of forward and backward propagations. The accuracies of the developed MLPNNs model as a function of epoch numbers are plotted in Fig. 4*C* and 4*D*.

As described in the section above, the training data set was a collection of concentration gradient under different E_{appl} and $\boldsymbol{k}$ value. We used average mean squared error (AMSE) and standard deviation (SD), more detailly defined below from equations (8) to (11). to quantify the ability of machine learning models to correctly reveal the connection between gradient curves and the two impacting factors, E_{appl} and $\boldsymbol{k}$. In model selection, we selected multiple-layer perceptron neural networks for gradient prediction due to the low AMSE and SD value gradient prediction (Fig. S18).

For both O_2 and H_2O_2 gradient, the predicted curve included 20 local concentration data along the wire array from Z = 5 μm to 100 μm every 5 μm. Z is defined as the distance from the bottom of wire array. In O_2 gradient prediction, MSE and SD are defined as the following equations.

$$MSE_O = \frac{1}{20}\sum_{n=1}^{20}([O_2]_{NNs} - [O_2]_{real})^2 \quad (8)$$

$$SD_O = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(MSE_{O\,n} - AMSE_o)^2} \quad (9)$$

In equations (8) and (9), $[O_2]_{NNs}$ stands for the predicted oxygen concentration from MLPNNs predicting O_2 gradient on Pt-loaded wire array. $[O_2]_{real}$ is the oxygen concentration in the simulation data set. MSE_O is defined based on the average square of concentration difference over the whole gradient profile. $AMSE_o$ is the average MSE over the whole data set. We calculated the

standard deviation over data under a wide range of experimental conditions to evaluate the overall precision of predictions from MLPNNs predicting O_2 gradient on Pt-loaded wire array.

MSE and SD in H_2O_2 gradient were defined in a similar pattern. In equations (10) and (11), $[H_2O_2]_{NNs}$ stands for the predicted hydrogen peroxide concentration from MLPNNs predicting H_2O_2 gradient on Au-loaded wire array. $[H_2O_2]_{real}$ is the hydrogen peroxide concentration in the simulation data set. MSE_H is defined based on the average square of concentration difference over the whole gradient profile. $AMSE_H$ is the average MSE over the whole data set.

$$MSE_H = \frac{1}{20}\sum_{n=1}^{20}([H_2O_2]_{NNs} - [H_2O_2]_{real})^2 \quad (10)$$

$$SD_H = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(MSE_{H_n} - AMSE_H\right)^2} \quad (11)$$

Morphology prediction for desired O_2 gradient and H_2O_2 gradient

In morphology prediction for desired O_2 gradient, we assigned the E_{appl} = 0.5 V *vs.* RHE. Initially, MLPNNs predicting O_2 gradient on Pt-loaded wire array randomly selected one morphology in the morphology space and calculated the corresponding O_2 gradient. The similarity score between the calculated O_2 gradient and target O_2 gradient, S_O, was quantified by equation (12). In the prediction process, the neural networks would find out the top 10,000 morphologies with highest S_O values. In equation (12), $[O_2]_T$ was the local oxygen concentration in the target O_2 gradient and $[O_2]_{NNs}$ was the calculated local oxygen concentration. Similarity score, S_O, is the average relative error among a collection of different *z* values, C_O. $C_O = [5, 10, 15, 25, 35, 45, 70]$, unit is μm.

$$S_O = 1 - 100\% \times \frac{1}{7}\sum_{z\in C_O}^{N}\frac{|[O_2]_{NNs} - [O_2]_T|}{[O_2]_T} \quad (12)$$

In morphology prediction for desired H_2O_2 gradient, we assigned the E_{appl} = 0.45 V *vs.* RHE. Initially, MLPNNs predicting H_2O_2 gradient on Au-loaded wire array randomly selected one morphology in the morphology space and calculated the corresponding H_2O_2 gradient. The similarity score between the calculated H_2O_2 gradient and target H_2O_2 gradient, S_H, was quantified by equation (13). In the prediction process, the neural networks would find out the top 10000 morphologies with highest S_H values. In equation (13), $[H_2O_2]_T$ was the local hydrogen peroxide concentration in the target H_2O_2 gradient and $[H_2O_2]_{NNs}$ was the calculated local hydrogen peroxide concentration. Similarity score, S_H, is the average relative error among a collection of different *z* values, C_H. $C_H = [5, 20, 30, 40, 55, 70, 80]$, unit is μm.

$$S_H = 1 - 100\% \times \frac{1}{7} \sum_{z \in C_H}^{N} \frac{|[H_2O_2]_{NNs} - [H_2O_2]_T|}{[H_2O_2]_T} \quad (13)$$

In the sliced mapping of H_2O_2 similarity score, the similarity score at *L* = 5 μm was the average of from *L* = 0 μm to 10 μm. The similarity score at *L* = 20 μm was the average of from *L* = 10 μm to 30 μm. The similarity score at *L* = 40 μm was the average of from *L* = 30 μm to 50 μm. The similarity score at *L* = 60 μm was the average of from *L* = 50 μm to 70 μm. The similarity score at *L* = 80 μm was the average of from *L* = 70 μm to 90 μm. The similarity score at *L* = 95 μm was the average of from *L* = 90 μm to 100 μm.

References

1. M. T. Madigan, J. M. Martinko, J. Parker, “Microbial Ecosystems” in Brock Biology of Microorganisms. 15th edition, M. T. Madigan, J. M. Martinko, J. Parker, (Pearson Prentice Hall, Upper Saddle River, NJ, 2018), pp. 651–686.
2. R. L. Morris, T. M. Schmidt, Shallow breathing: bacterial life at low O_2. *Nat. Rev. Microbiol.* **11,** 205−212 (2013).
3. P. S. Stewart, M. J. Franklin, Physiological heterogeneity in biofilms. *Nat. Rev. Microbiol.* **6,** 199−210 (2008).
4. T. J. Battin, K. Besemer, M. M. Bengtsson, A. M. Romani, A. I. Packmann, The ecology and biogeochemistry of stream biofilms. *Nat. Rev. Microbiol.* **14,** 251−263 (2016).
5. I. Guzmán-Soto et al., Mimicking biofilm formation and development: Recent progress in *in vitro* and *in vivo* biofilm models. *iScience* **24,** 102443 (2021).
6. J. M. Berg, J. L. Tymoczko, L. Stryer, G. J. Gatto Jr, “Oxidative Phosphorylation” in Biochemistry. 8th edition, J. M. Berg, J. L. Tymoczko, L. Stryer, G. J. Gatto Jr, (Macmillan, 2015), pp. 523–564.
7. D. de Beer, P. Stoodley, F. Roe, Z. Lewandowski, Effects of biofilm structures on oxygen distribution and mass transport. *Biotechnol. Bioeng.* **43,** 1131−1138 (1994).
8. J. A. Imlay, Where in the world do bacteria experience oxidative stress? *Environ. Microbiol.* **21,** 521−530 (2019).
9. Z. Lu, R. Sethu, J. A. Imlay, Endogenous superoxide is a key effector of the oxygen sensitivity of a model obligate anaerobe. *Proc. Natl. Acad. Sci. U.S.A.* **115,** E3266 (2018).

10. D. Klementiev Alexander, Z. Jin, M. Whiteley, R. Parsek Matthew, Micron scale spatial measurement of the O_2 gradient surrounding a bacterial biofilm in real time. *mBio* **11,** e02536−20 (2020).

11. L. S. Downing, R. Nerenberg, Effect of oxygen gradients on the activity and microbial community structure of a nitrifying, membrane-aerated biofilm. *Biotechnol. Bioeng.* **101,** 1193−1204 (2008).

12. D. Dar, N. Dar, L. Cai, D. K. Newman, Spatial transcriptomics of planktonic and sessile bacterial populations at single-cell resolution. *Science* **373,** eabi4882 (2021).

13. M. Carmona et al., Monitoring cytosolic H_2O_2 fluctuations arising from altered plasma membrane gradients or from mitochondrial activity. *Nat. Commun.* **10,** 4526 (2019).

14. C. Virgile et al., Engineering bacterial motility towards hydrogen-peroxide. *PLoS One* **13,** e0196999 (2018).

15. D. M. Lewis, M. R. Blatchley, K. M. Park, S. Gerecht, O_2-controllable hydrogels for studying cellular responses to hypoxic gradients in three dimensions *in vitro* and *in vivo*. *Nat. Protoc.* **12,** 1620−1638 (2017).

16. P. C. Thomas, S. R. Raghavan, S. P. Forry, Regulating oxygen levels in a microfluidic device. *Anal. Chem.* **83,** 8821−8824 (2011).

17. M. Polinkovsky, E. Gutierrez, A. Levchenko, A. Groisman, Fine temporal control of the medium gas content and acidity and on-chip generation of series of oxygen concentrations for cell cultures. *Lab Chip* **9,** 1073−1084 (2009).

18. K. Campbell, A. Groisman, Generation of complex concentration profiles in microchannels in a logarithmically small number of steps. *Lab Chip* **7,** 264−272 (2007).

19. M. L. Rexius-Hall, G. Mauleon, A. B. Malik, J. Rehman, D. T. Eddington, Microfluidic platform generates oxygen landscapes for localized hypoxic activation. *Lab Chip* **14,** 4688−4695 (2014).

20. J. L. Moran, P. M. Wheat, N. A. Marine, J. D. Posner, Chemokinesis-driven accumulation of active colloids in low-mobility regions of fuel gradients. *Sci. Rep.* **11,** 4785 (2021).

21. J. F. Lo, E. Sinkala, D. T. Eddington, Oxygen gradients for open well cellular cultures via microfluidic substrates. *Lab Chip* **10,** 2394−2401 (2010).

22. H.-H. Hsu, P.-L. Ko, H.-M. Wu, H.-C. Lin, C.-K. Wang, Y.-C. Tung, Study 3D endothelial cell network formation under various oxygen microenvironment and hydrogel composition combinations using upside-down microfluidic devices. *Small* **17,** 2006091 (2021).

23. M. R. Blatchley, F. Hall, S. Wang, H. C. Pruitt, S. Gerecht, Hypoxia and matrix viscoelasticity sequentially regulate endothelial progenitor cluster-based vasculogenesis. *Sci. Adv.* **5,** eaau7518 (2019).

24. Y.-A. Chen et. al., Generation of oxygen gradients in microfluidic devices for cell culture using spatially confined chemical reactions. *Lab Chip* **11,** 3626−3633 (2011).

25. M. D. Brennan, M. L. Rexius-Hall, L. J. Elgass, D. T. Eddington, Oxygen control with microfluidics. *Lab Chip* **14,** 4305−4318 (2014).

26. D. B. Weibel, G. M. Whitesides, Applications of microfluidics in chemical biology. *Curr. Opin. Chem. Biol.* **10,** 584−591 (2006).

27. A. J. Bard, L. R. Faulkner, "MASS TRANSFER BY MIGRATION AND DIFFUSION" in Electrochemical Methods: Fundamentals and Applications, 2nd edition, A. J. Bard, L. R. Faulkner, (Wiley, New York, NY, 2000), pp. 137−155.

28. J. Huskens, S. O. Krabbenborg, Electrochemically Generated Gradients, *Angew. Chem. Int. Ed.*, **53,** 9152−9167 (2014)

29. S. Lu, X. Guan, C. Liu, Electricity-powered artificial root nodule. *Nat. Commun* **11,** 1505 (2020).

30. A. Géron, "Introduction to Artificial Neural Networks with Keras" in Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems, A. Géron, (O'Reilly Media, Inc., Sebastopol, CA, 2019), pp. 279−313.

31. A. Mansouri Tehrani et al., Machine learning directed search for ultraincompressible, superhard materials. *J. Am. Chem. Soc.* **140,** 9844−9853 (2018).

32. J. Newman, K. E. Thomas-Alyea, "NUMERICAL SOLUTION OF COUPLED, ORDINARY DIFFERENTIAL EQUATIONS" in Electrochemical systems, 3rd edition, J. Newman, K. E. Thomas-Alyea, (Wiley, New York, NY, 2012), pp. 611−634.

33. E. J. F. Dickinson, H. Ekström, E. Fontes, COMSOL Multiphysics®: Finite element software for electrochemical analysis. A mini-review. *Electrochem. Commun.* **40,** 71−74 (2014).

34. B. B. Hoar, S. Lu, C. Liu, Machine-learning-enabled exploration of morphology influence on wire-array electrodes for electrochemical nitrogen fixation. *J. Phys. Chem. Lett.* **11,** 4625−4630 (2020).

35. A. Kulkarni, S. Siahrostami, A. Patel, J. K. Nørskov, Understanding catalytic activity trends in the oxygen reduction reaction. *Chem. Rev.* **118,** 2302−2312 (2018).

36. D. Zhao et al., Atomic site electrocatalysts for water splitting, oxygen reduction and selective oxidation. *Chem. Soc. Rev.* **49,** 2215−2264 (2020).

37. J.-C. Dong et al., In situ Raman spectroscopic evidence for oxygen reduction reaction intermediates at platinum single-crystal surfaces. *Nat. Energy* **4,** 60−67 (2019).

38. T. Kumeda, H. Tajiri, O. Sakata, N. Hoshi, M. Nakamura, Effect of hydrophobic cations on the oxygen reduction reaction on single–crystal platinum electrodes. *Nat. Commun.* **9,** 4378 (2018).

39. S. C. Perry et al., Electrochemical synthesis of hydrogen peroxide from water and oxygen. *Nat. Rev. Chem.* **3,** 442−458 (2019).

40. J. L. Terrell et al., Bioelectronic control of a microbial community using surface-assembled electrogenetic cells to route signals. *Nat. Nanotechnol.* **16,** 688−697 (2021).

41. C. Xiang, A. C. Meng, N. S. Lewis, Evaluation and optimization of mass transport of redox species in silicon microwire-array photoelectrodes. *Proc. Natl. Acad. Sci. U.S.A.* **109,** 15622−15627 (2012).

42. C. Xiang et al., Modeling, simulation, and implementation of solar-driven water-splitting devices. *Angew. Chem. Int. Ed.* **55,** 12974−12988 (2016).

43. X.-d. Wang, O. S. Wolfbeis, Optical methods for sensing and imaging oxygen: materials, spectroscopies and applications. *Chem. Soc. Rev.* **43,** 3666−3761 (2014).

44. J. Varchola et al., Temperature and oxygen-concentration dependence of singlet oxygen production by RuPhen as induced by quasi-continuous excitation. *Photochem. Photobiol. Sci.* **13,** 1781−1787 (2014).

45. N. I. Kolev, "Solubility of O_2, N_2, H_2 and CO_2 in water" in Multiphase flow dynamics 4: Turbulence, Gas Adsorption and Release, Diesel Fuel Properties, N. I. Kolev, (Springer, Berlin Heidelberg, 2012), pp. 209−239.

46. M. Ackermann, A functional perspective on phenotypic heterogeneity in microorganisms. *Nat. Rev. Microbiol.* **13,** 497−508 (2015).

47. J. J. Harrison, H. Ceri, R. J. Turner, Multimetal resistance and tolerance in microbial biofilms. *Nat. Rev. Microbiol.* **5,** 928−938 (2007).

48. J. A. Imlay, The molecular mechanisms and physiological consequences of oxidative stress: lessons from a model bacterium. *Nat. Rev. Microbiol.* **11,** 443−454 (2013).

49. N. Schürmann et al., Myeloperoxidase targets oxidative host attacks to *Salmonella* and prevents collateral tissue damage. *Nat. Microbiol.* **2,** 16268 (2017).

50. L. C. Seaver, J. A. Imlay, Hydrogen peroxide fluxes and compartmentalization inside growing *Escherichia coli. J. Bacteriol.* **183,** 7182−7189 (2001).

51. G. H. Seong, J. Heo, R. M. Crooks, Measurement of enzyme kinetics using a continuous-flow microfluidic system. *Anal. Chem.* **75,** 3161−3167 (2003).

52. A. Sen, J. A. Imlay, How microbes defend themselves from incoming hydrogen peroxide. *Front. Immunol.* **12,** 667343 (2021).

53. M. F. Lagadec, R. Zahn, V. Wood, Characterization and performance evaluation of lithium-ion battery separators. *Nat. Energy* **4,** 16−25 (2019).

54. P. Peljo, M. D. Scanlon, T. J. Stockmann, Simulations employing finite element method at liquid|liquid interfaces. *Curr. Opin. Electrochem.* **7,** 200−207 (2018).

55. C. Lin, R. G. Compton, Understanding mass transport influenced electrocatalysis at the nanoscale via numerical simulation. *Curr. Opin. Electrochem.* **14,** 186−199 (2019).

56. T. Dietterich, Overfitting and undercomputing in machine learning. *ACM Comput. Surv.* **27,** 326−327 (1995).

57. R. Stocker, Marine microbes see a sea of gradients. *Science* **338,** 628−633 (2012).

58. N. S. Jakubovics, S. R. Gill, M. M. Vickerman, P. E. Kolenbrander, Role of hydrogen peroxide in competition and cooperation between *Streptococcus gordonii* and *Actinomyces naeslundii*. *FEMS Microbiol. Ecol.* **66,** 637−644 (2008).

59. Z. Lu, J. A. Imlay, When anaerobes encounter oxygen: mechanisms of oxygen toxicity, tolerance and defence. *Nat. Rev. Microbiol.* **19,** 774−785 (2021).

60. E. D. Melton, E. D. Swanner, S. Behrens, C. Schmidt, A. Kappler, The interplay of microbially mediated and abiotic reactions in the biogeochemical Fe cycle. *Nat. Rev. Microbiol.* **12,** 797−808 (2014).

61. M. E. Hernandez, D. K. Newman, Extracellular electron transfer. *Cell. Mol. Life Sci.* **58,** 1562−1571 (2001).

62. A. Price-Whelan, L. E. P. Dietrich, D. K. Newman, Rethinking 'secondary' metabolism: physiological roles for phenazine antibiotics. *Nat. Chem. Biol.* **2,** 71−78 (2006).

63. E. K. Perry, L. A. Meirelles, D. K. Newman, From the soil to the clinic: the impact of microbial secondary metabolites on antibiotic tolerance and resistance. *Nat. Rev. Microbiol.* **20,** 129−142 (2022).

64. D. K. Nordstrom, Thermochemical redox equilibria of ZoBell's solution. *Geochim. Cosmochim. Acta* **41,** 1835–1841 (1977).

65. C. E. ZoBell, Studies on redox potential of marine Sediments. *AAPG Bull.* **30,** 477–513 (1946).

66. G. Shalev, S. W. Schmitt, G. Brönstrup, S. Christiansen, Maximizing the ultimate absorption efficiency of vertically-aligned semiconductor nanowire arrays with wires of a low absorption cross-section. *Nano Energy* **12,** 801−809 (2015).

67. H. Alaeian, A. C. Atre, J. A. Dionne, Optimized light absorption in Si wire array solar cells. *J. Opt.* **14,** 024006 (2012).

68. L. Yu et al., Understanding Light Harvesting in Radial Junction Amorphous Silicon Thin Film Solar Cells. *Sci. Rep.* **4,** 4357 (2014).

69. G. Gotti, D. Evrard, K. Fajerwerg, P. Gros, Oxygen reduction reaction features in neutral media on glassy carbon electrode functionalized by chemically prepared gold nanoparticles. *J. Solid State Electrochem.* **20,** 1539−1550 (2016).

70. C. Villena, J. Losada, B. Alonso, C. M. Casado, M. P. G. Armada, Easy preparation of electrode surfaces with dispersed size-controlled Au nanoparticles by electrodeposited PPI-dendrimers as templates. *J. Electrochem. Soc.* **164,** H396−H406 (2017).

Supplementary Information

Supplementary Text

Fluorescence intensity correction in H_2O_2 concentration calibration experiment: In the H_2O_2 concentration calibration experiment, H_2O_2 solution mixed with 2× working solution (PBS solution consisting of 0.4 U/mL HRP and 240 μM Amplex Red) in a mixing tee (IDEX, U-466). In the H_2O_2 gradient measurement experiment, the mixing area was the working electrode area. In the calibration experiment setup, it required extra time for liquid to travel from mixing tee to the working electrode resulting in higher fluorescence emission intensity in calibration experiments under the same H_2O_2 concentration. The relationship between increased fluorescence intensity and extra time was revealed by measuring the I_f difference between the upstream and downstream of the wire array electrode (Fig. S16). In each calibration experiment, I_f was corrected based on the relationship.

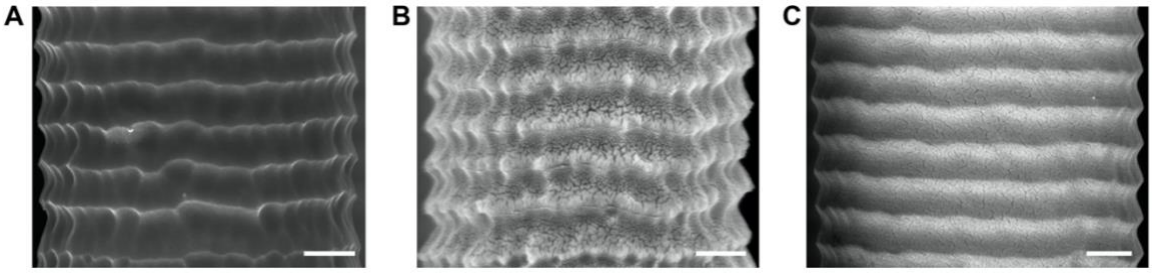


Figure S1. Scanning electron microscopy (SEM) images for the surface of wire array electrode without the deposition of Au/Pt catalysts (*A*), loaded with Pt electrocatalyst (*B*), and loaded with Au electrocatalyst (*C*). Scale bar, 500 nm.

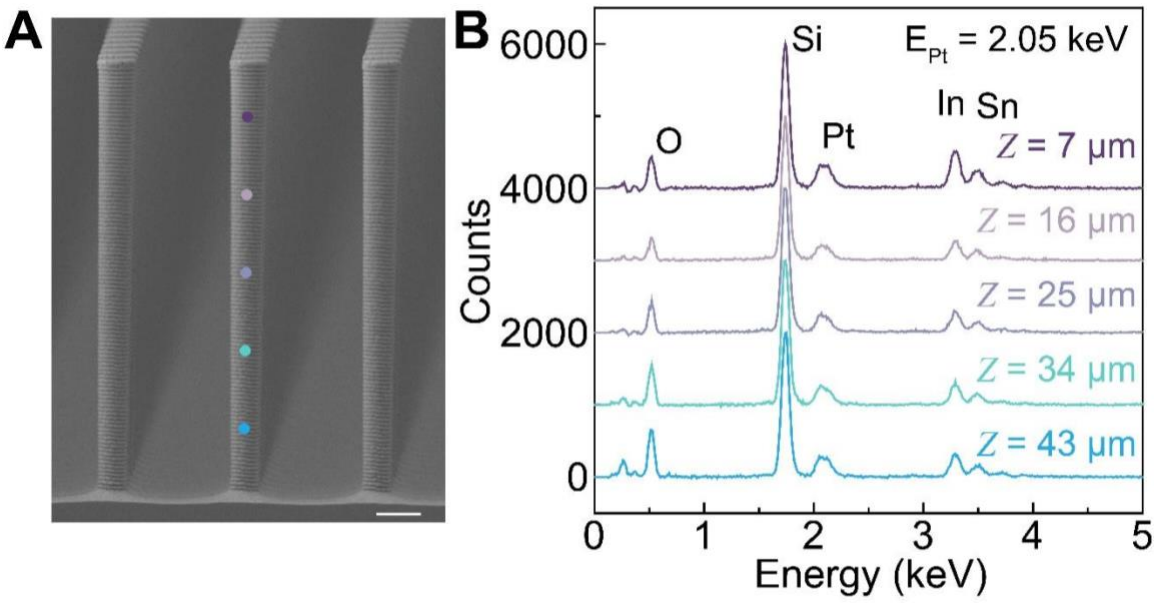


Figure S2. Morphological characterization of Pt-coated wire array electrodes with ***k*** = (15, 4, 50) (*A*) The cross-sectional scanning electron microscopy (SEM) image depicting the side view of wire array with ***k*** = (15, 4, 50) (*B*) Energy-dispersive X-ray spectroscopy (EDS) spectra taken at different height along the Pt-coated wire array in (*A*). A homogenous distribution of Pt, In and Sn was observed along the prepared microwire array. Scale bar, 5 μm.

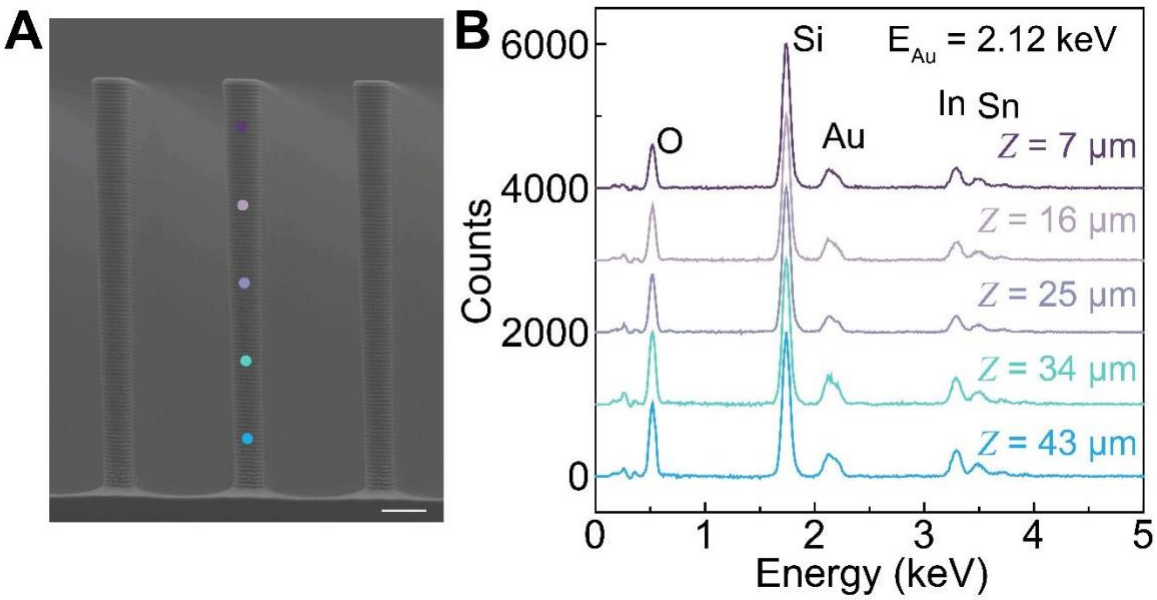


Figure S3. Morphological characterization of Au-coated wire array electrodes with ***k*** = (15, 4, 50) (*A*) The cross-sectional scanning electron microscopy (SEM) image depicting the side view of wire array with ***k*** = (15, 4, 50) (*B*) Energy-dispersive X-ray spectroscopy (EDS) spectra taken at different height along the Pt-coated wire array in (*A*). A homogenous distribution of Au, In and Sn was observed along the prepared microwire array. Scale bar, 5 μm.

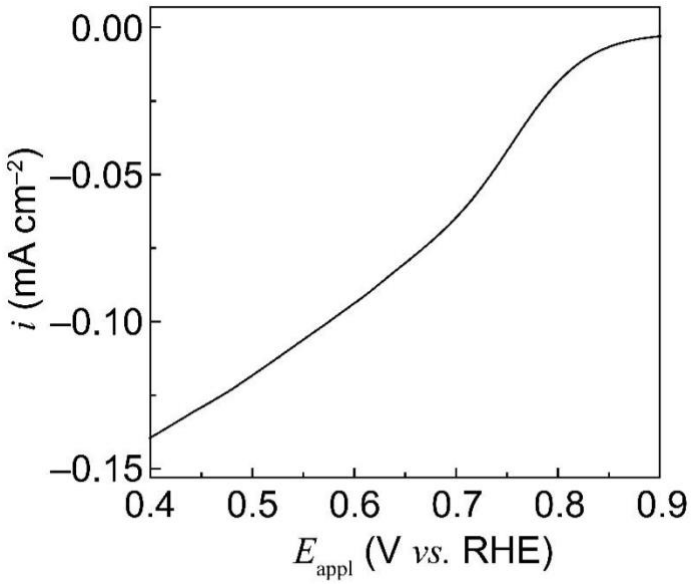


Figure S4. Linear scan voltammetry (LSV) measurement on Pt-coated wire array electrodes from 0.9 V *vs.* RHE to 0.4 V *vs.* RHE. The calculated Tafel slope was 110 mV/decade and the estimated exchange current density was estimated to be 3×10^{-7} mA/cm^2. RHE, reversible hydrogen electrode.

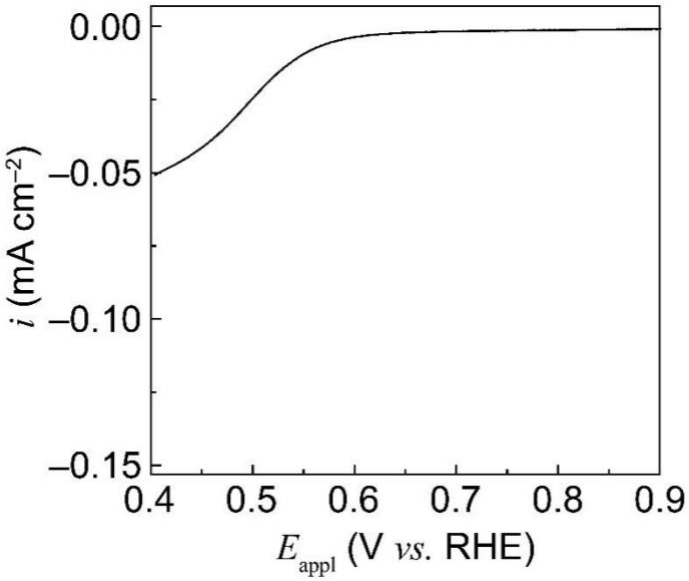


Figure S5. Linear scan voltammetry (LSV) measurement on Au-coated wire array electrodes from 0.9 V *vs.* RHE to 0.4 V *vs.* RHE. The fitted slope (176 mV/decade) supported electron transfer number and exchange current density of two-electron oxygen reduction reaction (ORR) and four-electron ORR on Au-coated wire array electrodes in simulation.

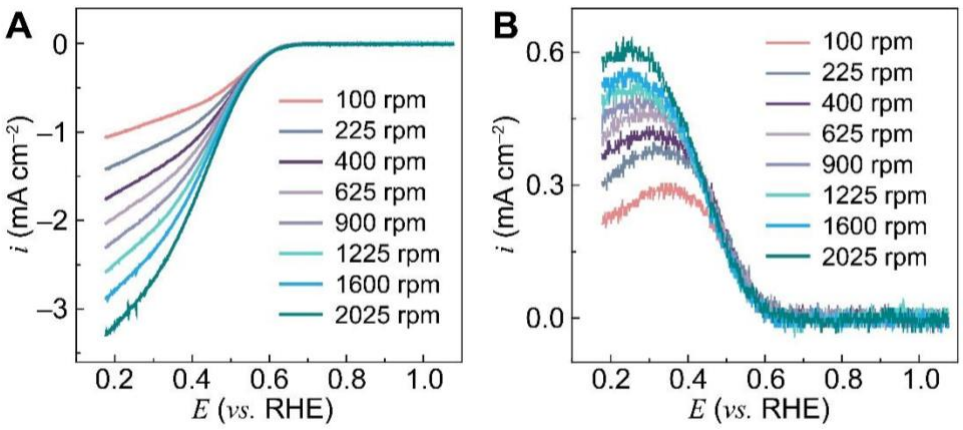


Figure S6. Rotating ring-disk electrode (RRDE) voltammetry on Au-coated working electrodes. (*A*) Voltammetry measurement on an Au-coated disk electrode. (*B*) Voltammetry measurement on a Pt ring electrode. The electrochemical potential applied on the Pt ring electrode was 1.9 V *vs.* RHE during the measurement.

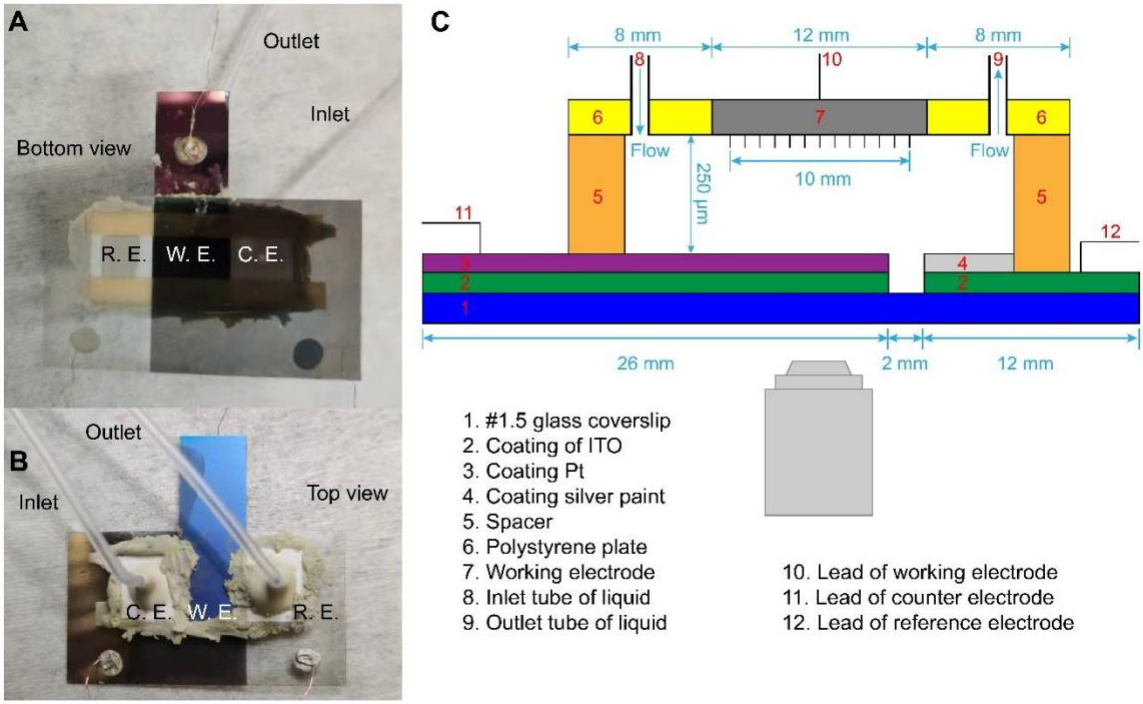


Figure S7. Photo of a fluidic device for a confocal microscope. (*A*) The bottom view of the fluidic device. (*B*) The top view of the fluidic device. (*C*) The device setup for confocal microscope. The objective lens is shown here in a setting of an inverted confocal laser scanning microscope. W. E., working electrode. C. E., counter electrode. R. E., reference electrode.

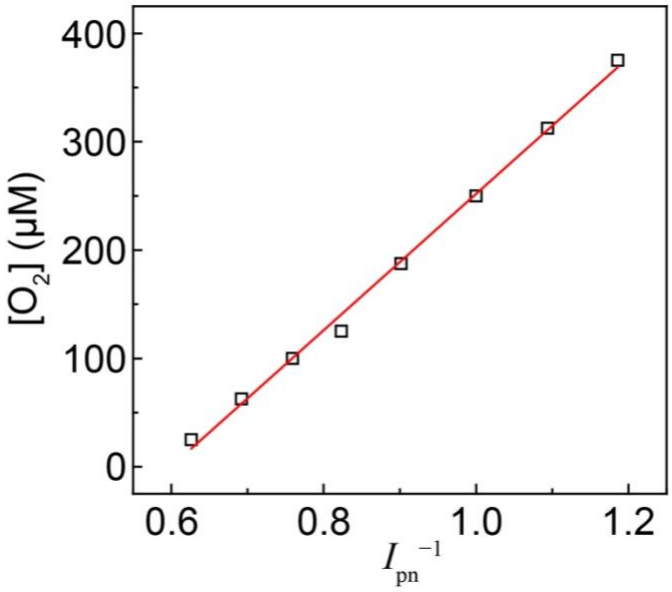


Figure S8. Plot of inverse of the measured normalized phosphorescence intensity I_{pn}^{-1} under various O_2 concentration. The I_{pn} of tris(1,10-phenanthroline)ruthenium(II), $Ru(Phen)_3^{2+}$ (Phen = 1,10-Phenanthroline) from 600 nm to 640nm (see "Materials and Methods" section) was measured under different O_2 concentration, which functioned as the standard curve for $[O_2]$ calibration.

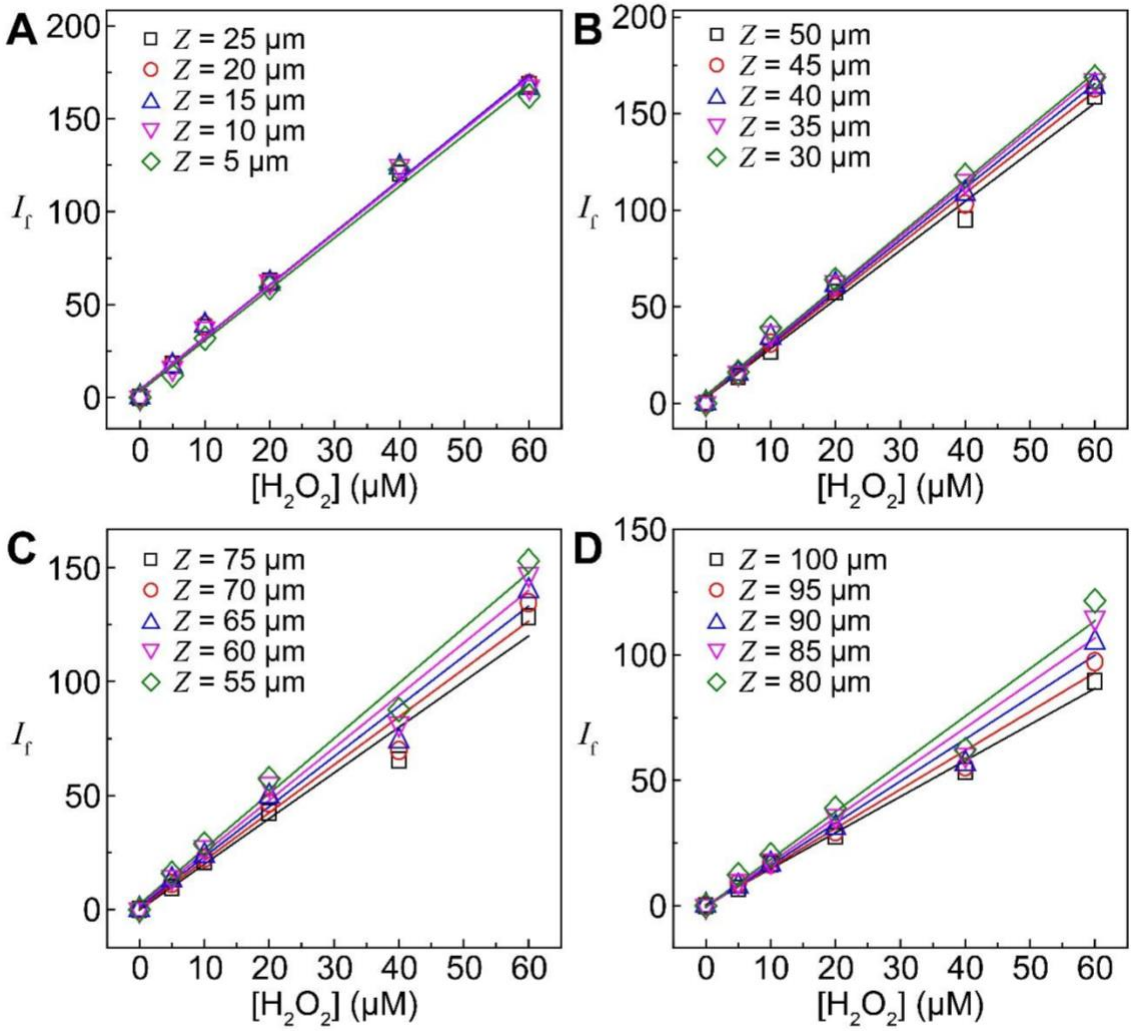


Figure S9. Plot of measured fluorescence intensity (I_f) under various H_2O_2 concentration on Au-coated wire array electrodes with morphology of $\boldsymbol{k}$ = (15, 4, 50). The local fluorescence intensity of resorufin from 590 nm to 650 nm I_f (see "Materials and Methods" section) was measured under different H_2O_2 concentration. Z is the distance between measured region and the base of the wire array. The relationship between I_f and local H_2O_2 concentration ($[H_2O_2]$) was used for $[H_2O_2]$ calibration on Au-coated wire array electrodes. Morphological vector $\boldsymbol{k}$ = (P, D, L) presented the wire arrays' periodicity (P), diameter (D), and length (L), respectively, and unit is μm. (*A*) Calibration curves for region from Z = 5 μm to 25 μm. (*B*) Calibration curves for region from Z = 30 μm to 50 μm. (*C*) Calibration curves for region from Z = 55 μm to 75 μm. (*D*) Calibration curves for region from Z = 80 μm to 100 μm.

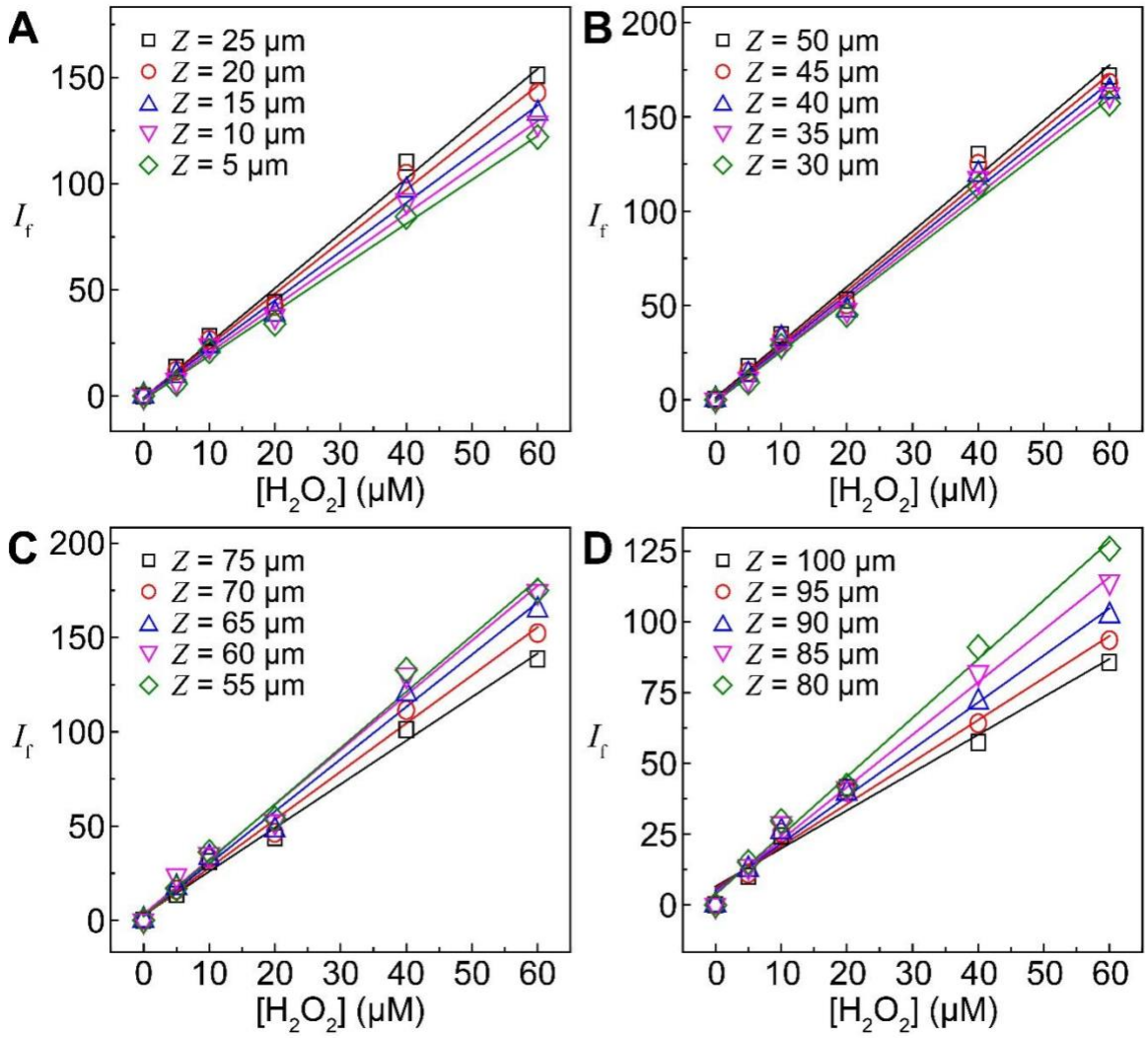


Figure S10. Plot of measured I_f under various H_2O_2 concentration on Au-coated wire array electrodes with morphology of $\boldsymbol{k}$ = (30, 3, 50). The local fluorescence intensity of resorufin from 590 nm to 650 nm I_f (see "Materials and Methods" section) was measured under different H_2O_2 concentration. Z is the distance between measured region and the base of the wire array. The relationship between I_f and $[H_2O_2]$ was used for $[H_2O_2]$ calibration on Au-coated wire array electrodes. (*A*) Calibration curves for region from Z = 5 μm to 25 μm. (*B*) Calibration curves for region from Z = 30 μm to 50 μm. (*C*) Calibration curves for region from Z = 55 μm to 75 μm. (*D*) Calibration curves for region from Z = 80 μm to 100 μm.

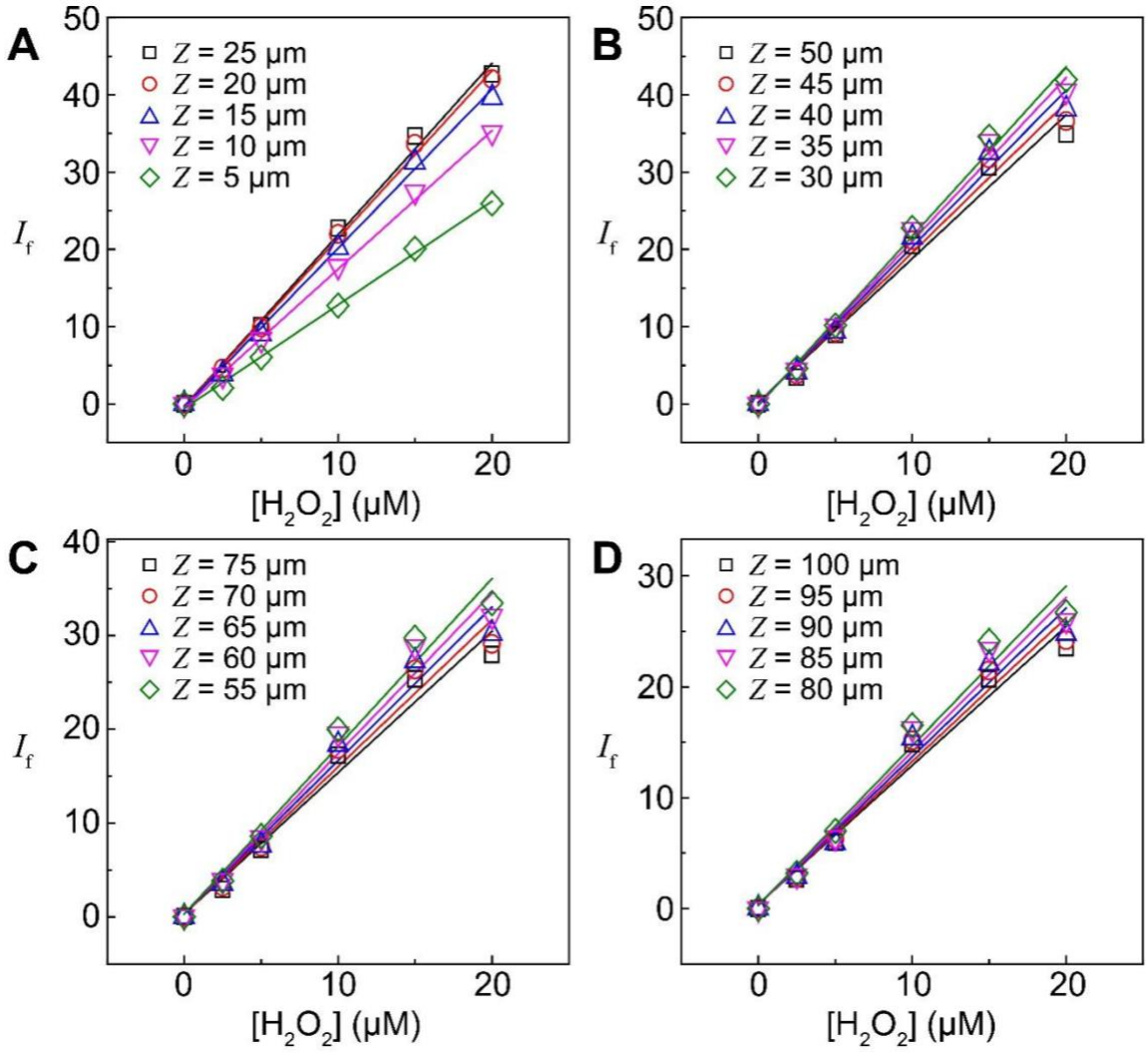


Figure S11. Plot of measured I_f under various H_2O_2 concentration on Au-coated wire array electrodes with morphology of $\boldsymbol{k}$ = (15, 4, 20). The local fluorescence intensity of resorufin from 590 nm to 650 nm I_f (see "Materials and Methods" section) was measured under different H_2O_2 concentration. Z is the distance between measured region and the base of the wire array. The relationship between I_f and $[H_2O_2]$ was used for $[H_2O_2]$ calibration on Au-coated wire array electrodes. (*A*) Calibration curves for region from Z = 5 μm to 25 μm. (*B*) Calibration curves for region from Z = 30 μm to 50 μm. (*C*) Calibration curves for region from Z = 55 μm to 75 μm. (*D*) Calibration curves for region from Z = 80 μm to 100 μm.

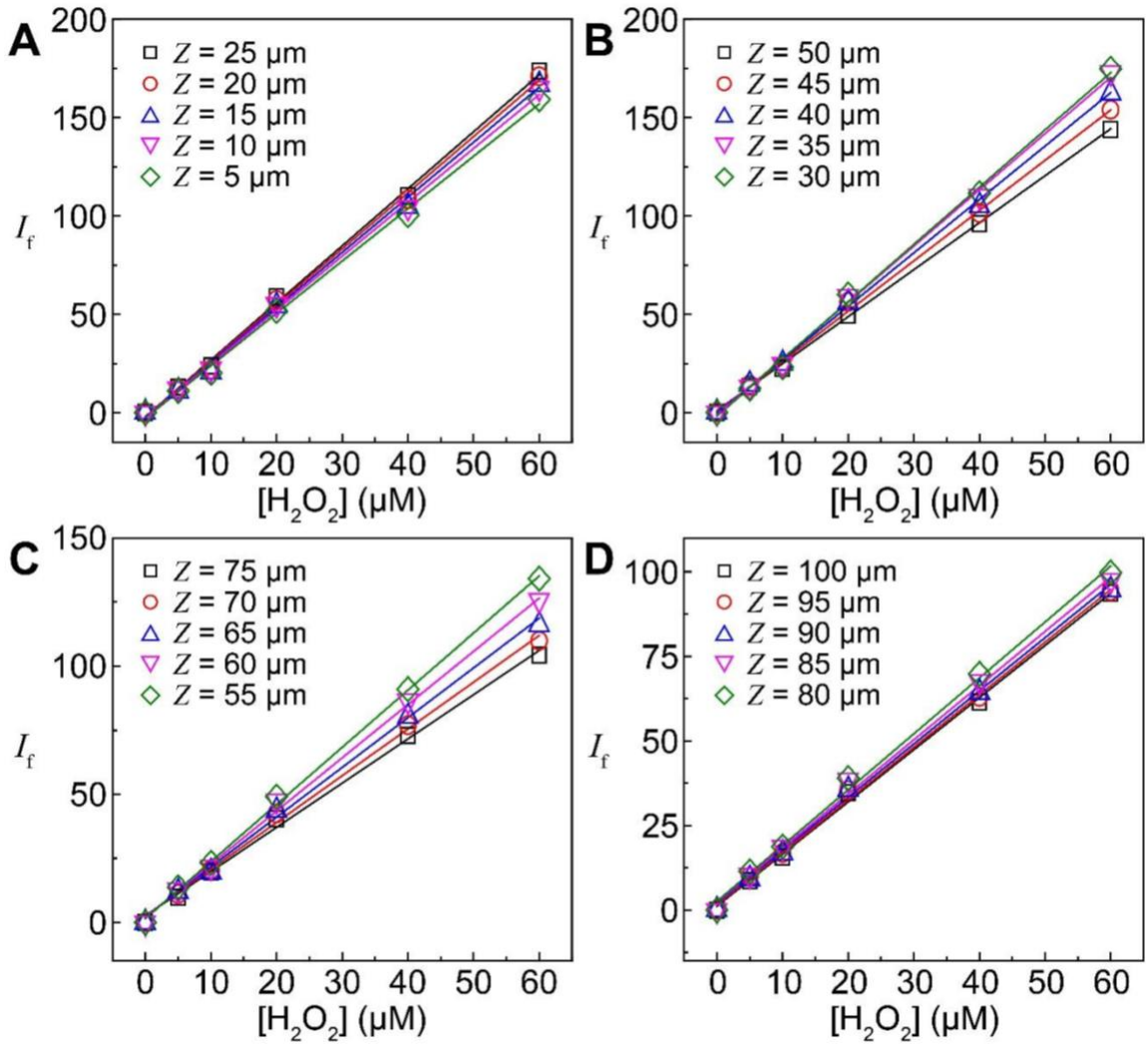


Figure S12. Plot of measured I_f under various H_2O_2 concentration on Au-coated wire array electrodes with morphology of ***k*** = (17, 3, 30). The local fluorescence intensity of resorufin from 590 nm to 650 nm I_f (see "Materials and Methods" section) was measured under different H_2O_2 concentration. *Z* is the distance between measured region and the base of the wire array. The relationship between I_f and [H_2O_2] was used for [H_2O_2] calibration on Au-coated wire array electrodes. (*A*) Calibration curves for region from *Z* = 5 μm to 25 μm. (*B*) Calibration curves for region from *Z* = 30 μm to 50 μm. (*C*) Calibration curves for region from *Z* = 55 μm to 75 μm. (*D*) Calibration curves for region from *Z* = 80 μm to 100 μm.

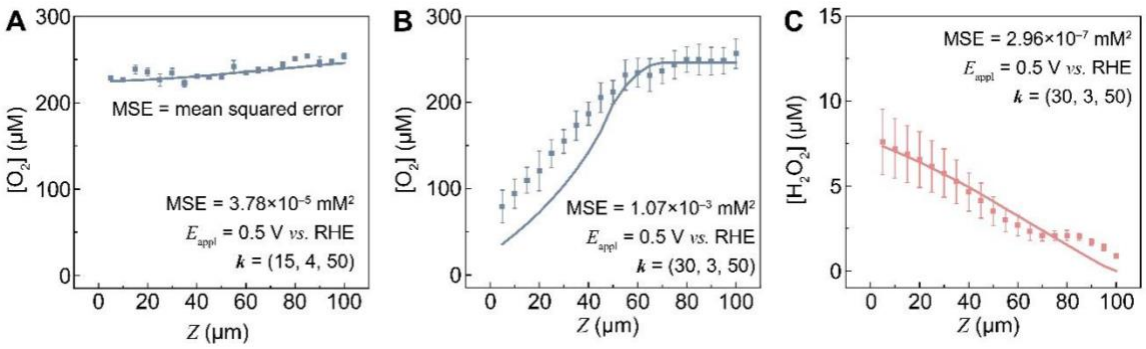


Figure S13. Simulated and experimental O_2 gradient and H_2O_2 gradient on Pt/Au-loaded wire array electrodes under E_{appl} = 0.5 V *vs.* RHE. (*A*) Simulated and experimental O_2 gradient on Au-loaded wire array electrodes with $\boldsymbol{k}$ = (15, 4, 50). (*B*) Simulated and experimental O_2 gradient on Pt-loaded wire array electrodes with $\boldsymbol{k}$ = (30, 3, 50). (*C*) Simulated and experimental H_2O_2 gradient on Au-loaded wire array electrodes with $\boldsymbol{k}$ = (30, 3, 50). All the dots are experimental data and all the lines are corresponding simulated gradients. Although O_2 gradients on Au-loaded wire array electrodes (a) are small perturbations (around 0.2 μM/μm), the simulation still achieved low mean squared error. MSE, mean squared error. Error bars represent standard deviations across multiple separate measurements in the device ($n \geq 3$).

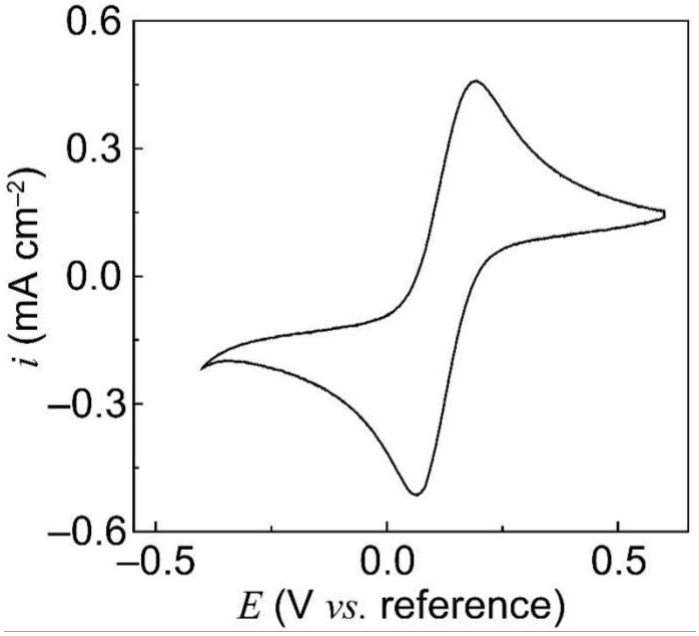


Figure S14. Cyclic voltammetry measurement in ZoBell's solution using Ag-pseudo reference electrode. Besides, a platinum wire and a glassy carbon electrode were used as the counter electrode and the working electrode respectively. The cyclic voltammetry measurement was conducted from 0.6 V *vs.* reference to – 0.4 V *vs.* reference at a scan rate of 20 mV/s.

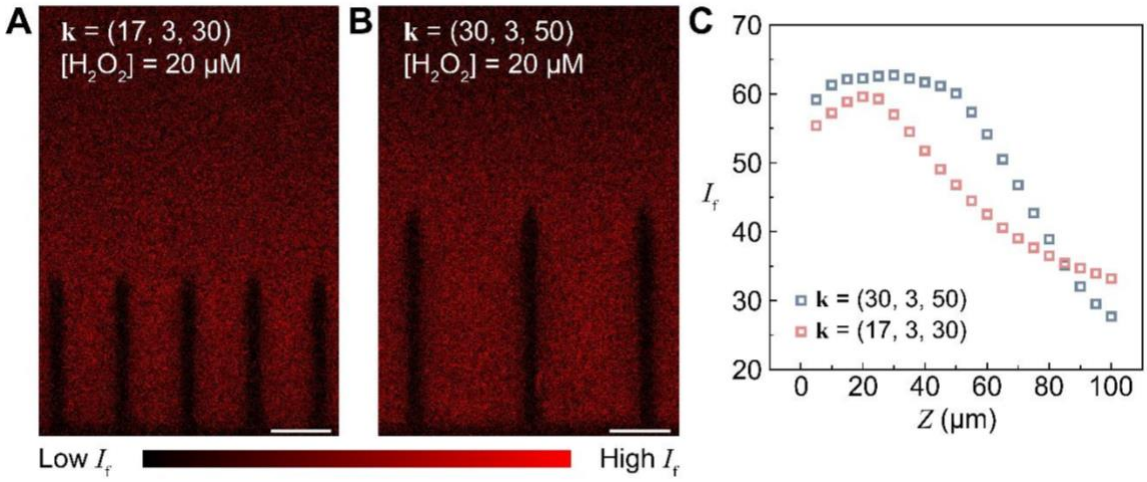


Figure S15. Fluorescence intensity I_f measurement under different morphologies. (*A*) Confocal image with ***k*** = (17, 3, 30) when $[H_2O_2]$ = 20 μM (*B*) Confocal image with ***k*** = (30, 3, 50) when $[H_2O_2]$ = 20 μM (*C*) I_f mapping comparison between 2 morphologies when $[H_2O_2]$ = 20 μM, suggesting I_f is related to both morphology and location. Scale bar, 15 μm.

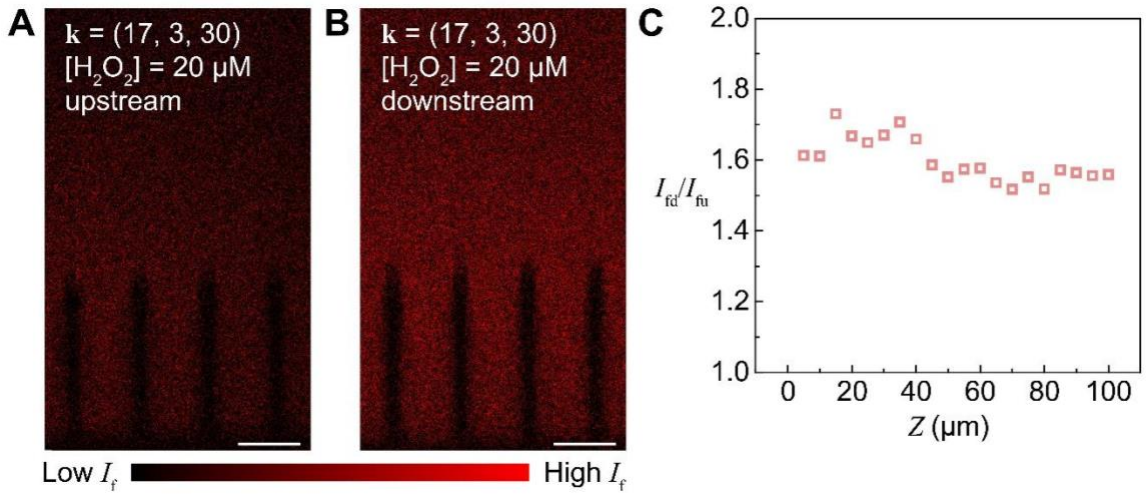


Figure S16. I_f difference between the upstream and downstream of the wire array electrode. (*A*) Confocal image with $\boldsymbol{k}$ = (17, 3, 30) when $[H_2O_2]$ = 20 μM taken at the upstream of the wire array electrode (*B*) Confocal image with $\boldsymbol{k}$ = (17, 3, 30) when $[H_2O_2]$ = 20 μM taken at the downstream of the wire array electrode (*C*) The ratio between fluorescence intensity at the downstream and fluorescence intensity at the upstream (I_{fd}/I_{fu}) at different location. The average ratio is around 1.6. The travel time from upstream to downstream was calculated as 0.84 s, suggesting I_f will increase by 1.9 times each second.

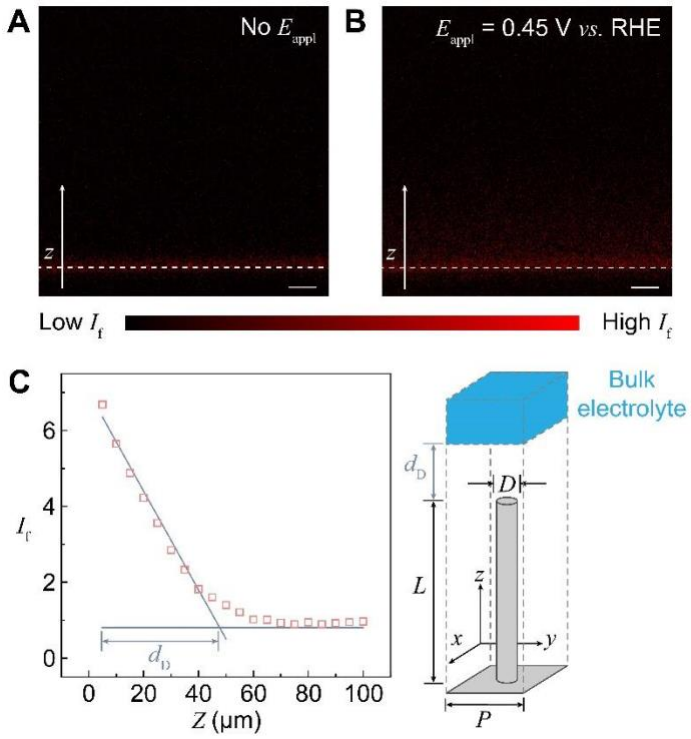


Figure S17. Fluorescence mapping on Au-coated planar electrode. (*A*) The fluorescence intensity ($\boldsymbol{I_f}$) mapping on Au-coated planar electrode without E_{appl} (*B*) The I_f mapping on Au-coated planar electrode when E_{appl} = 0.45 V *vs.* RHE (*C*) Plotting of local I_f against the local distance from the base of the electrode (*Z*) on Au-coated planar electrode with E_{appl} = 0.45 V *vs.* RHE. The diffusion layer thickness, d_D, is determined by the intersection of the two fitted red lines. On Au-coated electrodes, d_D was 50 μm. Scale bar, 15 μm.

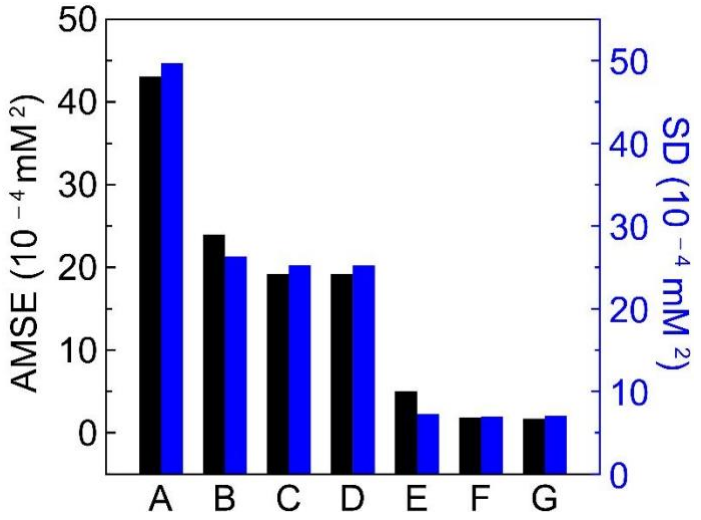


Figure S18. Machine learning model selection for O_2 gradient prediction on Pt-loaded wire array electrodes. The average mean square error (AMSE) and standard deviation (SD) of mean square error (MSE) in various machine learning approaches for O_2 gradient prediction on Pt-loaded wire array electrodes (A is Random Forest Regressor, B is Support Vector Regression, C is Bayesian Ridge Regression, D is Stochastic Gradient Descent, E is Adaboost Regressor, F is K-Nearest Neighbors Regressor, G is Multiple-layer Perceptron Neural Networks).

Table S1: Phosphate-buffered saline (PBS)

Component	Concentration (g/L)
NaCl	8
KCl	0.2
Na_2HPO_4	1.44
KH_2PO_4	0.24

Chapter 4. Applications of bulk generation of simulated electrochemical data for automated mechanism assignment in cyclic voltammetry

This chapter is a version of Hoar, B. B.; Zhang, W.; Xu, S.; Deeba, R.; Costentin, C.; Gu, Q.; Liu, C. "Electrochemical mechanistic analysis from cyclic voltammograms based on deep learning." *ACS Measurement Science Au* **2022**, *2(6)*, 595–604

Abstract

For decades, employing cyclic voltammetry for mechanistic investigation has demanded manual inspection of voltammograms. Here we report a deep-learning-based algorithm that automatically analyzes cyclic voltammograms and designates a probable electrochemical mechanism among five of the most common ones in homogenous molecular electrochemistry. The reported algorithm will aid researchers' mechanistic analyses, utilize otherwise elusive features in voltammograms, and experimentally observe the gradual mechanism transitions encountered in electrochemistry. An automated voltammogram analysis will aid the analysis of complex electrochemical systems and promise autonomous high-throughput research in electrochemistry with minimal human interference.

Introduction

Cyclic voltammetry is one of the most common electrochemical characterization techniques and it generates valuable mechanistic information for redox-active chemical systems.[1-3] For decades, cyclic voltammetry has been indispensable for electrochemical applications in sensing, energy-storage, chemical transformations, and beyond; however, the general protocol of initial mechanistic analysis after experiments has remained largely unchanged since its

inception.[3,4] Researchers manually inspect the shapes and variations of cyclic voltammograms under multiple different scan rates (v), sometimes with different reactant concentrations, and subsequently hypothesize a qualitative mechanism consisting of interfacial charge transfers (E step) and/or solution reactions (C steps),[1,2] before sometimes extracting quantitative kinetic information via additional experiments and/or numerical simulations.[5,6] However, such manual inspection demands extensive researcher training, potentially incurs human bias, and is not compatible with automated testing needed for high-throughput screening. An algorithm that automatically analyzes cyclic voltammograms and qualitatively categorizes electrochemical systems into mechanisms with a specific combination of E and/or C steps will help alleviate the aforementioned challenges in manual analysis of cyclic voltammograms.

We envision that machine-learning algorithms such as those of deep-learning (DL) are capable of aiding mechanism categorization in cyclic voltammetry. In electrochemistry, the kinetics of E and/or C steps formulate the set of partial differential equations (PDE) and boundary conditions that dictate the i-E characteristics recorded in the cyclic voltammograms under a collection of different v values ($\{v, i(E)\}_n$, n, number of different v values) (Fig. 1a).[1,2,7] Such a mathematically bijective function between electrochemical mechanisms ($\{E_i, C_j\}$) and the electrochemically accessible parameter space of the combined voltammograms $\{v, i(E)\}_n$ suggests that it is feasible to employ DL algorithms to designate discrete mechanisms from sufficiently sampled cyclic voltammograms with minimal ambiguity (Fig. 1a). Indeed, this bijective relationship enables numerical simulations based on finite-element methods,[7,8] and recently by artificial neural networks,[9-11] to be used as a tool to efficiently search the parameter space of cyclic voltammograms and fit kinetic parameters upon a mechanism determined *a priori* by manual inspections of voltammogram. Yet, electrochemical mechanistic investigations remain commonly

trial-and-error because the determination of the aforementioned *a priori* mechanism before any quantitative studies still relies on manual inspection.

We posit that a set of cyclic voltammograms simulated from finite-element methods based on pre-set mechanism designations will suffice in the first-order approximation for the establishment of a DL model that analyzes cyclic voltammograms and qualitatively categorizes mechanisms provided a large enough sampling of $\{v, i(E)\}_n$. Research by Bond and coworkers tested the concept of DL-based automatic analysis of a single simulated voltammogram for a selection of 3 mechanism types with overall accuracies just below 90%.[12,13] The exploration of this concept with experimental data, which commonly includes multiple voltammograms at different v values, has yet to be conducted for mechanistic studies. We advocate the broader use of DL-based analysis and hypothesize that the bijective relationship between $\{E_i, C_j\}$ and $\{v, i(E)\}_n$ enables the establishment of DL algorithms that detect and utilize subtle voltammogram features, ones not commonly used as mechanistic discriminants by humans, and observe the evolution of "edge" cases when two mechanisms co-exist and/or one mechanism is transitioning into another one, challenging scenarios for manual analysis in cyclic voltammetry. The established algorithms can be continuously refined and improved from experimental data, potentially including data contributed from the electrochemistry community. The algorithms will find their use in analyzing complex mechanistic scenarios and addressing the current paucity of automatic, high-throughput mechanistic analysis in electrochemistry.

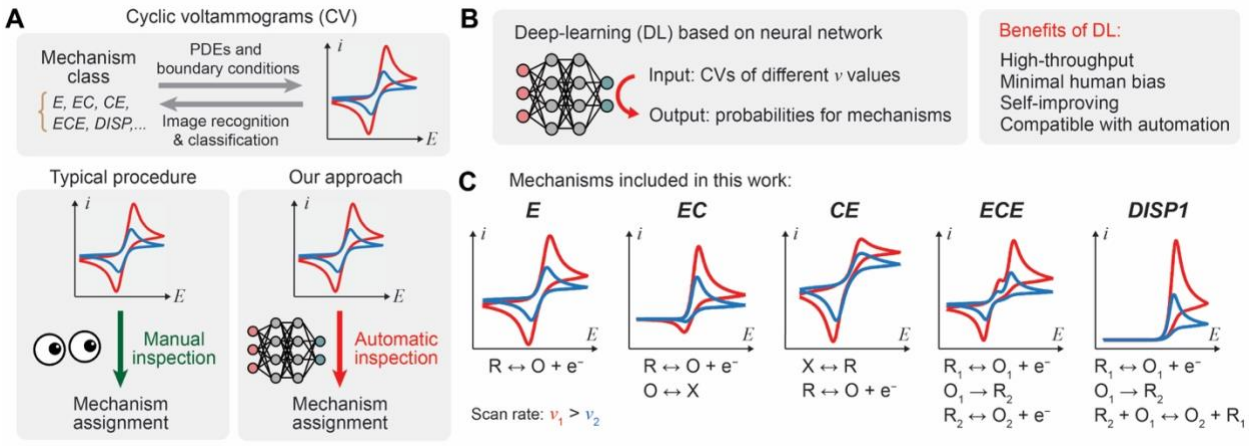


Figure 1. A deep-learning (DL) algorithm of automatically analyzes for cyclic voltammetry. **A**, The bijective relationship between electrochemical mechanism and cyclic voltammograms, and the comparison between manual inspection of voltammograms and our approach. **B**, The function of DL algorithm and its proposed benefits. **C**, The five common molecular electrochemical mechanisms included in the DL algorithm. PDEs, partial differential equations.

In this work, we demonstrate a DL algorithm for automatic mechanistic analysis for cyclic voltammetry (Fig. 1b). State-of-art DL algorithms using residual neural networks (ResNet) architecture[14] are established to analyze cyclic voltammograms at different scan rates $\{v, i(E)\}$ and yield the electrochemical system's probability towards five of the most common stoichiometric homogenous mechanisms in electrochemistry textbooks (Fig. 1c):[1,2,15,16] a single electron transfer with any level of reversibility (*E*), an *E* step followed by a *C* step with any level of reversibility (*EC*), an *E* step preceded by a *C* step (*CE*), a system of two *E* steps connected by an irreversible rate-limiting C_i step with the second *E* step being more thermodynamically facile than the first one (*ECE*), and a two-electron transfer that is similar to *ECE* yet the second *E* step is replaced by a solution disproportionation reaction (*DISP1*). We demonstrate DL's capability of accurately designating mechanisms in simulated and experimental scenarios, unveiling potential new features

in the voltammograms elusive to manual inspection, as well as semi-quantitatively observing the gradual transitions of electrochemical mechanisms. The developed algorithm will be applicable to analyze complex electrochemical systems when competing mechanisms are intertwined together. In conjunction with robotic experimentation,[17,18] the demonstration of automatic mechanistic analysis in cyclic voltammetry presents the possibility of automated high-throughput research to investigate mechanisms in electrochemical systems with minimal human intervention.

Results

The data of cyclic voltammograms were sanitized and transformed into two-dimensional matrices suitable for DL algorithms of ResNet architecture. While cyclic voltammograms are typically presented as images in literature, much of the white space in voltammograms contains little (if any) information. Hence, similar to the case of electrocardiogram,[19] a two-dimensional matrix of $\{v, i(E)\}_n$, rather than the images of cyclic voltammograms as in the works of Bond and coworkers,[12,13] is employed to store electrochemical information for the DL-based analysis (Fig. 2a). For the ease of training a ResNet, in each set of $\{v, i(E)\}_n$ the current densities i in voltammograms were normalized as $i_{\text{normalized}}$ against the largest i among all voltammograms in $\{v, i(E)\}_n$, with $i_{\text{normalized}}$ in the forward scan designated as positive values. The electrochemical potentials E were adjusted so that the position of 0 V of the adjusted electrochemical potential (E_{adjusted}) roughly corresponds to the potential of studied redox couple (see Methods). Such data processing ensures a generally readable format of cyclic voltammograms despite the large variations in experimental testing conditions.

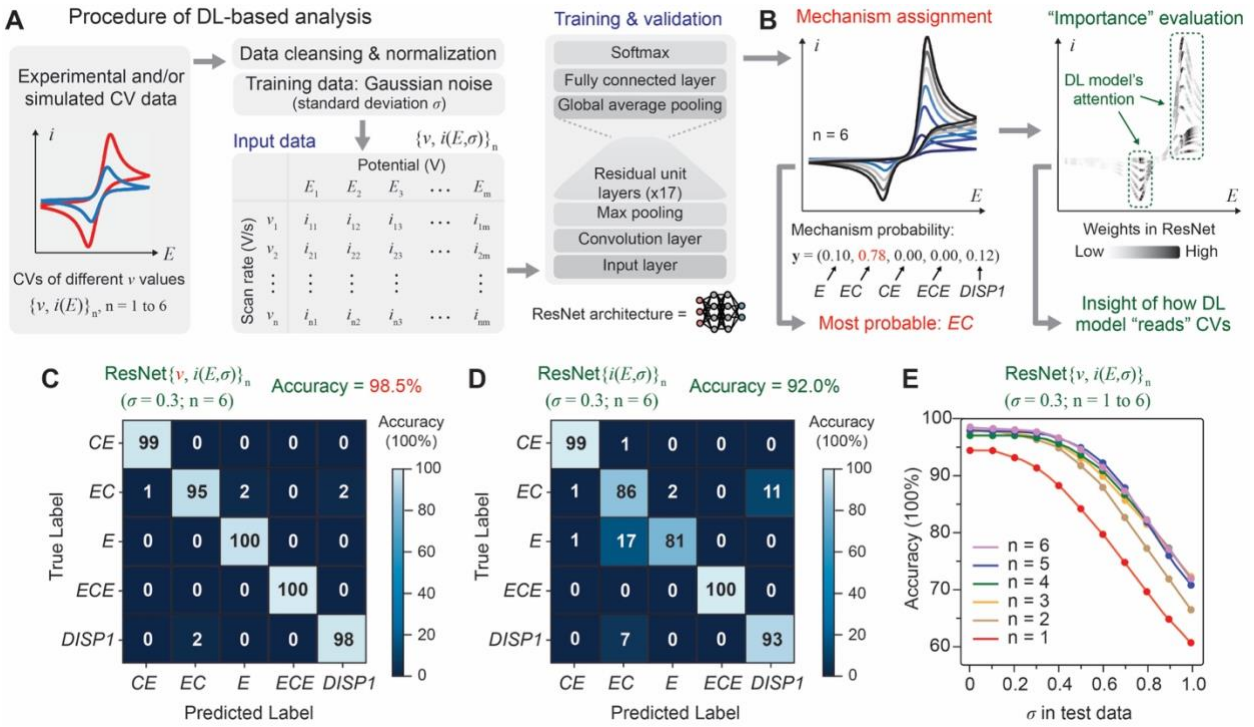


Figure 2. The established DL algorithm of ResNet architecture for cyclic voltammetry. **A**, The structure of input data and the convolutional neural networks of residual neural network (ResNet) architecture. **B**, The output and utility of DL model for mechanism designation. **C** and **D**, The confusion matrix of DL model trained by simulated cyclic voltammograms with (**C**) and without (**D**) explicit values of scan rate (v) as input data. v, scan rate; σ, the standard deviation of the Gaussian noise; n, the number of v values. **e**, the accuracies of the DL model in **c** when tested with simulated voltammograms with varying values of n and σ.

Cyclic voltammograms based on the targeted mechanisms (Fig. 1c) were numerically simulated as the training set for deep neural networks via the finite-element method (see Methods). Numerical models of PDEs, boundary conditions, and initial conditions were constructed based on mechanisms' definitions in textbooks (see Supplementary Information).[1,2] Moreover, the numerical models' parameters, including but not limited to, the numbers and values of scan rate v (n = 1 to 6), electrodes' double layer capacitance reported in literature,[20] standard rate constant of

interfacial charge transfer in the concentration-dependent Butler-Volmer equation following Nicholson's formalism ($\psi \in [10, 0.3]$) in the *E* step,[3] and the equilibrium constants and forward/backward rate constants in the *C* step based on Savéant's definitions,[2] were incorporated into the simulations and carefully constrained with practical and fundamental considerations. The parameters were randomly sampled (see Table S1 and Supplementary Information) for a comprehensive yet even exploration of the mechanism's corresponding domain in the $\{v, i(E)\}_n$ space, i.e. the corresponding kinetic zone diagrams.[2] As the experimental voltammograms typically contain Gaussian-type noise due to background and instrumentation,[21] Gaussian noise of varying degrees of standard deviation σ relative to the maximal current densities were added to the simulated voltammograms, resulting in the training set $\{v, i(E, \sigma)\}_n$ (n = 1 to 6) (examples in Fig. S1). The addition of Gaussian noise not only better reflects the realistic electrochemical data but also increases the algorithm's tolerance towards noises in automatic mechanism categorization[14] (*vide infra*).

We chose ResNet[14], a widely used network architecture evolved from convolutional neural networks,[22,23] to extract intrinsic features from high-dimensional data. The ResNet architecture utilizes skip connections within its convolutional layers for deeper networks for greater feature extraction. Such architecture is critical for training successes as it alleviates the problems of both vanishing and exploding gradients during the training process which mitigates the risk of training failure while maintaining the low overfitting risk inherent of convolutional architectures.[14] The neural network is trained to take either the first or the second cycle of voltammograms for the same electrochemical system at various numbers of different scan rates ($\{v, i(E, \sigma)\}_n$, n = 1 to 6) and yields the vector $\mathbf{y} = \{y_1, y_2, y_3, y_4, y_5\}$ (Fig. 2b, see Methods), in which each component in **y** is a surrogate of the electrochemical system's probability or fraction towards mechanisms of *E*, *EC*,

CE, *ECE*, and *DISP1*, respectively. The classification process is completed by designating the mechanism of the largest component in **y** as the most probable or most prominent one for the studied electrochemical system. Furthermore, the proposed model can estimate the "importance" of different parts of the voltammograms to the prediction (Fig. 2b), by visualizing the relative magnitudes of gradients of the logits on input data after feeding through the neural network. We hypothesize that such a visual guide of algorithm's "importance" will illustrate how DL models analyze cyclic voltammograms and offer a comparative study between manual inspection and the ResNet-based one.

State-of-art ResNet models of 18 residual learning layers, i.e. ResNet-18, were trained and validated by simulated cyclic voltammograms $\{v, i(E, \sigma)\}_n$. Here the use of a minimal ResNet-18 model is commensurate with our study when smaller neural networks are desired.[14] When $n = 6$ and $\sigma = 0.3$, satisfactory accuracy (>90%) was achieved among $\{v, i(E, \sigma)\}_n$ when more than 3,000 electrochemical systems were included for each mechanism type in the training set (Fig. S2a). After 1000 epochs of training to improve accuracy (Fig. S2b), a voting process that contains eight ResNet-18 models, designated as ResNet$\{v, i(E, \sigma)\}_n$, achieved an overall accuracy of 98.5% for $\{v, i(E, \sigma)\}_n$ ($n = 6$, $\sigma = 0.3$) and generated a confusion matrix with nearly zero off-diagonal components (Fig. 2c). In comparison, alternative machine-learning algorithms[24] including linear classification, the vanilla multilayer perceptron (MLP), the MLP using attention mechanism to aggregate the extracted features of each curve ("MLP & attention mechanism"), and the MLP sharing same parameters/weights on the first layer ("MLP & parameter sharing") only yielded lower accuracies of 88.7%, 89.6%, 92.3%, and 90.7%, respectively (Fig. S2c to S2f). We also built a DL model under the same protocol without the v values as input, ResNet$\{i(E, \sigma)\}_n$ ($n = 6$, $\sigma = 0.3$), in which the model's inputs contain the n number of voltammograms but not the exact v

values. Only an overall accuracy of 92.0% was achieved and the corresponding confusion matrix contains noticeable non-zero off-diagonal entries (Fig. 2d). Consistent with manual inspection, the exact values of v are critical in DL to fully differentiate electrochemical mechanisms.

The established DL model is remarkably resilient to appreciable degrees of noises in the simulated cyclic voltammograms. The prediction accuracy of the DL model trained by $\{v, i(E, \sigma)\}_n$ ($n = 6$, $\sigma = 0.3$) was tested by simulated cyclic voltammograms ($n = 6$) with varying values of σ ranging from 0.0 to 1.0. The overall accuracy remains mostly constant and higher than 95% until $\sigma \geq 0.5$ (purple trace in Fig. 2e). Even at $\sigma = 1.0$ when the simulated voltammograms are barely recognizable by manual inspection (Fig. S1), an overall accuracy of more than 70% was achieved. Such a tolerance towards noises in cyclic voltammograms is remarkable in comparison to the DL models trained when $\sigma = 0.0$ (no noise) and 0.1, since in the latter two models (trained with $\sigma = 0.0$ and 0.1) the overall accuracies gradually drop below 40% and 80%, respectively, at $\sigma = 0.5$ of the testing data (Fig. S2g and S2h). Such a gradual decline in accuracy with added noise is indicative of a robust and well fit model, since an overfit model would be expected to perform poorly when added noise in the testing set deviates from the added noise level that it was trained on. The addition of Gaussian noise in model training increases the robustness and sensitivity of the established DL model against data noise that may not be tolerable by manual analysis.[14]

We also evaluated how the value of n, i.e., the number of cyclic voltammograms, affects the overall accuracy of DL models of ResNet architecture. The accuracies of DL models trained by $\{v, i(E, \sigma)\}_n$ (n = 1 to 6, $\sigma = 0.3$) were tested by simulated voltammograms (n = 1 to 6, respectively) of σ values ranging from 0.0 to 1.0 (Fig. 2e). Interestingly, when $\sigma < 0.5$ in the testing data, there is no distinguishable differences in overall accuracies when n = 2 to 6 while the accuracies are noticeably lower when n = 1 (Fig. 2e). Consistent with the diagnostic value of

voltammograms' evolution across different zones in the kinetic zone diagram,[2] cyclic voltammograms of at least two different *v* values are sufficient for the DL algorithm to accurately designate reaction mechanism within the parameter space defined in the training set.

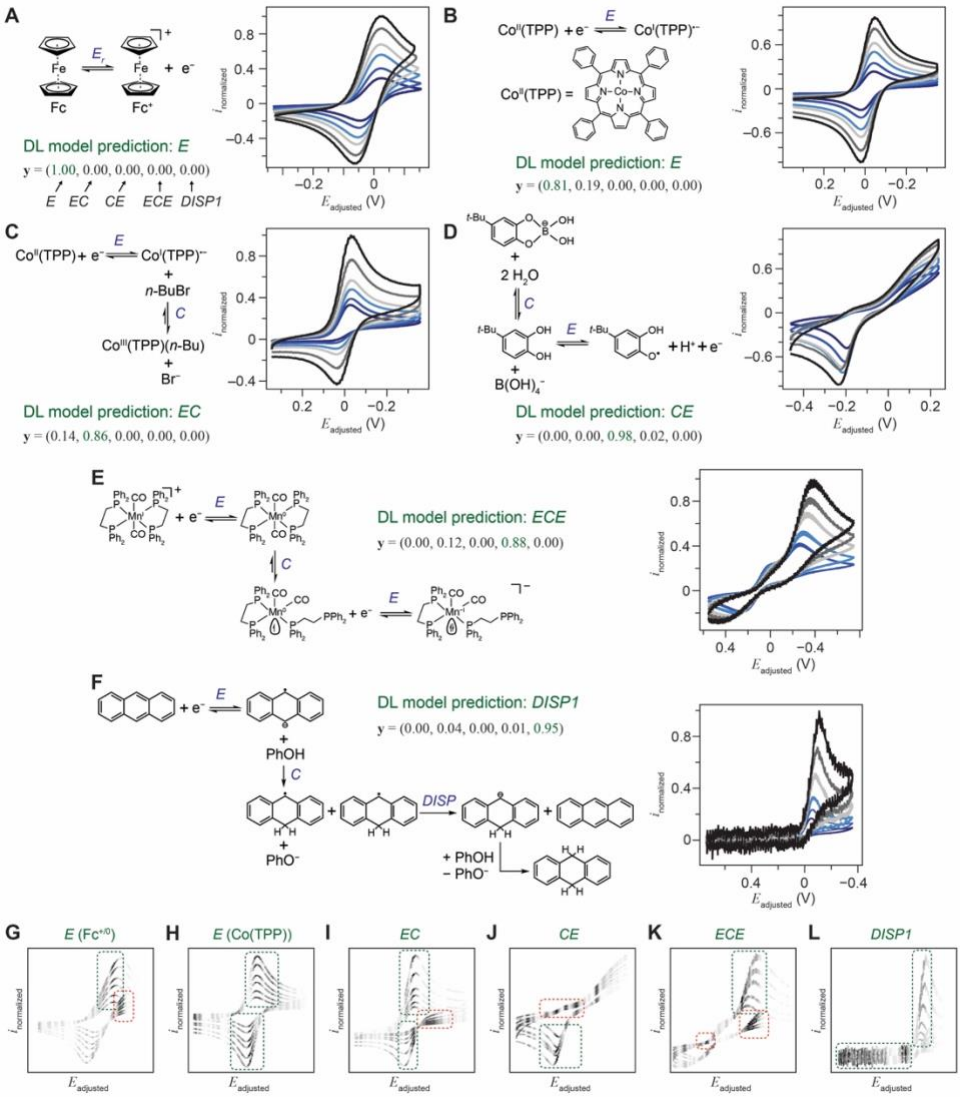


Figure 3. Application of DL model to experimental scenarios. The mechanisms, cyclic voltammograms, **y** vectors from DL model's predictions, and the "importance" plots for 1 mM ferrocene/ferrocenium redox couple (**A, G**), 1 mM tetraphenylporphyrin cobalt(II) (Co^{II}(TPP)) (**B, H**), 1 mM Co^{II}(TPP) and 50 mM 1-bromobutane (*n*-BuBr) (**C, I**), 1 mM 4-*tert*-butylcatechol in 10 mM pH 9.2 boric acid buffer (**D, J**), 5 mM *trans*-$Mn(CO)_2(\eta^2\text{-DPPE})_2^+$ (DPPE, 1,2-bis(diphenylphosphino)ethane) (**E, K**), and 1 mM anthracene and

0.1 M phenol (**F, L**). $i_{\text{normalized}}$, normalized current density with the forward scan in the positive direction. E_{adjusted}, electrochemical potential shifted to center the redox features. The "importance" towards the DL model in expected (green) and somewhat unexpected (red) parts in the voltammograms are highlighted. 3 mm glassy carbon; Ag^+/Ag reference except **D** (Ag/AgCl, 3M KCl); Pt wire counter electrode. DMF in Ar except **D** (water) and **E** (THF). 0.1 M *n*-Bu_4NClO_4 in **A** and **E**, 0.1 M *n*-Bu_4NPF_6 in **B**, **C**, and **E**, and 90 mM KCl in **D**. The voltammograms of the second cycles, notwithstanding **f** (first cycle), are displayed and analyzed. *iR* corrected. v = 0.05, 0.1, 0.2, 0.3, 0.5, 0.7 V/s in **A**, **B**, **C**, and **D**; v = 0.1, 0.2, 0.3, 0.5, 0.7 V/s in **E**; v = 0.05, 0.1, 0.2, 0.5, 1, 2 V/s in **F**. Darker traces in **G** to **L** indicates higher "importance" in the DL model.

Encouraged by the ResNet-based model's accuracy of simulated cyclic voltammograms, we applied the established DL models trained by $\{v, i(E, \sigma)\}_n$ (n = 1 to 6, σ = 0.3) to exemplary experimental scenarios mostly based on the voltammograms of the second cycle. ResNet-based DL model is capable of accurately predicting the *E* mechanism in the ferrocene/ferrocenium (Fc/Fc^+, 1 mM) redox couple in dimethylformamide (DMF) [25] (Fig. 3a) and the single-electron Cobalt (II/I) redox couple with tetraphenylporphyrin cobalt(II) (Co^{II}(TPP), 1 mM) as the starting compound in the absence of any electrophiles in DMF [26-28] (Fig. 3b). The DL model accurately recognizes an *EC* mechanism, or more precisely the E_rC_i variant when both the thermodynamic and kinetic propensity of the forward *C* step is large enough to be considered an irreversible C_i step, where the addition of 1-bromobutane (*n*-BuBr, 50 mM) as an electrophile to 1 mM Co^{II}(TPP) in DMF leads to the formation of tetraphenylporphyrin cobalt(III) *n*-butyl (Co^{III}(TPP)(*n*-Bu)) (Fig. 3c).[28,29] The model also accurately recognizes a *CE* mechanism for the first single-electron oxidation of 4-*tert*-butylcatechol (1 mM) in the aqueous buffer of boric acid (pH = 9.2, 10 mM), where a reversible *C* step is needed to dissociate the thermodynamically favored catechol-borate

adduct into the electrochemically accessible catechol[30,31] (Fig. 3d). The **y** vector output from the DL model designates *ECE* mechanism as the most probable one in the case of 5 mM *trans*-$Mn(CO)_2(\eta^2\text{-DPPE})_2^+$ (DPPE, 1,2-bis(diphenylphosphino)ethane) in tetrahydrofuran (THF). This designation is consistent with prior determinations of a *ECE* process with minimal *DISP* contribution,[32,33] where an intramolecular ligand rearrangement exists between the two single-electron reductions in which the second reduction is thermodynamically more favored than the first one (Fig. 3e). Last, from the first-cycle voltammograms, the DL model accurately designates the *DISP1* mechanism in the net two-electron reduction of 1 mM anthracene in DMF with the presence of 0.1 M phenol as proton donor, where the protonation of the anion radical (*C* step) after the first one-electron reductive *E* step is the rate-determining step followed by the disproportionation reaction (*DISP* step)[15,34] (Fig. 3f). The successful mechanism designation by the DL models for model experimental systems suggests the practicality of utilizing DL for automatic analysis in cyclic voltammetry.

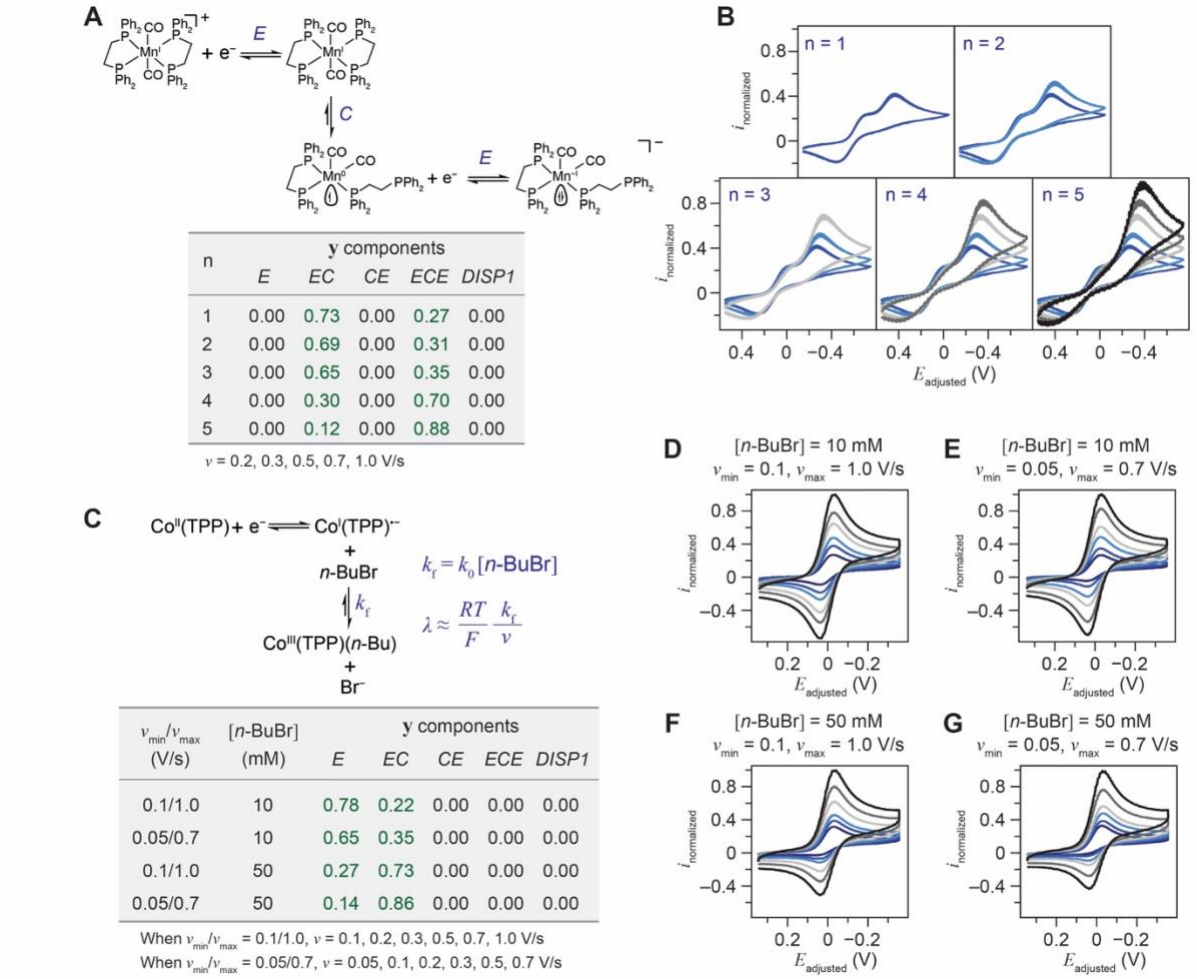
Discussion

There are appreciable similarities between the analytic processes of the established DL algorithm and human inspection. The accuracy decrease from 98.5% to 92.0% when the *v* values were not included as input in the DL model (Fig. 2c and 2d) is consistent with manual inspection, when more definitive mechanism assignment is feasible when the explicit *v* values are included in the analysis.[1-3] Indeed, the exclusion of *v* values as model input deteriorates the model's accuracies mostly by misassigning *E* as *EC* and misassigning *EC* as *DISP1* (increasing from 0% chance to 17% and from 2% to 11%, respectively, from Fig. 2c to Fig. 2d). These misassignments are common when *v* information is missing in manual analysis, owing to the gradual transition of cyclic voltammograms between *E* and *EC* in the kinetic zone diagrams as well as the similarity

between the one-electron EC_i and two-electron *DISP1* processes.[2] The ResNet-based model's dependence of prediction accuracies on n (Fig. 2e) is understandable yet informative. It is common in qualitative mechanistic studies to obtain cyclic voltammograms under multiple v values (i.e. $n > 1$) and compare the voltammograms' evolution.[4] Yet in practice the number of v values needed for mechanism determination seems ill-defined. What the DL model suggests is that statistically in most scenarios two cyclic voltammograms of different v values will suffice and there are diminishing returns of prediction accuracy when $n \geq 2$ within the parameter space defined in our training set of simulated voltammograms. When $n = 2$, we derived the corresponding mathematical requirements for the two v values to be satisfied in the training data set hence empirically offer good accuracy of mechanistic prediction from our DL model (See in Supplementary Information). Such mathematical relationships will be helpful for researchers when deciding experimental parameters in cyclic voltammetry.

Plotting the algorithm's "importance" towards parts of cyclic voltammograms suggests that there is additional information in subtle voltammogram features that may elude manual inspection. Fig. 3g to 3l plot the "importance" distributions in cyclic voltammograms from the DL model shown in Fig. 3a to 3f, respectively. Additional "importance" plots of exemplary simulated voltammograms are available in Fig. S3. While an understandable amount of the model's attention is attributed to presence or absence of primary redox peaks in voltammograms (green areas in Fig. 3g to 3l and Fig. S3), one noticeable feature unique to DL algorithm emerges. Appreciable amount of "importance" of the DL model is frequently assigned to the reverse scan roughly beneath the redox peaks (red areas in Fig. 3g to 3l and Fig. S3), an area typically not carefully examined in manual inspection. We propose that such seldomly examined regions in cyclic voltammograms contains useful mechanistic information and ought to be better utilized for mechanistic studies,

probably by DL-based automatic analysis thanks to the algorithm's sensitivity. Additional investigations will be conducted to examine such an argument with a more rigorous and systematic analysis.



n	y components				
	E	*EC*	*CE*	*ECE*	*DISP1*
1	0.00	0.73	0.00	0.27	0.00
2	0.00	0.69	0.00	0.31	0.00
3	0.00	0.65	0.00	0.35	0.00
4	0.00	0.30	0.00	0.70	0.00
5	0.00	0.12	0.00	0.88	0.00

v = 0.2, 0.3, 0.5, 0.7, 1.0 V/s

v_{min}/v_{max} (V/s)	[*n*-BuBr] (mM)	y components				
		E	*EC*	*CE*	*ECE*	*DISP1*
0.1/1.0	10	0.78	0.22	0.00	0.00	0.00
0.05/0.7	10	0.65	0.35	0.00	0.00	0.00
0.1/1.0	50	0.27	0.73	0.00	0.00	0.00
0.05/0.7	50	0.14	0.86	0.00	0.00	0.00

When v_{min}/v_{max} = 0.1/1.0, v = 0.1, 0.2, 0.3, 0.5, 0.7, 1.0 V/s
When v_{min}/v_{max} = 0.05/0.7, v = 0.05, 0.1, 0.2, 0.3, 0.5, 0.7 V/s

Figure 4. Semi-qualitative analysis of cyclic voltammograms with DL algorithm. **A**, the mechanism and **y** vector components for the *ECE* scenario of *trans*-$Mn(CO)_2(\eta^2\text{-DPPE})_2^+$. **B**, Cyclic voltammograms of 5 mM *trans*-$Mn(CO)_2(\eta^2\text{-DPPE})_2^+$ for different *n* values under the same conditions in Fig. 3f. **C**, The mechanism, kinetic factors, and **y** vector components for the *EC* scenario of $Co^{II}(TPP)$ and *n*-BuBr. **D** to **G**, Cyclic voltammograms of 1 mM $Co^{II}(TPP)$ under different *v* ranges and *n*-BuBr concentrations ([*n*-BuBr]) under the same conditions in Fig. 3c. The voltammograms of the second cycles are displayed and analyzed.

In the predominantly *ECE* case of *trans*-$Mn(CO)_2(\eta^2\text{-DPPE})_2^+$ in THF,[32,33] the evolution of **y** vector as *n* increases illustrates how more definitive mechanism determination in edge cases will benefit from cyclic voltammograms of multiple *v* values (Fig. 4a and 4b). With additional voltammograms, the E_rC_r component in the yielded **y** vector decreases from 0.73 ($n = 1$) to 0.12 ($n = 5$) while the *ECE* one increases from 0.27 ($n = 1$) to 0.88 ($n = 5$) (Fig. 4b). Such changes in **y** components as a function of *n* is consistent with the reported challenges in differentiating E_rC_r and *ECE*/*DISP1* mechanisms with a small number of *v*.[2]

The output **y** vector from the DL model can be utilized to provide additional semi-quantitative analysis in addition to its function of designating the most probable mechanism. We propose that non-zero probabilities/fractions for mechanisms other than the most probable/prominent one suggest either a competing reaction or a gradual transition from one mechanism to another. In the one-electron reduction of $Co^{II}(TPP)$ in the presence of *n*-BuBr (Fig. 4c), the forward *C* step, namely the nucleophilic attack of $Co^{I}(TPP)^{\bullet-}$ towards *n*-BuBr, is pseudo-first-order on the concentration of *n*-BuBr ([*n*-BuBr]) and the forward rate constant k_f is proportional to [*n*-BuBr].[28,29] The corresponding dimensionless kinetic parameter λ, proportional to k_f/v (Fig. 4c), measures the competition between the solution reaction and redox specie's diffusion from/to the electrode surface.[2] As [*n*-BuBr] increases or *v* decreases, λ is larger and leads to more pronounced "irreversibility" from the *C* step in the voltammograms. Experimentally, we found that the cyclic voltammograms display a smaller tendency of reoxidation of $Co^{I}(TPP)^{\bullet-}$ when [*n*-BuBr] increases from 10 mM (Fig. 4d and 4e) to 50 mM (Fig. 4f and 4g), and when the range of *v* values decreases from $v \in [0.1, 1.0]$ V/s (Fig. 4d and 4f) to $v \in [0.05, 0.7]$ V/s (Fig. 4e and 4g). Accordingly, as shown in Fig. 4c, the *EC* component in the yielded **y** vectors gradually increased from 0.22 for Fig. 4d to 0.86 for Fig. 4f, while the component for E_r correspondingly

decreases from 0.78 to 0.14. Such results indicate that the DL model is capable of semi-quantitatively detecting the extent and gradual increase of *C* step, which could be valuable when studying systems with undesirable deactivation in redox cycling or desirable chemical transformation, amid an E_r system.

We speculate that the DL algorithm based on a set of voltammograms is capable of addressing, at least partly, the mechanistic ambiguity noted as "heterogeneous equivalent" by Feldberg and coworkers,[35] in which different reaction mechanisms may yield similar voltammograms that are indistinguishable within measurement errors. While the issue of "heterogeneous equivalent" is mostly prominent under a single voltammogram due to the limited information available from electrochemical measurement, our DL algorithm is based on a bijective relationship between mechanism and a more informative-rich set of voltammograms under different scan rates ($\{v, i(E)\}_n$). Hence, we contend that the issue of mechanistic ambiguity will be alleviated if not mitigated in the DL algorithm, which will be further evaluated in future studies. Yet the *EC* case of one-electron reduction of Co^{II}(TPP) in the presence of *n*-BuBr is already sufficiently exemplary (Fig. 4c to 4g). In Fig. 4d, the voltammogram at the largest scan rate *v* (1.0 V/s, black trace) can be seemingly interpreted as a quasi-reversible *E* process, while mechanistic ambiguity is resolved by the data at smaller *v* values. Meanwhile, the DL algorithm assigns 0.78 and 0.22 for *E* and *EC* process, acknowledging the possible existence of mechanistic ambiguity yet offering statistically meaningful diagnostic results useful for researchers. Such a probability-driven approach avoids the pitfalls of deterministic mechanistic assignment and will be beneficial towards addressing the issue of mechanistic ambiguity in the long run.

Conclusion

Electrochemical analysis for mechanistic investigation has relied heavily on manual inspection, which demands extensive training and may be prone to human bias and misinterpretation due to researchers' prior experience. In this work, we demonstrated a DL algorithm based on ResNet architecture that automatically analyzes cyclic voltammograms and, congruent with manual inspection, suggests the most probable mechanism among five of the most common mechanisms in homogenous molecular electrochemistry. Potentially being more sensitive and capable of detecting subtle elusive features at least within our parameter range (Table S1), the established DL model can also semi-quantitatively analyze competing pathways and observe the gradual transition from one mechanism to another. Additional factors that are known to impact the voltammograms, such as the 50/60 Hz noise of power-line frequency, will be incorporated into the DL model to increase its utility in practical applications. The established DL algorithm will be further refined with experimental data, being specifically targeted for algorithm development or contributed to by the general electrochemistry community via an open-access online platform. Efforts of expanding the types of electrochemical mechanisms, including homogenous, heterogeneous, stoichiometric and catalytic transformations, will advance the utility of the algorithm. In the long run, this DL-based approach will aid if not replace extensive manual mechanistic inspections in electrochemistry. We propose that such automatic analysis will find its advantages of analyzing complex reaction schemes that may be beyond the capacity of manual analysis, such as the square diagrams with the possibility of concerted pathways in systems of proton-coupled electron transfer.[5] The semi-quantitative output of the developed DL algorithm also offers a mathematically quantified feature, which can be the subject of Bayesian optimization that seeks to maximize electrochemical transformations with optimal experimental conditions. In conjunction with automatic robotic experimentation,[17,18] an autonomous high-throughput

electrochemistry research will become feasible, where Bayesian optimization strives to maximize certain transformations with intelligently varied experimental conditions (e.g. reactant type and concentrations) in an iterative fashion, by "learning" the parameter space via our DL algorithm and deciphering the partition of various reaction pathways measured in cyclic voltammetry.

Methods

Finite-element simulation of cyclic voltammograms

Finite-element simulations of cyclic voltammograms were conducted using COMSOL Multiphysics v5.5. The modules of Electrochemistry and Chemical Reaction Engineering were used for a one-dimensional model under the supporting electrolyte assumption with a time-dependent solver specialized for cyclic voltammetry, using an adaptive mesh with a maximal mesh size of 41 μm and a growth rate of 1.3. COMSOL simulations were iterated using COMSOL LiveLink™ which implements MATLAB R2020b. Random samples of variables were realized by Python 3 scripts and fed to COMSOL via MATLAB for the simulations of at least five consecutive cycles in cyclic voltammetry. Additional sanitization was implemented after COMSOL simulation to ensure the simulated cyclic voltammograms not only satisfy the corresponding mechanism but also are electrochemically accessible. A total of about 15,000 valid simulated cases, each containing cyclic voltammograms up to 6 different v values, were conducted. The detailed model information and the constrains of variables for each specific mechanism type are provided in the Supplementary Information.

Establishment of machine-learning algorithm

Simulated and experimental data was sanitized and translated in to the two-dimensional matrix $\{v, i(E)\}_n$ as reported in the main text (Fig. 2a) before the implementation of machine

learning. For each data point that is comprised of either simulated or experimental cyclic voltammograms at n number of v values ($\{v, i(E)\}_n$), the current densities i in voltammograms were normalized as $i_{\text{normalized}}$ against the largest i among all voltammograms in $\{v, i(E)\}_n$, with $i_{\text{normalized}}$ in the forward scan designated as positive value. The electrochemical potentials E were adjusted so that the adjusted electrochemical potential E_{adjusted} = 0 V roughly corresponds to the potential of studied redox couple. For voltammograms in which irreversibility precludes an accurate determination of redox potential, a rough estimate is automatically conducted based on the largest slope of the first rising redox peak. Interpolation and/or imputation of the *i-E* characteristics were conducted so that the two-dimensional matrix as input of the machine-learning model (Fig. 2a) does not explicitly contain the information of E. Therefore, the use of E_{adjusted} in this work is mostly for presentation purpose because the absolute values of E_{adjusted} are not inputs of the machine-learning model hence are not directly relevant to the automatic mechanistic analysis.

Machine-learning code was implemented on Jupyter notebooks using Python3 code. PyTorch machine learning frameworks were used to implement the various neural networks discussed in this work. The DL algorithms are trained to take either the first or second cycles of voltammograms for the same electrochemical system at various numbers of different scan rates ($\{v, i(E, \sigma)\}_n$, n = 1 to 6, σ = 0.0 to 1.0) and yields the vector $\mathbf{y} = \{y_1, y_2, y_3, y_4, y_5\}$. Because different starting potentials of voltammograms create additional variations for the first cycle of the voltammograms in both simulated and experimental scenarios, algorithms trained by the second cycles of voltammograms are used for the results reported here. Data of normalized cyclic voltammograms were processed by python API OpenCV to a three-dimensional tensor/matrix with a size of $\{6 \times 3 \times 500\}$ and labels $\{n, y, m\}$. Here, n has a dimension of six correlating to the

number simulated scan rates, m has a dimension of 500 correlating to the 500 potential values used during resizing, and y has a dimension of three corresponding to i_{for}, i_{rev}, and v_n which are the forward (i_{for}) and reverse (i_{rev}) normalized current values for scan rate v_n at potential m. When training models with $n < 6$, the size of the tensor remains the same and empty regions are filled with zeros so that the tensor size remained $\{6 \times 3 \times 500\}$ for all models. The input tensor was evaluated using a kernel/filter of size $\{6 \times 3 \times 7\}$. The kernel only views the data present in the tensor and the filter is not changed with different values of n. In terms of the hyperparameters of the ResNet-18 model, the standard learning rate of 1×10^{-3} and standard weight decay of 1×10^{-5} were used for all trainings. It takes approximately one hour to train the ResNet model and just seconds to predict the class of one sample. Training was performed on a machine using an Intel® Xeon® Silver 4214 CPU with 126 GB of RAM and an Nvidia GeForce RTX 2080 Ti GPU. A terminal training data accuracies of 99.95% during the establishment of our ResNet-18 model was achieved, and is in line with the reported test data accuracy of 98.5% (Fig. 2c). Such an alignment of accuracies during model training and testing is an indication of a model that generalizes well and is not overfit. Graphs were generated using the MatPlotLib library and the PyPlot module. Training data was input with stochastically added noise, after the raw/pre-noise data was normalized to have a global absolute current of 1, increase the robustness of the model to the noise encountered in real experimental data.

Final classifications were dictated by eight trained ResNet-18 models voting on final classification to decrease the effect of randomness in individual model training. For each simulated or experimental set of cyclic voltammograms, 8 individually trained ResNet-18 models will provide their **y** values. The values of yielded **y** vectors are subsequently averaged and the maximal component in the **y** vector is chosen as the final predicted type of electrochemical mechanism.

Relatively small, yet non-zero variability of predictions exist among 8 ResNet models. Statistically, the standard deviations of the yielded **y** vectors for a specific mechanism type are: 2.3×10^{-3} (*CE*), 9.3×10^{-3} (*EC*), 7.6×10^{-2} (*E*), 1.3×10^{-4} (*ECE*), 3.5×10^{-5} (*DISP*) (n = 100 for each mechanism class).

Experiments of electrochemical characterization

The tetraphenylporphyrin cobalt(II) (Co^{II}(TPP)) (80%), tetra-*n*-butylammonium hexafluorophosphate (*n*-Bu_4NPF_6) (98%) and tetra-*n*-butylammonium perchlorate (*n*-Bu_4NClO_4) (98%) were purchased from TCI America; anhydrous diethyl ether was purchased from Fisher Scientific; ferrocenium (Fc^+) hexafluorophosphate (98%) was purchased from Santa-Cruz Biotechnology; 1-bromobutane (*n*-BuBr) (99%), anhydrous *N*,*N*-dimethylformamide (DMF), anhydrous benzene, anhydrous acetonitrile, anhydrous tetrahydrofuran (THF), anhydrous dichloromethane, anhydrous pentane, dimanganese(0) decacarbonyl (98%), ethylenebis(diphenylphosphine) (99%), boric acid (99.5%), potassium chloride (99%), sodium hydroxide (99%) and 4-*tert*-butylcatechol (97%, HPLC) were purchased from Sigma-Aldrich. All the chemicals were used as received unless otherwise specified below. Bu_4NPF_6 and Bu_4NClO_4 salts were recrystallized from ethanol before use. Co^{II}(TPP) was recrystallized from methylene chloride before use. *n*-BuBr was fractionally distilled over $CaSO_4$ under N_2 at atmospheric pressure. The second fraction was collected at 102 °C and was dried over molecular sieves before use. THF was dried over molecular sieves before use. 4-*tert*-butylcatechol was distilled under reduced pressure and was allowed to recrystallize under vacuum at room temperature as a white crystalline solid before use.

The Mn complex [*trans*-$Mn(CO)_2(DPPE)_2$]PF_6 was synthesized according to a published procedure with some modifications.[32] Dimanganese(0) decacarbonyl (0.2 g, 0.5 mmol) and DPPE

(DPPE = ethylenebis(diphenylphosphine), 0.4 g, 1 mmol) were dissolved in 10 mL of benzene and the solution was refluxed under N_2 for 4 hrs. The [*trans*-$Mn(CO)_2(DPPE)$][$Mn(CO)_5$] salt was formed and collected as a yellowish solid. A portion of this solid (0.11 g, 0.1 mmol) was dissolved in 3 mL acetonitrile and 1 equivalent ferrocenium hexafluorophosphate (0.033 g, 0.1 mmol) was added to this solution and the reaction mixture was stirred vigorously for 30 min. Layering diethyl ether over this reaction mixture afforded an orange-yellow solid, which upon further recrystallization with dichloromethane/pentane afforded an orange-yellow crystalline solid (0.06 g, 59%). ^{31}P NMR ($CDCl_3$): δ 77.9 ppm (s) and −144.3 ppm (m).

Experiments of cyclic voltammetry were performed at room temperature using a CH Instruments 630D potentiostat. Solutions in organic solvents were performed under an Ar atmosphere in a glovebox (Vigor SG1200/750TS), while aqueous experiments were performed under N_2 atmosphere. *iR* corrections were conducted with positive feedback compensations for the ohmic drop. Ag/Ag^+ pseudo-reference electrode was calibrated against Fc^+/Fc redox after electrochemical measurements.

References

1. Bard, A. J. & Faulkner, L. R. *Electrochemical Methods: Fundamentals and Applications*. 2 edn, (John Wiley & Sons, inc., 2001).
2. Savéant, J.-M. & Costentin, C. *Elements of Molecular and Biomolecular Electrochemistry: An Electrochemical Approach to Electron Transfer Chemistry*. 2nd edn, (John Wiley & Sons, Inc., 2019).
3. Nicholson, R. S. Theory and Application of Cyclic Voltammetry for Measurement of Electrode Reaction Kinetics. *Anal Chem* **37**, 1351−1355 (1965).
4. Elgrishi, N. *et al.* A Practical Beginner's Guide to Cyclic Voltammetry. *J Chem Educ* **95**, 197−206 (2018).
5. Costentin, C., Robert, M. & Savéant, J.-M. Update 1 of: Electrochemical Approach to the Mechanistic Study of Proton-Coupled Electron Transfer. *Chem Rev* **110**, PR1−PR40 (2010).
6. Sandford, C. *et al.* A synthetic chemist's guide to electroanalytical tools for studying reaction mechanisms. *Chem Sci* **10**, 6404−6422 (2019).
7. Newman, J. & Thomas-Alyea, K. E. *Electrochemical Systems*. 3rd edn, (Wiley, 2004).
8. Rudolph, M., Reddy, D. P. & Feldberg, S. W. A Simulator for Cyclic Voltammetric Responses. *Anal Chem* **66**, 589A−600A (1994).
9. Gundry, L. *et al.* Recent advances and future perspectives for automated parameterisation, Bayesian inference and machine learning in voltammetry. *Chem Commun* **57**, 1855−1870 (2021).
10. Chen, H., Katelhon, E. & Compton, R. G. Predicting Voltammetry Using Physics-Informed Neural Networks. *J Phys Chem Lett* **13**, 536−543 (2022).

11. Chen, H., Katelhon, E., Le, H. & Compton, R. G. Use of Artificial Intelligence in Electrode Reaction Mechanism Studies: Predicting Voltammograms and Analyzing the Dissociative CE Reaction at a Hemispherical Electrode. *Anal Chem* **93**, 13360−13372 (2021).

12. Kennedy, G. F., Zhang, J. & Bond, A. M. Automatically Identifying Electrode Reaction Mechanisms Using Deep Neural Networks. *Anal Chem* **91**, 12220−12227 (2019).

13. Gundry, L., Kennedy, G., Bond, A. M. & Zhang, J. Inclusion of multiple cycling of potential in the deep neural network classification of voltammetric reaction mechanisms. *Faraday Discuss* **233**, 44−57 (2022).

14. He, K., Zhang, X., Ren, S. & Sun, J. Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770−778 (2016).

15. Amatore, C., Gareil, M. & Savéant, J. M. Homogeneous vs. heterogeneous electron transfer in electrochemical reactions: Application to the electrohydrogenation of anthracene and related reactions. *J Electroanal Chem Interf Electrochem* **147**, 1−38 (1983).

16. Evans, D. H. Solution electron-transfer reactions in organic and organometallic electrochemistry. *Chem Rev* **90**, 739−751 (1990).

17. Eyke, N. S., Koscher, B. A. & Jensen, K. F. Toward Machine Learning-Enhanced High-Throughput Experimentation. *Trends Chem* **3**, 120−132 (2021).

18. Dave, A. *et al.* Autonomous Discovery of Battery Electrolytes with Robotic Experimentation and Machine Learning. *Cell Rep Phys. Sci.* **1**, 100316 (2020).

19. Zihlmann, M., Perekrestenko, D. & Tschannen, M. Convolutional recurrent neural networks for electrocardiogram classification *2017 Computing in Cardiology (CinC).* 1−4 (2017)

20. Yoon, Y., Yan, B. & Surendranath, Y. Suppressing Ion Transfer Enables Versatile Measurements of Electrochemical Surface Area for Intrinsic Activity Comparisons. *J Am Chem Soc* **140**, 2397−2400 (2018).

21. Gao, R., Edwards, M. A., Harris, J. M. & White, H. S. Shot noise sets the limit of quantification in electrochemical measurements. *Curr Opin Electrochem* **22**, 170−177 (2020).

22. Zeiler, M. D. & Fergus, R. Visualizing and Understanding Convolutional Networks *Computer Vision – ECCV 2014.* 818−833 (2014).

23. Sermanet, P. *et al.* OverFeat: Integrated Recognition, Localization and Detection using Convolutional Networks. *International Conference on Learning Representations (ICLR) (Banff)* (2013).

24. Géron, A. *Hands-on Machine Learning with Scikit-Learn, Keras, and Tensorflow*. 2 edn, (O'Reilly, 2019).

25. Tsierkezos, N. G. Cyclic Voltammetric Studies of Ferrocene in Nonaqueous Solvents in the Temperature Range from 248.15 to 298.15 K. *J Solution Chem* **36**, 289−302 (2007).

26. Felton, R. H. & Linschitz, H. Polarographic Reduction of Porphyrins and Electron Spin Resonance of Porphyrin Anions1. *J Am Chem Soc* **88**, 1113−1116 (1966).

27. Whitlock, H. W. & Bower, B. K. Cobalt (I) meso-tetraphenylporphyrin. *Tetrahedron Lett* **6**, 4827−4831 (1965).

28. Lexa, D., Savéant, J. M. & Soufflet, J. P. Chemical catalysis of the electrochemical reduction of alkyl halides: Comparison between cobalt-tetraphenyl porphin and vitamin B_{12} derivatives. *J Electroanal Chem Interf Electrochem* **100**, 159−172 (1979).

29. Lexa, D., Mispelter, J. & Saveant, J. M. Electroreductive alkylation of iron in porphyrin complexes. Electrochemical and spectral characteristics of σ-alkylironporphyrins. *J Am Chem Soc* **103**, 6806−6812 (1981).

30. Bailey, S. I. & Ritchie, I. M. The effect of borate buffer on ortho quinone electrochemistry. *Electrochim Acta* **32**, 1027−1033 (1987).

31. Rafiee, M. & Nematollahi, D. Electrochemical study of catechol–boric acid complexes. *Electrochim Acta* **53**, 2751−2756 (2008).

32. Kuchynka, D. J. & Kochi, J. K. Facile rearrangement and electron transfer of 19-electron radicals from the reduction of the bischelated manganese carbonyl cation $Mn(CO)_2[PPh_2(CH_2)_2PPh_2]_2^+$. *Inorg Chem* **27**, 2574−2581 (1988).

33. Kuchynka, D. J. & Kochi, J. K. Equilibrium of 17-electron and 19-electron organometallic radicals derived from carbonylmanganese anions and cations. *Inorg Chem* **28**, 855−863 (1989).

34. Amatore, C. & Savéant, J. M. Electrochemical hydrogenation of aromatic hydrocarbons: Discrimination between ECE and disproportionation mechanisms by double potential step chronoamperometry. *J Electroanal Chem Interf Electrochem* **107**, 353−364 (1980).

35. Ruić, I. & Feldberg, S. The heterogeneous equivalent: A method for digital simulation of electrochemical systems with compact reaction layers. *J Electroanal Chem Interf Electrochem* **50**, 153−162 (1974).

Supplementary Information

General considerations for the model of cyclic voltammetry

We established a time-dependent one-dimensional model under the supporting electrolyte assumption for the COMSOL-based finite-element simulation of cyclic voltammograms.[1-3] The model numerically simulates the oxidative electrochemical systems, in which only the reduced species (*R*) are present in the solution before cyclic voltammetry. Only the oxidative electrochemical processes are needed in the training model thanks to the process of data pre-treatment and sanitization discussed in the main text. Below are the boundary and initial conditions in specific mechanistic scenarios. The ranges in variable's values and the sampling method (linearly or logarithmically) are discussed and summarized in Table S1. As shown below, the range in variable's values could be interdependent. Such interdependence and random sampling are implemented by python 3 scripts.

Partial differential equations

$$\frac{\partial C_i}{\partial t} = D_i \frac{\partial^2 C_i}{\partial x^2} + f_i \quad (1)$$

Here f_i denotes the mechanism-specific function that describes any possible *C* step in the solution. $f_i = 0$ denotes the absence of any homogenous *C* steps.

The diffusion coefficients D_i, sampled logarithmically, are assumed to be same *D* for all the molecular redox species in the solution. The hypothesis of constant *D* values is reasonable for two reasons: (1) The assumed reversible and quasi-reversible *E* step suggest a small reorganization energy λ and the resultant a small change of the molecular structure. (2) The value of *D* is relatively

insensitive to the changes of chemical identities since the scaling relationship between D and molecular weight is relatively weak ($D \propto MW^{-\frac{1}{3}}$) based on the Stokes-Einstein relationship.[4]

The initial concentration of the reduced species $C_{R,i}$ is linearly sampled from 0.1 mM to 100 mM with additional constraints listed below.

Boundary and initial conditions

Diffusion layer assumption[1,3] was implemented in the simulation. A finite diffusion layer L is implemented so that $x = 0$ denotes the electrode and $x = L$ denotes the boundary diffusion layer. In Nicholson's formalism of cyclic voltammetry[5] and presented below for single-electron transfer from O to R with the period of triangular voltage wave as λ, function $S_\lambda(t)$ describes the temporal concentration variation of O in the presence of diffusion for each period of triangular voltage wave,

$$D_O \frac{\partial [O]}{\partial x} = k_s \left(\frac{C_{O,i}}{C_{R,i}}\right)^{-\alpha} [S_\lambda(t)]^{-\alpha} \left[[O]_{x=0} - \frac{C_{O,i}}{C_{R,i}} S_\lambda(t) [R]_{x=0}\right]$$

$$S_\lambda(t) = \begin{cases} e^{-at}, t < \lambda \\ e^{at-2a\lambda}, t > \lambda \end{cases}, a = \frac{Fv}{RT}$$

in which k_s is the standard rate constant of interfacial charge transfer and $\alpha = 0.5$ is the transfer coefficient.

The above expression suggests that the characteristic time constant of diffusional behavior is $\frac{RT}{Fv}$ for $S_\lambda(t)$. Therefore, in our simulation, the thickness of the diffusion layer L is adaptively chosen so that the L is more than six times of the characteristic length scale of diffusion within the noted characteristic time constant when T = 298.15 K (same below).

$$L = 6\sqrt{D \cdot \frac{RT}{Fv}} \quad (2)$$

Here the scan rate *v* is evenly sampled both logarithmically and linearly between 0.01 to 2 V/s. Additional algorithms to sample *n* number of different *v* values in the same simulated electrochemical systems is extensively discussed below.

In addition to the Faradaic processes simulated below, capacitive double-layer charging events are also simulated in all mechanistic scenarios, with double-layer capacitance C_{dl} randomly sampled linearly between 5 to 35 μF/cm^2 based on literature values.[6] The capcacitive current i_{dl} is simulated based the following equation,

$$i_{dl} = C_{dl}v \quad (3)$$

The current model does not include uncompensated resistance hence the *iR* drop. We contend that any serious mechanistic electrohcemical analysis should all be based on experimental data whose *iR* drop has been much minimized, if not completely mitigated, through judicious instrument setting during the experimental characterization. Moreover, as the reported deep-learning (DL) algorithm does not intend to evaluate the reversibility of charge transfer within the *E* mechanism, the possible convolution between *iR* drop and quasi-reversible charge transfer will not negatively impact the practical utility of DL algorithm. Nonetheless, future versions of DL algorithms will consider including the impact of *iR* drops in the training data.

_E_ mechanism

E: $R \overset{E}{\leftrightarrow} O + e^-$

$f_R = 0,\ f_O = 0$ in eq. (1).

$[R]_{t=0} = C_{R,i}, [O]_{t=0} = 0$

The thermodynamic potential of the *O/R* redox couple is $E_{O/R}$ = 0 V versus an arbitrary reference electrode. The cyclic voltammograms are simulated with a potential window in which the anodic bound $E_{window,a}$ is linearly sampled between 0.5 and 1 V vs. NHE and the cathodic bound $E_{window,c}$ is linearly sampled between −0.5 and −1 V vs. NHE. The starting potential of the cyclic voltammogram E_{start} is linearly sampled between −0.2 V vs. NHE and $E_{window,c}$. Such an arrangement of E_{start} ensures that there is minimal transient current at the beginning of voltage sweep.

Concentration-dependent Butler-Volmer equation[1] is employed to define the *E* step at the electrode interface.

$$i(t) = i_0 \left\{ \frac{[R]_{x=0}}{0.5 \cdot C_{R,i}} \exp\left[\frac{\alpha F}{RT}\left(E - E_{O/R}\right)\right] - \frac{[O]_{x=0}}{0.5 \cdot C_{R,i}} \exp\left[-\frac{(1-\alpha)F}{RT}\left(E - E_{O/R}\right)\right]\right\} \quad (4)$$

Here, $\alpha = 1/2$ and $0.5 \cdot C_{R,i}$ denotes the equilibrium concentration when $E = E_{O/R}$ and $C_R = C_O$. The exchange current density i_0 is logarithmically sampled with the upper-bound $i_{o,upper}$ and lower-bound $i_{o,lower}$.

The transfer coefficient α here is assumed to be a constant value (0.5) for 3 reasons: (1) Most of the elemental charge transfer *E* step in practical application possess a value of α quite close to 0.5.[1,2] Indeed, many of the classical analysis in electrochemical mechanism, such as the evaluation of Tafel slope in electrocatalysis,[1,2] can only be practically usable by assuming α = 0.5. (2) Since the transfer coefficient α only alters the shape the waveform,[1,5] we conjecture that the impact of α could be limited in the context of differentiating *E*, *EC*, *CE*, *ECE* and *DISP* mechanisms, whose

waveforms will possess much more significant variations. (3) For the semi-quantitative analysis based on our DL model as shown in Fig. 4, the relative values of **y** vectors yielded from DL algorithms remain comparable within the same electrochemical system, because the values of α in the individual *E* step remain the same under different experimental conditions. Nonetheless, in an updated version of our DL algorithm, α will be considered a variable in the future.

Following the Nicholson's formalism in cyclic voltammetry,[5] i_0 is dependent on the standard rate constant of surface change transfer k_s:

$$i_0 = k_s \cdot 0.5FC_{R,i} = \psi\sqrt{\frac{\pi FvD}{RT}} \cdot 0.5FC_{R,i} \quad (5)$$

We chose $\psi \in [10, 0.3]$ following the Nicholson's formalism[5], which corresponding to a peak separation ΔE_p = 62 ~ 120 mV in the cyclic voltammograms. We note that the upper bound of ψ values is high enough that the resultant scenarios resemble the Nernstian scenario in cyclic voltammetry in which the interfacial charge transfer is fast enough to ensure a Nernstian equilibrium for the redox species in the immediate proximity near the electrode.[1]

$$i_{0,upper} = 10 \cdot 0.5FC_{R,i}\sqrt{\frac{\pi FvD}{RT}} \quad (6a) \qquad i_{0,lower} = 0.3 \cdot 0.5FC_{R,i}\sqrt{\frac{\pi FvD}{RT}} \quad (6b)$$

As we want to ensure detectable peaks in cyclic voltammograms, additional constraint about the minimal concentration of redox *O*/*R* ($C_{R,i}$) are needed. We estimated the current densities of the redox peaks based on Randle-Sevcik equation[1,7] and ensure that the estimated current densities are approximately 5 times or more of the background current density from the capacitive double-layer charging/discharging.[1,8]

$$C_{R,i,min} = \max\left\{0.1\ mM, \frac{5}{0.446F} C_{dl} v \sqrt{\frac{RT}{FvD}}\right\} \quad (7)$$

_EC_ mechanism

E: $R \overset{E}{\leftrightarrow} O + e^-$

C: $O \overset{C}{\leftrightarrow} A$

$f_R = 0,\ f_O = k_b[A] - k_f[O],\ f_A = k_f[O] - k_b[A]$ in eq. (1).

$K_{O/A} \equiv \frac{k_f}{k_b}$

$[R]_{t=0} = C_{R,i},\ [O]_{t=0} = 0, [A]_{t=0} = 0.$

Most of the constraints in the *EC* mechanism are the same as the E_r mechanism with the following additional constraints.

The equilibrium constant of the *C* step $K_{O/A}$ is logarithmically sampled between $10^{0.5}$ ~ 10^3.

The kinetic rate constant of *C* step in the forward direction k_f is logarithmically sampled within the following upper and lower bound so that $\log_{10}\lambda \equiv log_{10}\left(\frac{RT}{F}\frac{k_f+k_b}{v}\right) \in [-2, 4.5]$.

$$k_{f,upper} = 10^{4.5} \cdot \frac{F}{RT} \cdot \frac{vK_{O/A}}{K_{O/A}+1} \quad (8a) \qquad k_{f,lower} = 10^{-2} \cdot \frac{F}{RT} \cdot \frac{vK_{O/A}}{K_{O/A}+1} \quad (8b)$$

The above ranges of $K_{O/A}$ and k_f values captures all of the possible variations in the *EC* mechanism as illustrated in Fig. 2.1 of Savéant's textbook (DO, KO, KG, KP, KE, and DE zones),[2]

before the small value of $K_{O/A}$ leads to situations that are indeed the E_r mechanism and presented at the very upper part of that figure.

Because of the resultant potential shifts of redox peaks in the *E*C mechanism, the E_{start} is now linearly sampled between $E_{window,c}$ and $-0.2 - 0.059 log_{10} K_{O/A}$ V vs. NHE.

_CE_ mechanism

C: $A \overset{C}{\leftrightarrow} R$

E: $R \overset{E}{\leftrightarrow} O + e^-$

$f_A = k_b[R] - k_f[A]$, $f_R = k_f[A] - k_b[R]$, $f_O = 0$ in eq. (1).

$$K_{R/A} \equiv \frac{k_f}{k_b}$$

$$[A]_{t=0} = C_{R,i} \frac{1}{K_{R/A} + 1},\ [R]_{t=0} = C_{R,i} \frac{K_{R/A}}{K_{R/A} + 1}, [O]_{t=0} = 0.$$

Most of the constraints in the *CE* mechanism are the same as the E_r mechanism with the following additional constraints.

The equilibrium constant of the *C* step $K_{R/A}$ is logarithmically sampled between $10^{-3} \sim 10^{-0.5}$.

The kinetic rate constant of *C* step in the forward direction k_f is logarithmically sampled within the following upper and lower bound so that $\log_{10} \lambda \equiv log_{10} \left(\frac{RT}{F} \frac{k_f + k_b}{v} \right) \in [-2, 4.5]$.

$$k_{f,upper} = 10^{4.5} \cdot \frac{F}{RT} \cdot \frac{v K_{R/A}}{K_{R/A} + 1} \qquad (9a)$$

$$k_{f,lower} = 10^{-2} \cdot \frac{F}{RT} \cdot \frac{v K_{R/A}}{K_{R/A} + 1} \qquad (9b)$$

The above ranges of $K_{R/A}$ and k_f values captures all of the possible variations in the *CE* mechanism as illustrated in Fig. 2.8 of Savéant's textbook (DO, KO, KG, KP, KE, and DE zones),[2] before the large value of $K_{R/A}$ leads to situations that are indeed the *E* mechanism and presented at the very upper part of that figure.

Because of the resultant potential shifts of redox peaks in the *CE* mechanism, the anodic bound of electrochemical window $E_{\text{window,a}}$ is now linearly sampled between $0.5 + 0.059log_{10}K_{R/A}$ and $1.0 + 0.059log_{10}K_{R/A}$ V vs. NHE, and the E_{start} is now linearly sampled between $E_{\text{window,c}}$ and $-0.2 + 0.059log_{10}K_{R/A}$ V vs. NHE.

ECE mechanism

E: $R1 \overset{E}{\leftrightarrow} O1 + e^-$

C: $O1 \overset{C}{\rightarrow} R2$

E: $R2 \overset{E}{\leftrightarrow} O2 + e^-$

$f_{R1} = 0,\ f_{O1} = -k[O1],\ f_{R2} = k[O1], f_{O2} = 0$ in eq. (1).

$[R1]_{t=0} = C_{R,i},\ [O1]_{t=0} = 0, [R2]_{t=0} = 0, [O2]_{t=0} = 0.$

The *E* step between *R1* and *O1* follows the same definition of *R* and *O* in the *E* mechanism.

The kinetic rate constant *k* of the *C* step is logarithmically sampled with the following constraints so that $\log_{10}\lambda \equiv log_{10}\left(\frac{RT}{F}\frac{k}{v}\right) \in [-1, 3]$

$$k_{upper} = 10^3 \frac{F}{RT} v \qquad (10a) \qquad\qquad k_{lower} = 10^{-1} \frac{F}{RT} v \qquad (10b)$$

The selection of above *k* range covers almost all of the possible variations in the *ECE* mechanism as illustrated in Fig. 2.9 and Fig. 10a of Savéant's textbook.[2]

The *E* step between *R2* and *O2* are defined with its thermodynamic redox potential $E_{O2/R2}$ linearly sampled between −0.7 and −0.18 V vs. NHE. The electrochemical kinetics of the E_r step is defined as a concentration dependent Butler-Volmer process illustrated in eq. (4). The standard rate constant of surface change transfer $k_{s,O2/R2}$ and the corresponding exchange current density $i_{0,O2/R2}$ is defined and sampled similarly as the i_0 in the E_r mechanism.

$$i_{0,O2/R2} = k_{s,O2/R2} \cdot 0.5FC_{R,i} = \psi' \sqrt{\frac{\pi F v D}{RT}} \cdot 0.5FC_{R,i} \qquad (11)$$

When ψ' was chose as $\psi' \in [10, 0.3]$, we have,

$$i_{0,O2/R2,upper} = 10 \cdot 0.5FC_{R,i} \sqrt{\frac{\pi F v D}{RT}} \quad (12a) \qquad i_{0,O2/R2,lower} = 0.3 \cdot 0.5FC_{R,i} \sqrt{\frac{\pi F v D}{RT}} \quad (12b)$$

In order to accommodate the additional redox features, the E_{start} is now linearly sampled between $E_{window,c}$ and −0.6 V vs. NHE.

***DISP1* mechanism**

E: $R1 \overset{E}{\leftrightarrow} O1 + e^-$

C: $O1 \overset{C}{\rightarrow} R2$

$DISP$: $R2 + O1 \xrightarrow{DISP} O2 + R1$

$f_{R1} = k_{DISP}[R2][O1]$, $f_{O1} = -k[O1] - k_{DISP}[R2][O1]$, $f_{R2} = k[O1] - k_{DISP}[R2][O1]$, $f_{O2} = k_{DISP}[R2][O1]$ in eq. (1).

$[R1]_{t=0} = C_{R,i}$, $[O1]_{t=0} = 0, [R2]_{t=0} = 0, [O2]_{t=0} = 0$.

The *E* step between *R1* and *O1* follows the same definition of *R* and *O* in the *E* mechanism.

The E_{start} and the kinetic rate constant *k* of the *C* step follow the same definition of *k* in the *ECE* mechanism.

The kinetic rate constant k_{DISP} of the *DISP* step is logarithmically sampled with the following constraints.

$$k_{DISP,upper} = 10^{2}\frac{k^{3/2}}{C_{R,i}}\sqrt{\frac{RT}{Fv}} \quad (13a) \qquad\qquad k_{DISP,upper} = 10^{-2}\frac{k^{3/2}}{C_{R,i}}\sqrt{\frac{RT}{Fv}} \quad (13b)$$

The above definition of k_{DISP} covers almost the full phase diagram noted by Savéant and coworkers[2,9] since the corresponding p_{ECE}^{DISP}, defined below, is within the range of [10^{-2}, 10^{2}].

$$p_{ECE}^{DISP} = \frac{k_{DISP}}{k^{3/2}}C_{R,i}\sqrt{\frac{Fv}{RT}} \quad (14)$$

The above definition indeed may also include scenarios that is similar, but not quite the same, to the *DISP2* mechanism,[9] when the *DISP* step is slow and rate-limiting (yet the limiting case of *DISP2* mechanism requires a reversible pre-equilibrium for the *C* step between *O1* and *R2*). Such slight ambiguity of simulated voltammograms in the training data will be addressed in future versions of the algorithm.

Additional considerations when sampling scan rate *v*

In the sampling of simulated cyclic voltammograms, variables intrinsic to the chemistry of the electrochemical systems are first sampled either linearly or logarithmically. Variables related to the electrochemical testing conditions, including E_{start}, $E_{window,a}$, $E_{window,c}$, and v are sampled subsequently. Particular attention was paid the sampling of v since multiple chemistry-intrinsic variables are also dependent on the v values as shown in Table S1. Because we aim to obtain up to 6 simulated cyclic voltammograms with different v values (n = 6), an iterative process of variable samplings was implemented in the python 3 scripts as shown below.

Step 1: After the initial generation of random combinations of chemistry-related variables listed in Table S1, a medium scan rate v_{medium} is linearly sampled between 0.1 to 0.5 V/s, the range of v mostly commonly used in cyclic voltammetry. We note that as shown below v_{medium} serves as a temporal variable in the selection of v values and is not numerically used in simulation. The current densities of the redox peak is roughly estimated based on the Randle-Sevcik equation.[1,6]

$$i_{medium} = 0.446C_{R,i}\sqrt{\frac{FDv_{medium}}{RT}} \tag{15}$$

Step 2. The maximal and minimal scan rate v_{max} and v_{min} were randomly selected based on the following constraints.

$$v_{max} = min\left\{2, \frac{RT}{\pi FD}\left(\frac{10i_0}{FC_{R,i}}\right)^2, \frac{FD}{RT}\left(\frac{0.446FC_{R,i}}{5C_{dl}}\right)^2, 10^{0.6}\cdot i_{medium}\right\} \tag{16a}$$

$$v_{min} = max\left\{0.01, \frac{RT}{\pi FD}\left(\frac{i_0}{10FC_{R,i}}\right)^2, 10^{-0.6}\cdot i_{medium}\right\} \tag{16b}$$

The above constraints ensure that v_{max} and v_{min} are within the ranges of 0.01 to 2 V/s, the peak separations in the Nicholson's formalism[5] will not deviate too much from the targeted values separation (ΔE_p = 62 ~ 120 mV), the voltammograms at maximal scan rates won't lead to indistinguishable redox peaks due to capacitive double-layer charging/discharging, and there is significant differences, $10^{0.6}$ ~ 4 fold difference in current densities, among the n number of simulated cyclic voltammograms.

Step 3. If $v_{max} - v_{min} \leq 0.5\ V/s$, go back to Step 1 again. Otherwise, proceed to Step 4.

Step 4. 4 more additional v values are linearly or logarithmically sampled between the v_{max} and v_{min}, leading to 6 values of v in total (n = 6).

Additional discussion about the number of voltammograms *n* needed for mechanism determination

As discussed in the main text and presented in Fig. 2e, when $n \geq 2$ the prediction accuracies of DL models trained by $\{v, i(E, \sigma)\}_n$ (n = 1 to 6, σ = 0.3) more or less remain equally satisfactory (> 95%). Such results suggest that within the tested set of simulated voltammograms, statistically on average there is diminishing returns of prediction accuracy when $n > 2$. Here we provide additional discussion and illustrate the parameter range when such a statement is applicable, given the defined parameter space of simulated voltammograms provided in the Supplementary Information and listed in Table S1.

In section "Additional considerations when sampling scan rate v" of the Supplementary Information, we reported the procedures of selecting the maximal and minimal values of v in the training set of simulated voltammograms. When n = 2, we propose that there exists the following

approximate relationship between the two v values (v_{high} and v_{low}) and the medium value v_{medium}:

$$v_{medium} \approx \frac{v_{high} + v_{low}}{2} \in [0.1, 0.5]\ V/s \tag{17}$$

Here the "≈" sign suggests that the above relationship is a statistically approximation given that v_{high} and v_{low} are randomly sampled around the value of v_{medium}.

Hence, the approximate range of v_{high} and v_{low} can be defined as:

$$v_{high} = min\left\{2, \frac{RT}{\pi FD}\left(\frac{10 i_0}{F C_{R,i}}\right)^2, \frac{FD}{RT}\left(\frac{0.446 F C_{R,i}}{5 C_{dl}}\right)^2, 10^{0.6} \cdot 0.446 C_{R,i}\sqrt{\frac{FD\left(v_{high} + v_{low}\right)}{2RT}}\right\}$$

$$v_{low} = max\left\{0.01, \frac{RT}{\pi FD}\left(\frac{i_0}{10 F C_{R,i}}\right)^2, 10^{-0.6} \cdot 0.446 C_{R,i}\sqrt{\frac{FD\left(v_{high} + v_{low}\right)}{2RT}}\right\}$$

$$v_{max} - v_{min} \leq 0.5\ V/s \tag{18a, 18b, 18c}$$

Equation (18a), (18b), and (18c) provide an approximate empirical range of v_{high} and v_{low} values in order to satisfy the defined training data set of voltammograms and hence offer good accuracy of mechanistic prediction based on our DL model. The above relationships indicate that v_{high} and v_{low} values are dependent on the redox species' concentration ($C_{R,i}$) and diffusion coefficient (D), the electrodes' double-layer capacitance (C_{dl}), and the exchange current density (i_0) hence the standard rate constant of interfacial charge transfer (k_s) based on equation (5). A combination of experimental parameters ($C_{R,i}$ and C_{dl}) and redox's intrinsic properties (D and k_s) determines the values of v_{high} and v_{low} for effective discernment of electrochemical mechanisms.

Table S1. The variables and the corresponding value ranges in the numerical simulation

Variable	Value ranges	Constraints
E_r mechanism		
v	0.01 ~ 2 V/s [a,b]	Default setting, see specific mechanisms below if noted
$C_{R,i}$	$C_{R,i,min}$ ~ 100 mM [a]	$C_{R,i,min} = \max\left\{0.1\ mM, \frac{5}{0.446F} C_{dl} v \sqrt{\frac{RT}{FvD}}\right\}$
D	1×10^{-7} ~ 1×10^{-4} cm^2/s [a,c]	
L	N.A.	$L = 6\sqrt{D \cdot \frac{RT}{Fv}}$
C_{dl}	5 ~ 35 μF/cm^2 [b]	
$E_{window,a}$	0.5 ~ 1 V vs. NHE [b]	Default setting, see specific mechanisms below if noted
$E_{window,c}$	−0.9 ~ −1.4 V vs. NHE [b]	
E_{start}	$E_{window,c}$ ~ −0.2 V vs. NHE [b]	Default setting, see specific mechanisms below if noted
i_0	$i_{0,lower}$ ~ $i_{0,upper}$ [a]	$i_{0,lower} = 0.3 \cdot 0.5FC_{R,i}\sqrt{\frac{\pi FvD}{RT}}$ $i_{0,upper} = 10 \cdot 0.5FC_{R,i}\sqrt{\frac{\pi FvD}{RT}}$ Default setting, see specific mechanisms below if noted
E_rC_r mechanism		
$K_{O/A}$	$10^{0.5}$ ~ 10^{3} [a]	
k_f	$k_{f,lower}$ ~ $k_{f,upper}$ [a]	$k_{f,lower} = 10^{-2} \cdot \frac{F}{RT} \cdot \frac{vK_{O/A}}{K_{O/A}+1}$ $k_{f,upper} = 10^{4.5} \cdot \frac{F}{RT} \cdot \frac{vK_{O/A}}{K_{O/A}+1}$
E_{start}	$E_{window,c} \sim -0.2 - 0.059log_{10}K_{O/A}$ V vs. NHE [b]	
C_rE_r mechanism		
$K_{R/A}$	10^{-3} ~ $10^{-0.5}$ [a]	
k_f	$k_{f,lower}$ ~ $k_{f,upper}$ [a]	$k_{f,lower} = 10^{-2} \cdot \frac{F}{RT} \cdot \frac{vK_{R/A}}{K_{R/A}+1}$ $k_{f,upper} = 10^{4.5} \cdot \frac{F}{RT} \cdot \frac{vK_{R/A}}{K_{R/A}+1}$
$E_{window,a}$	$0.5 + 0.059log_{10}K_{R/A}$ ~ $1.0 + 0.059log_{10}K_{R/A}$ vs. NHE [b]	
E_{start}	$E_{window,c} \sim -0.2 + 0.059log_{10}K_{R/A}$ V vs. NHE [b]	
***ECE* mechanism**		
k	k_{lower} ~ k_{upper} [a]	$k_{upper} = 10^3 \frac{F}{RT} v$

		$k_{lower} = 10^{-1}\frac{F}{RT}v$
$E_{O2/R2}$	−0.7 ~ −0.18 V vs. NHE [b]	
$i_{0,O2/R2}$	$i_{0,O2/R2,lower}$ ~ $i_{0,O2/R2,upper}$ [a]	$i_{0,O2/R2,upper} = 10 \cdot 0.5FC_{R,i}\sqrt{\frac{\pi FvD}{RT}}$ $i_{0,O2/R2,lower} = 0.3 \cdot 0.5FC_{R,i}\sqrt{\frac{\pi FvD}{RT}}$
E_{start}	$E_{window,c}$ ~ −0.6 vs. NHE [b]	
***DISP* mechanism**		
k	k_{lower} ~ k_{upper} [a]	$k_{upper} = 10^{3}\frac{F}{RT}v$ $k_{lower} = 10^{-1}\frac{F}{RT}v$
E_{start}	$E_{window,c}$ ~ −0.6 vs. NHE [b]	
k_{DISP}	$k_{DISP,lower}$ ~ $k_{DISP,upper}$ [a]	$k_{DISP,upper} = 10^{2}\frac{k^{3/2}}{C_{R,i}}\sqrt{\frac{RT}{Fv}}$ $k_{DISP,upper} = 10^{-2}\frac{k^{3/2}}{C_{R,i}}\sqrt{\frac{RT}{Fv}}$

[a] Variable values are randomly sampled logarithmically
[b] Variable values are randomly sampled linearly.

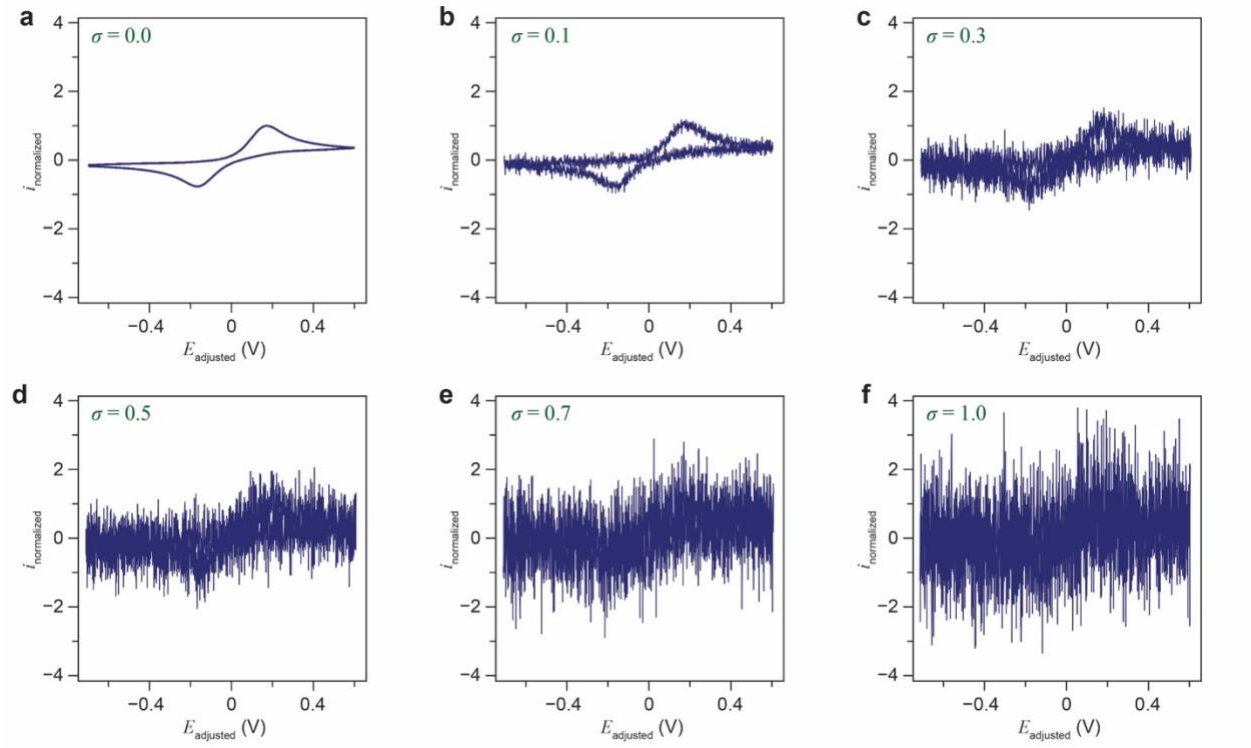


Figure S1. Exemplary simulated cyclic voltammograms with different levels of Gaussian noises. The same simulated cyclic voltammograms added with Gaussian noises whose standard deviations σ = 0.0 (**a**), 0.1 (**b**), 0.3 (**c**), 0.5 (**d**), 0.7 (**e**), and 1.0 (**f**). The voltammograms of the second cycles are displayed.

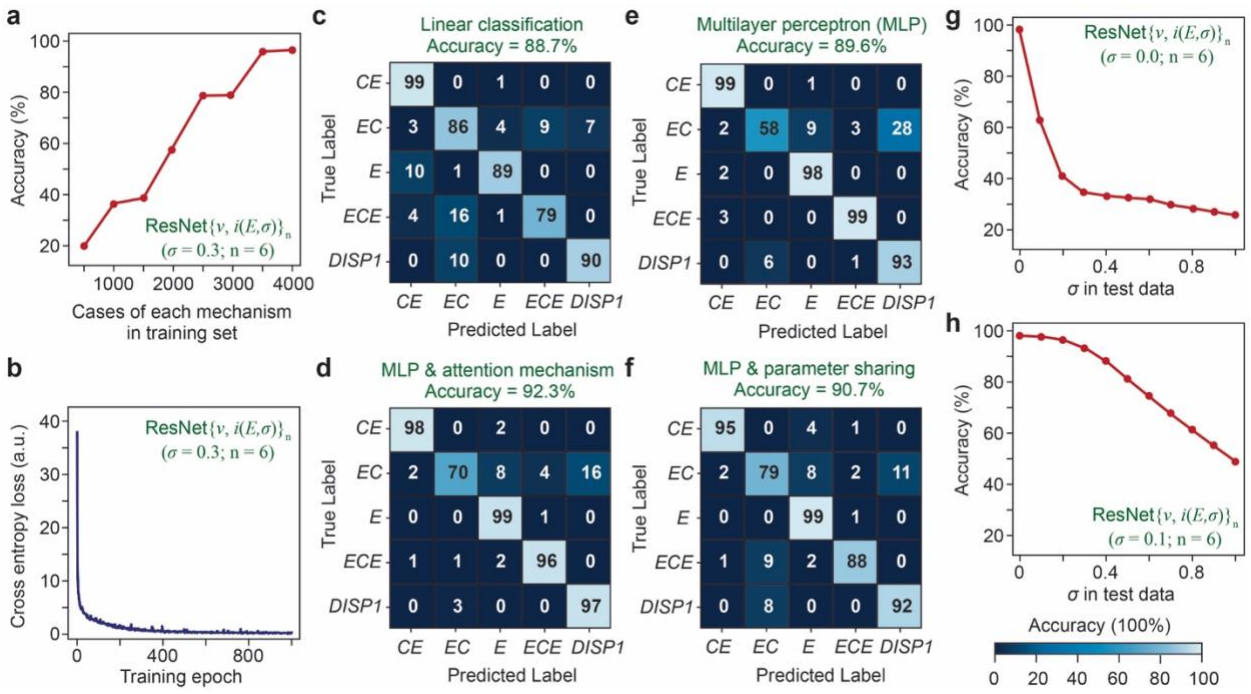


Figure S2. Training of machine-learning algorithms for cyclic voltammetry. **a**, Accuracies of RetNet{v, $i(E, \sigma)\}_n$ (σ = 0.3; n = 6) trained by varying numbers of simulated electrochemical systems in each mechanism in the training set. **b**, Cross entropy loss, a surrogate of the algorithms' accuracy in the training process, as a function of training epochs for ResNet{v, $i(E, \sigma)\}_n$ (σ = 0.3; n = 6). **c** to **f**, the confusion matrices and the overall accuracy of various machine-learning models trained by simulated cyclic voltammograms. Training data, {v, $i(E, \sigma)\}_n$ (σ = 0.3; n = 6). **g** and **h**, the accuracies of the DL models trained by {v, $i(E, \sigma)\}_n$ (σ = 0.0; n = 6) and v, $i(E, \sigma)\}_n$ (σ = 0.1; n = 6), respectively, when tested with simulated voltammograms with varying values of n and σ. Here the "linear classification" model is the simplest as it require that the high dimensional data be linearly separable. In the fully-connected MLP every node of every layer is connected to every node of the layers before it and after it. Instead, the ResNets architecture are not fully connected and further have residual layers in place which act as mechanisms to prevent training problems such as exploding and vanishing gradients that can occur during the learning process/optimization of network architectures.

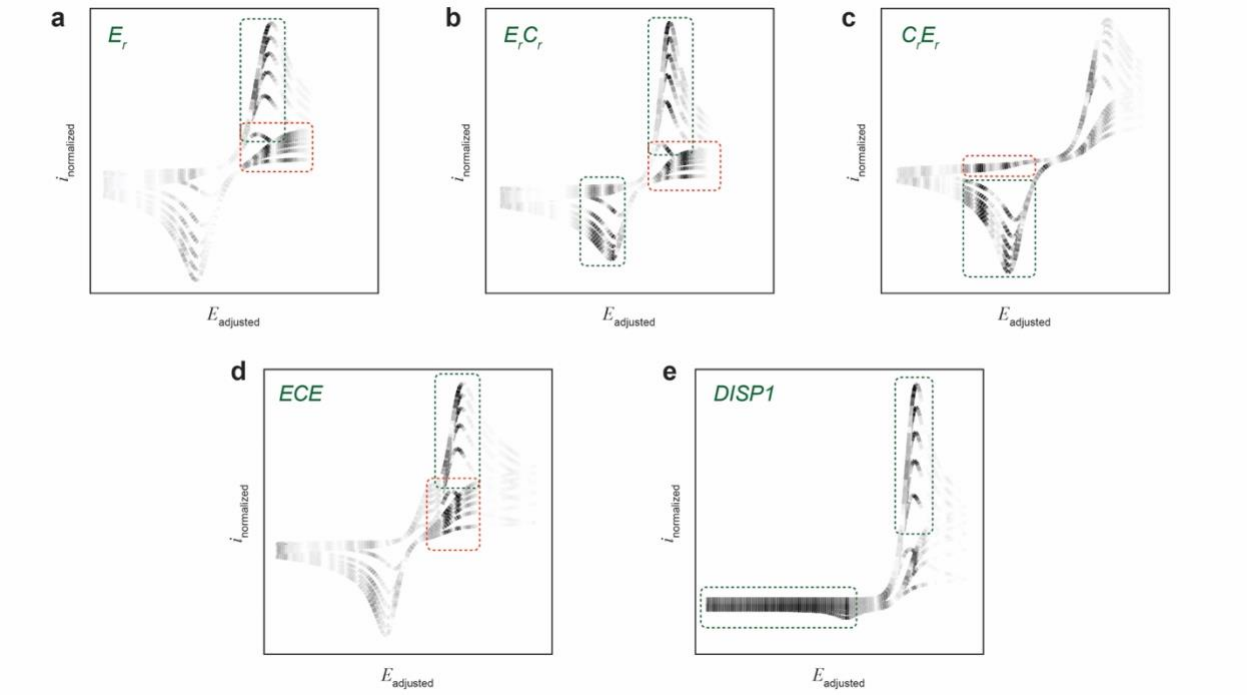


Figure S3. The "importance" plots of simulated cyclic voltammograms. Exemplary simulated cyclic voltammograms (σ = 0.0) and the corresponding "importance" in RetNet$\{v, i(E, \sigma)\}_n$ (σ = 0.3; n = 6) for *E* (**a**), *EC* (**b**), *CE* (**c**), *ECE* (**d**), and *DISP1* (**e**) mechanisms. The "importance" towards the DL model in expected (green) and somewhat unexpected (red) parts in the voltammograms are highlighted. The voltammograms of the second cycles are displayed and analyzed.

Supplementary References

1. Bard, A. J. & Faulkner, L. R. *Electrochemical Methods: Fundamentals and Applications*. 2 edn, (John Wiley & Sons, inc., 2001).
2. Savéant, J.-M. & Costentin, C. *Elements of Molecular and Biomolecular Electrochemistry: An Electrochemical Approach to Electron Transfer Chemistry*. 2nd edn, (John Wiley & Sons, Inc., 2019).
3. Newman, J. & Thomas-Alyea, K. E. *Electrochemical Systems*. 3rd edn, (Wiley, 2004).
4. Young, M. E., Carroad, P. A. & Bell, R. L. Estimation of diffusion coefficients of proteins. *Biotechnol Bioeng* **22**, 947−955 (1980).
5. Nicholson, R. S. Theory and Application of Cyclic Voltammetry for Measurement of Electrode Reaction Kinetics. *Anal Chem* **37**, 1351−1355 (1965).
6. Yoon, Y., Yan, B. & Surendranath, Y. Suppressing Ion Transfer Enables Versatile Measurements of Electrochemical Surface Area for Intrinsic Activity Comparisons. *J Am Chem Soc* **140**, 2397−2400 (2018).
7. Elgrishi, N. *et al.* A Practical Beginner's Guide to Cyclic Voltammetry. *J Chem Educ* **95**, 197−206 (2018).
8. Gao, R., Edwards, M. A., Harris, J. M. & White, H. S. Shot noise sets the limit of quantification in electrochemical measurements. *Curr Opin Electrochem* **22**, 170−177 (2020).
9. Amatore, C., Gareil, M. & Savéant, J. M. Homogeneous vs. heterogeneous electron transfer in electrochemical reactions: Application to the electrohydrogenation of anthracene and related reactions. *J. Electroanal Chem Interf Electrochem* **147**, 1−38 (1983).

Chapter 5. An object detection approach for the automated analysis of more complex cyclic voltammograms

This chapter is a version of unsubmitted work by Hoar, B.B.; Zhang, W.; Chen Y.; Sun, J.; Costentin, C.; Gu, Q.; Liu, C. "Custom one-dimensional Faster R–CNN object detection for the analysis of multi-event cyclic voltammograms" *in preparation*

Abstract

In the field of electrochemical analysis, mechanism assignment is fundamental to the understanding of a system. The mechanism class describing the general reaction scheme for a system is foundational to the subsequent quantitative evaluation of that system, but this class is often deduced from relatively subjective analysis of cyclic voltammograms. Fortunately, deep learning techniques provide the means for development of automated mechanism assignment algorithms that can support experimentalists in deduction of electrochemical steps in a system. Herein, we present a custom Faster R-CNN architecture, "CVNet," capable of assigning both mechanism and voltage window to electrochemical events within larger parent cyclic voltammograms. This algorithm, trained on simulated cyclic voltammetry data, can detect up to four individual electrochemical events from eight possible mechanism classes within a set of cyclic voltammetry data. The developed technique shows up to a 97.2% accuracy on detected training redox events and an overall inference F_1 score of 0.932, with over 96% of all electrochemical events detected. With this evolution, many of the constraints of previous automated electrochemical mechanism assignment approaches are removed, showing significant practical improvements.

Introduction

Cyclic voltammetry is one of the most fundamental electrochemical techniques.[1-3] In fact, one often doesn't need to look beyond the cover of electrochemical textbooks to see the famous "duck shaped" plots of cyclic voltammetry experiments.[2-4] Despite its foundational place in the pantheon of electroanalytical tools, there is no consistent heuristic for its use in mechanism assignment – perhaps the most common use of cyclic voltammetry (CV).[2, 5] In practice, the qualitative and semi-quantitative correlation between current-potential response as a function of scan rate $\{\nu, i(E)\}_n$[6] (ν for scan rate, n for number of measured scan rates) in CV is necessary not only for deduction of a system's constituent chemical (*C* steps) and electrochemical (*E* steps) reactions, it is indeed a prerequisite for downstream quantitative analyses, for example, the extraction of rate constants for a chemical step.[1, 2, 7-9] Despite cyclic voltammetry's ubiquity in research laboratories and fields such as sensing and energy-storage, the visual inspection of scan rate's influence on CV behavior remains the primary means of mechanism assignment. Reliance on manual inspection precludes any application in high-throughput systems, limits the usefulness of the technique for non-domain experts, and becomes difficult when cyclic voltammograms increase in complexity and noise.

Herein, we introduce an approach for automatic mechanism assignment that extends beyond single event CV analyses that we (Figure 1A) and others have demonstrated in recent years.[10, 11] In this work, our primary goal was to develop a technique for mechanism assignment and localization based on complex CV data (Figure 1B). Therefore, the primary action taken was to deduce the sets of parameters, boundary conditions, and equations that may be arbitrarily combined to generate cyclic voltammograms for each of the relevant mechanisms we aimed to identify within CV data. With the careful establishment of such sets, thousands of cyclic

voltammetry experiments were run *in silico*, providing the breadth of possible current-potential responses for each mechanism type while also ensuring a minimal degree of ambiguity such that mechanism assignment is possible. In this work, eight fundamental mechanisms[3] were directly considered (Figure 1C), expanding our previous work[11] to include reaction schemes indicative of heterogenous electrocatalysis (*T*), homogenous electrocatalysis (*EC'*), and reversible surface redox (SR) reactions. In many systems, it is not safe to assume that the $\{v, i(E)\}_n$ information will be constrained to represent one electrochemical mechanism. For example, one often encounters solvent windows or terminal hydrogen evolution reactions, where the *T* mechanism would appear alongside a redox event of more direct interest (e.g., CO_2 reduction), resulting in a familiar form of a multi-event CV.[12, 13]

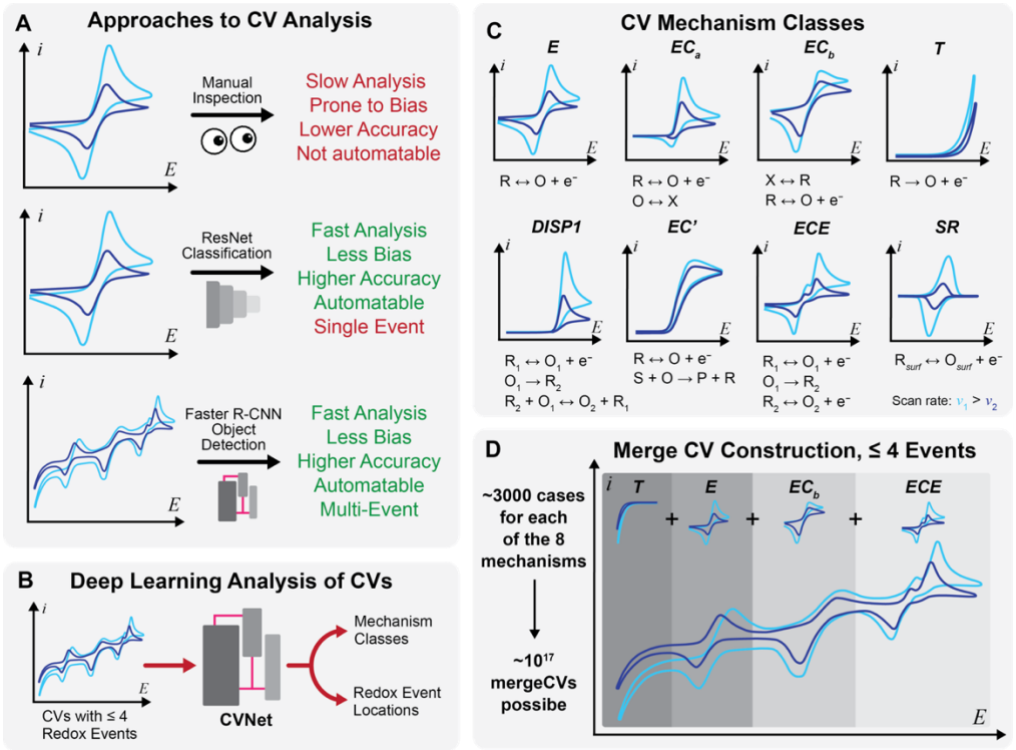
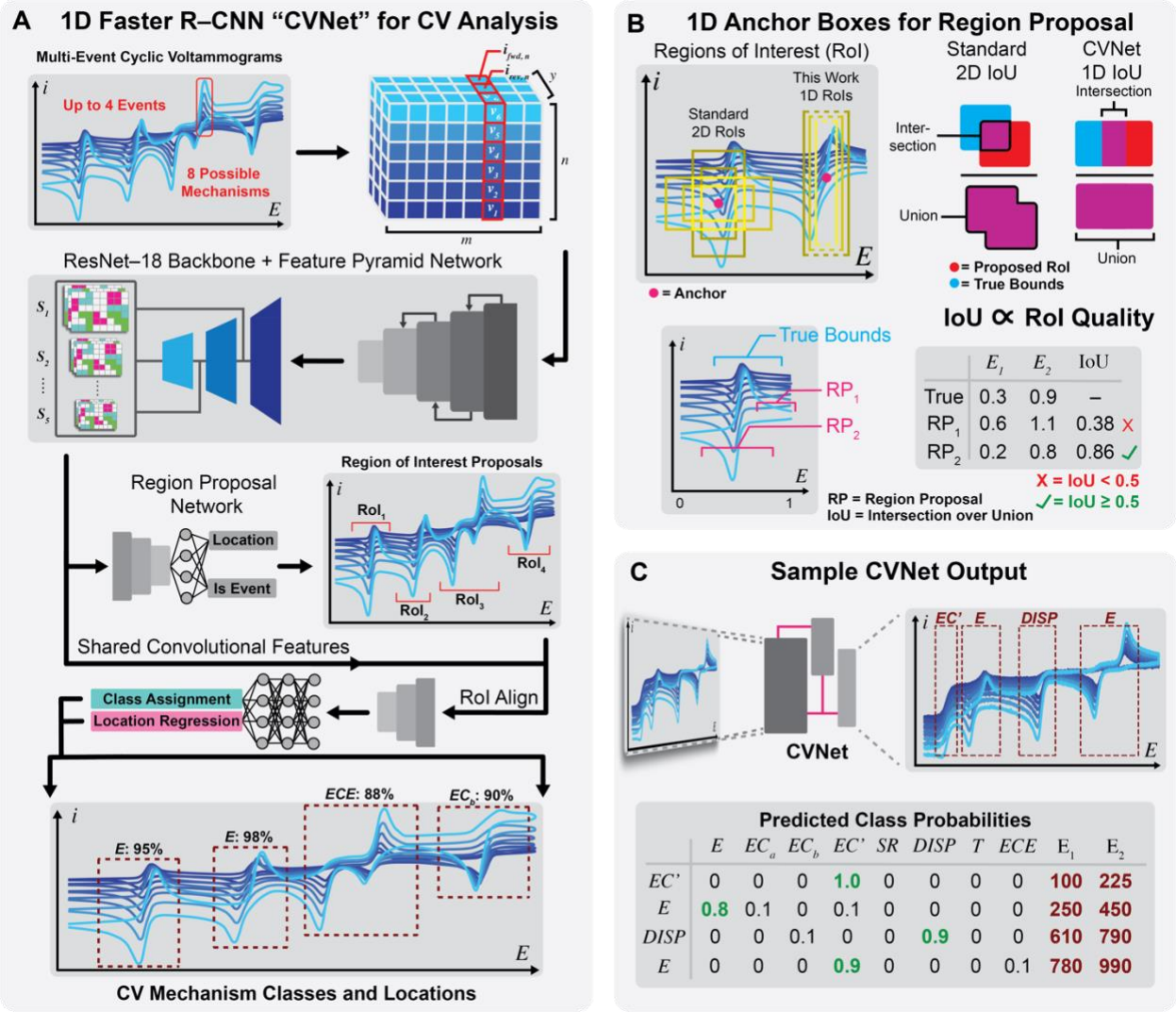


Figure 1. The evolution of CV analysis and the relevant electrochemical mechanisms considered in this work. **A** Qualitative comparisons between manual inspection versus our deep learning approaches to CV mechanism assignment. **B** Ideal qualitative output of deep learning mechanism assignment algorithm. **C**

The eight mechanism classes possibly represented in a complex CV. **D** Approximation of the numerical complexity of mechanism assignment in CVs with up to four independent electrochemical events and a simplified representation of the process of data generation.

The presence of multiple electrochemical mechanisms within a single cyclic voltammogram precludes the use of simple classification techniques as the number of possible classes with eight mechanisms and up to four mechanisms per CV would be over 4500 (Figure 1D). Thankfully, object detection algorithms may be considered as a mature technology for the elucidation of electrochemical mechanisms contributing to a convoluted $\{v, i(E)\}_n$ output. One such algorithm, Faster R–CNN, is well known within the artificial intelligence community to excel at classifying and localizing an arbitrary number of objects within an image.[14] Improving on frameworks with sluggish region proposal techniques, Faster R–CNN provided an online region proposal network (RPN), which greatly accelerated time to prediction; further, feature pyramid networks[15] for better object scale tolerance and improvements to algorithms that align the RPN's proposed regions of interest (RoI) have resulted in yet more powerful object detection architectures (Figure 2A). In this work, we have employed these tools along with the custom implementation of a 1D region proposal algorithm (Figure 2B) to obtain a highly effective means of CV mechanism enumeration from complex CV data. In addition to mechanism assignment, CVNet/Faster R–CNN must provide accurate localizations of voltage windows within a CV that contain redox events representing our eight mechanism classes (Figure 2C).



	E	EC_a	EC_b	EC'	SR	DISP	T	ECE	E_1	E_2
EC'	0	0	0	1.0	0	0	0	0	100	225
E	0.8	0.1	0	0.1	0	0	0	0	250	450
DISP	0	0	0.1	0	0	0.9	0	0	610	790
E	0	0	0	0.9	0	0	0	0.1	780	990

	E_1	E_2	IoU
True	0.3	0.9	–
RP_1	0.6	1.1	0.38 X
RP_2	0.2	0.8	0.86 ✓

Figure 2. Deep learning approach to multi-event CV analysis. **A** General structure of the Faster R-CNN/CVNet architecture, multi-event CV data is first evaluated to elucidate possible voltage ranges that contain mechanisms with subsequent classification and localization. **B** Comparison of region proposal in standard Faster R-CNN versus CVNet. One dimensional bounding boxes, or ranges, are calculated using the anchor box approach and evaluated for quality based on intersection over union (IoU). **C** Cartoon of terminal CVNet evaluation metrics. Match rate (MR) relates the percent of all electrochemical events confidently detected and classified, MR over threshold additionally requires correct classifications, and classification accuracy requires only that the max prediction is correct regardless of magnitude of prediction probability.

Herein, we discuss the initial results in the development of a custom 1D Faster R–CNN framework, CVNet. We foresee CVNet and its later evolutions as the foundation for powerful electroanalytical tools that enable the use of CV for experimentalists of all levels of familiarity, and ultimately as a technology that will encourage the use of cyclic voltammetry for high throughput experimentation and automation.

Results

Data was simulated via COMSOL Multiphysics v5.5 and sanitized and further refined with python for the conversion of simulated CVs into complex multi-event CVs with up to four constituent mechanisms (Supplementary Discussion). The data generated in this manner was amenable to object detection tasks, with the data matrices defining CVs directly fed to the custom CVNet architecture for mechanism assignment and localization. As a CV is fed through the network, it is first analyzed by the region proposal network for 1D RoI proposals which indicate voltage windows likely to contain one of our mechanism classes. During training, data exposed to this part of the network ideally finds regions that correspond to voltage windows predefined in our training data, and correctly identifies an object by assigning a binary classification confidence, where a confidence over 0.7 indicates a strong likelihood that an object is contained within the proposed window. Further, overlapping proposals were suppressed via non-max suppression such that a diversity of region proposals was generated along the global voltage window, hopefully detecting all events present in the CV. Through the course of region proposal and object detection, a few outcomes were possible (Figure 3A). Region proposals can ultimately align with true redox bounds (object detection true positive, tp_1) or not (object detection false positive, fp_1), and regions where known true redox bounds were not detected were implicitly false negatives (fn). In this work

there is no delineation between object detection and overall inference false negatives and *fn* contributes equally to the evaluation of object detection and overall inference metrics. Following object detection, proposed regions of interest were classified, where the best overall bounds (mechanism confidence ≥ 0.7 and IoU near 1) were presented as final outputs. For the classification task, true positives were defined as correct final classifications with good overlap with ground truth bounds (tp_2) and false positives were defined as incorrect final classifications (fp_2). For the evaluation of overall inference performance, any fp_1 that ultimately was classified as background by the classification portion of the network as dropped. Evaluation metrics (Figure 3B) showed strong performance by the object detection network and overall inference performance by the trained CVNet model. An overall F_1 score of 0.936 indicated strong overall performance with high precision and recall, which was further strengthened by an average IoU near one, where one was a perfect overlap of predicted bounds with ground truth bounds. Class by class accuracies and errors were derived from the left confusion matrix in Figure 3C, where 39% of all fp_1 results received final classifications of background/null (φ) and were subsequently dropped before calculating overall F_1. Class by class confusion minus background considerations were presented in the right confusion matrix of Figure 3C, where only tp_1 cases were considered for drawing an overall classification accuracy. While this value of 97.2% is slightly misleading as *fn* and fp_1 values were dropped from consideration, it did provide a valuable means of evaluating performance on an interpretable mechanism versus mechanism basis (e.g., if a true *DISP* event is present in the CV, it is clear to see which true mechanisms it is most likely confused with).

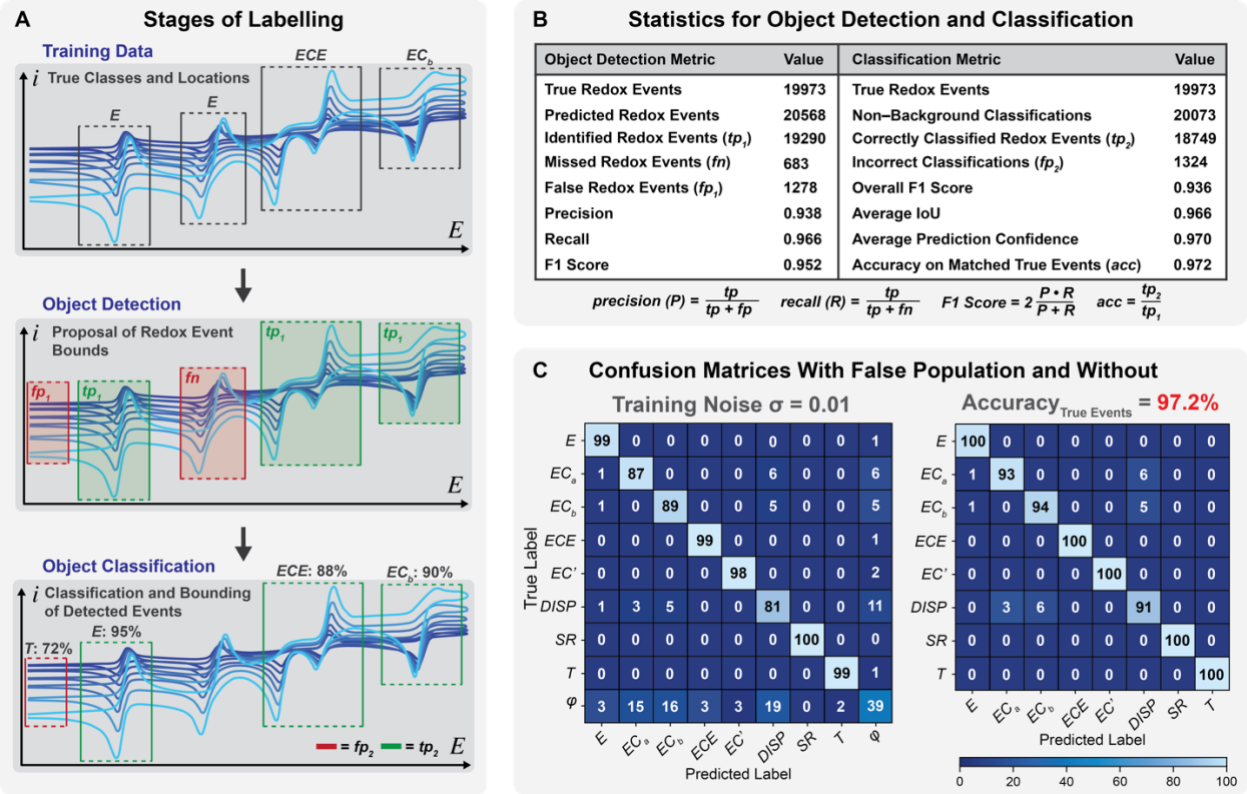


Figure 3. Data labelling throughout the inference process and types of labels assigned. **A** Labelling from training through final classification and types of negatives and positives assigned to events, the overall goal is to align predicted bounding intervals and mechanisms with those in the training data. **B** Various metrics associated with performance of the object detection and classification tasks where classification performance is downstream from object detection. **C** Confusion matrices for data trained and tested on $\sigma = 0.01$, on the left matrix, false negatives are presented in the far-right column and false positives in the bottom row. Object detection false positives (φ) that receive a final classification of φ are dropped from calculations. The matrix on the right provides performance evaluation as a function of classifications on the tp_1 sample.

Model inference was conducted with cases containing up to four events per CV (Figure 4A–D). While the model did perform well with all degrees of complexity, there was a slight decay in performance as the complexity of the CV increased as measured by number of constituent mechanisms. In terms of F_1 score as a function of events (n), the overall F_1 score decayed as follows {n:F_1 → 1:0.983, 2:0.963, 3:0.951, 4:0.937}. This decrease made sense, as the contribution of events with two neighbors (m) increases as n increases which was also shown to have a negative impact on F_1 {m:F_1 → 0:0.983, 1:0.934, 2:0.926}. Overall, however, the model performed both the mechanisms assignment and regression/localization tasks with high fidelity.

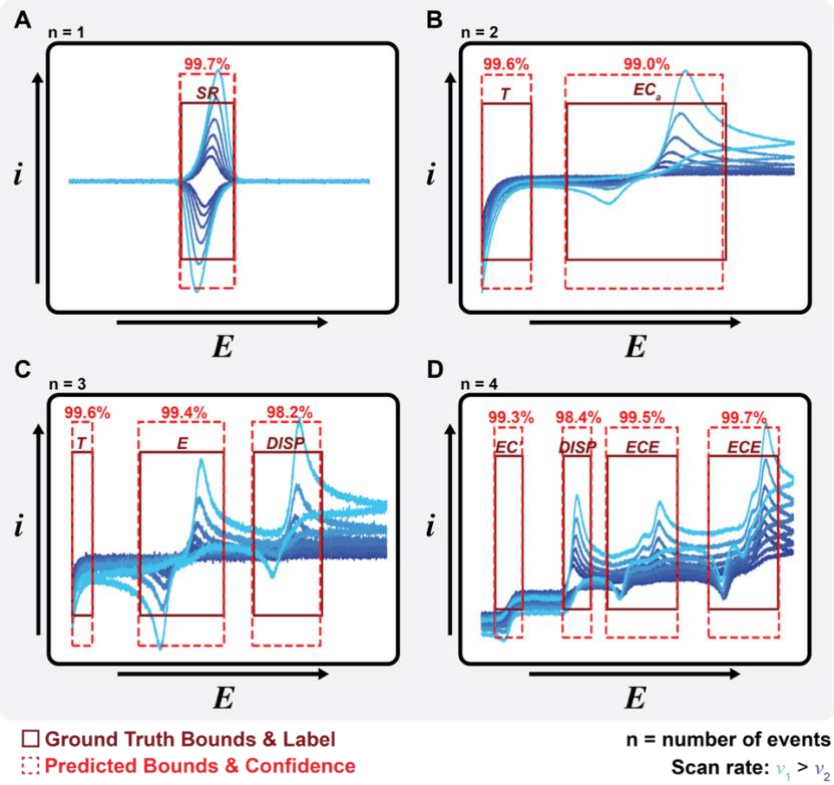


Figure 4. Inference results on CVs of various complexity encompassing all eight possible classes. **A** Ground truth bounding interval and voltage window shown in maroon and the predicted bounding interval and classification confidence shown in red for a single event CV representing the *SR* mechanism. **B** Results on a two event, **C** three event, **D** and four event CV.

Discussion

While the general performance of the model was sufficient, there were evaluations of interest beyond just simple performance. One key issue in cyclic voltammetry and electrochemical measurements in general is noise (Figure 5A). Measurement noise can greatly complicate the issue of mechanism assignment, but CVNet can learn to handle different levels of noise depending on its exposure to that noise during training, with some tolerance to noise outside what was directly applied in training. In general, CVNet models performed well on data containing noise up to the degree of noise on which they were trained (Figure 5A–F). The key takeaway inferred from the analyses of performance metric versus test noise curves is that no one model can be a catch all, and a model should be selected that was trained on data containing cases with noise relatively similar that of the data to be analyzed. High noise trained models can show fairly high F_1 scores and perform better than humans would as the noise levels become extreme (Figure S1), but performance on lower noise samples was somewhat diminished and the prevalence of false positives and false negatives in model predictions grew, although most object detection false positives (fp_1) were ultimately classified as the background class, which were dropped before final model results are presented and were not ultimately as damaging to model reliability as is implied by the pure (fp_1) and (fn) curves (Figure 5B–D). Ultimately, if experimental noise is low, a model trained on lower noise will have the best overall performance and should thus be used, but if noise is unavoidable in experiment, then higher noise models can provide a fairly high degree of performance that may be of value to experimentalists trying to extract information from noisy systems. As a gauge of true performance as noise grows, Figures 5E–F show that overall F_1 scores were more robust as their extreme training noise exposure grew, and impressively, IoU values stayed over 0.8 on average regardless of the combination of training noise and test noise, indicating

the regressed voltage windows containing redox events are reliable regardless of the intensity of measured noise.

In addition to noise, two other key aspects of the training data were of interest. When generating the training set, all individual CVs (which were combined to make the complex CVs) were originally scaled to have a possible global max current of 1 (unitless) and a possible minimum current of –1 (unitless). To account for this when merging single-event CVs to make complex CVs, each case was assigned a current scaling factor from (0.2–1.0) to ensure a more accurate representation of current responses as not every electrochemical event within a complex CV would be expected to be equally prominent. When this scaling factor was investigated for impact on model performance, it was seen that as noise in testing data increased and the weaker the current response of an event became, the more difficult it was to detect; however, as the noise reached more and more extreme values, the average/median values of missed redox events, as a function of scaling factor, converged on the global average scaling factor of 0.6, indicating that the data had reached a level of noise where all events were equally difficult to detect (Figure 5G). Besides scaling, the proximity of the next redox event (for $n \geq 2$ mechanisms per CV) would also be expected to be a challenge for classification, as two proximal events were likely to muddle and become difficult to analyze. For CVs with $n > 1$, the average distance to the next redox event was 443 mV. Box plots (Figure S2) show no relationship between errors as a function of nearest event and a dependence on test noise but do show a consistently lower average nearest events in the error sets, indicating that multiple events occurring in small voltage windows are more difficult to handle. In terms of classification performance, scaling and nearest event characteristics did not show any obvious interpretable trends that would be different from those mentioned for object detection (Figure S3). Finally, scan rate sampling procedures were shown to have an influence on

overall performance. In Figure 5H, arithmetic sampling takes any arithmetic sequence of scan rates from the original sequence of six can rates, while guided sampling chooses scan rates on a more ideal basis so as to promote accuracy but was less likely to be reflective of existing experimental practices. Particularly of note was the strong difference in performance of scan rate count n=2, where starting with six scan rates $\{v_1, v_2, v_3, v_4, v_5, v_6\}$ any combination of two scan rates may be selected for with the arithmetic sampling routine, but v_1 and v_6 were chosen for the guided sampling routine. The near identical performance of n=2 and n=6 for the guided sampling routine strongly implies that, typically, only two well selected scan rates are necessary for mechanism determination with this approach.

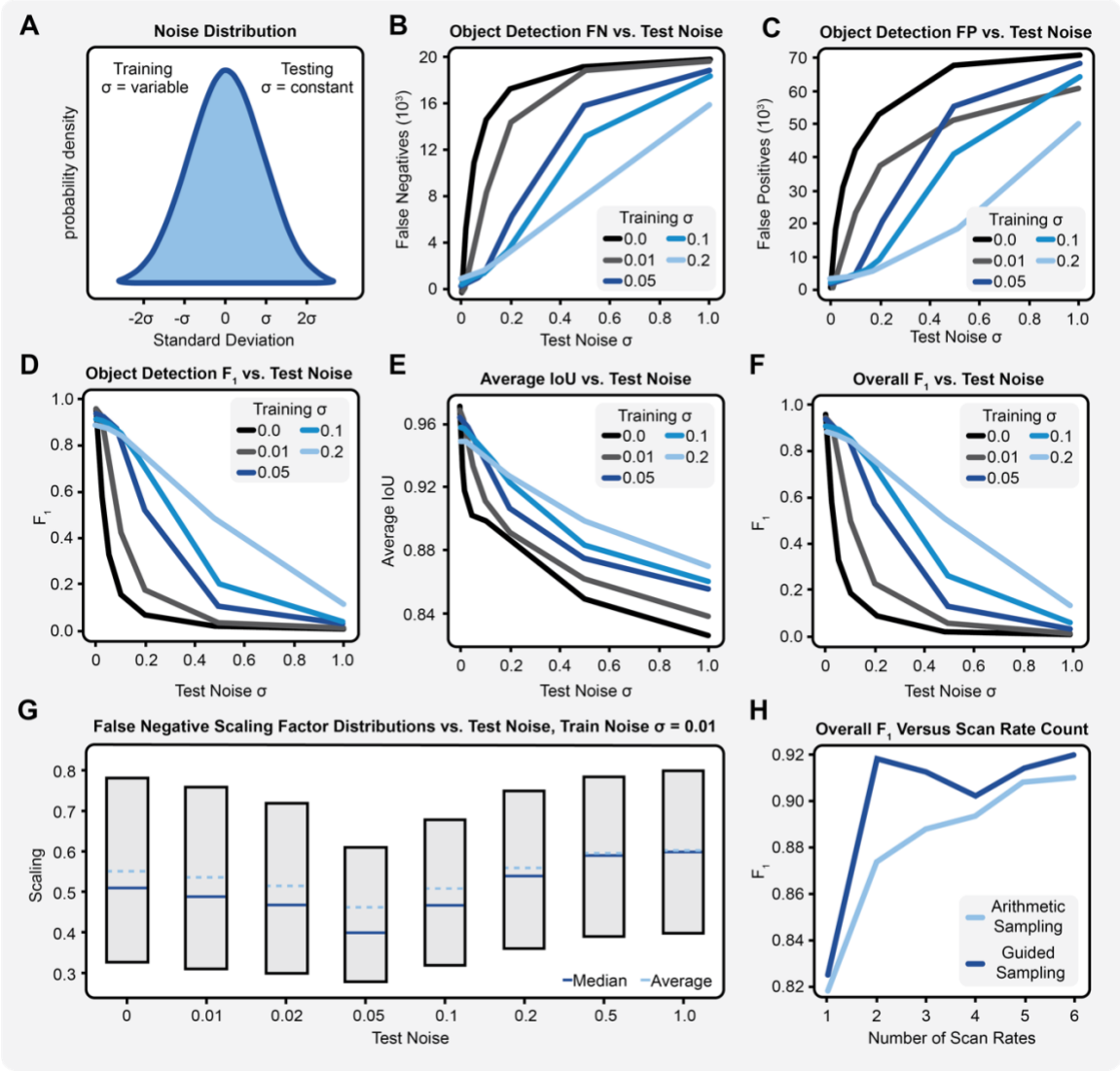


Figure 5. Impact of noise in testing data on metrics of interest to both object detection **B–D, G** and overall inference performance **E, F**. **A** Cartoon of noise distributions sampled from when adding noise to a CV. Training sets always contained data with variable noise, indicating that sigma could be any value from 0 to the value presented in the insets of the figures. Test noise used one pre-defined distribution to add noise to all samples, giving an even level of noise throughout the entire testing set. **B–F** The evolution of performance evaluation metrics with noise added to testing data. **G** Impact of scaling on the object detection tasks, 0.6 is the global scaling average. **H** Impact of scan rate sampling protocol on overall inference performance. Data using training and test noise factor $\sigma = 0.01$.

Continuing Work

While the model is trained and performing well on simulated data, some key investigations need to be done before the project is finished. First, an evaluation of information density within training data would be a valuable insight to study. By perturbing regions of the input data and measuring influence on model confidence, we could get a sense of which regions of data are most important for making inference decisions. This is valuable to obtain some sense of why decisions are made within CVNet, as deep learning algorithms are famously difficult to interpret. Finally, and most importantly, performance of the model on experimental data must be investigated for the tool to have any acceptance by the experimental research community. Performance on training data, especially when simulated, will likely be exaggerated due to the lack of noise and non-idealities found in real systems that were not directly accounted for in the training data.

References

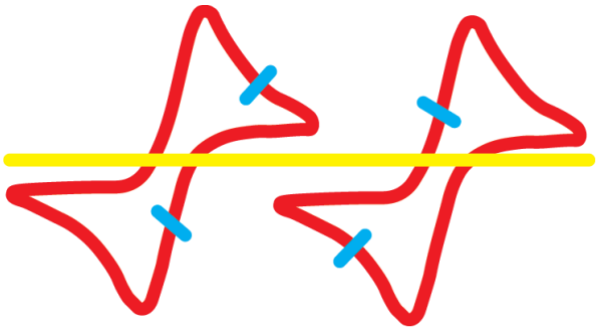
1. Nicholson, R. S., Theory and Application of Cyclic Voltammetry for Measurement of Electrode Reaction Kinetics. *Analytical Chemistry* **1965,** *37* (11), 1351-1355.
2. Allen J. Bard, L. R. F., *Electrochemical Methods Fundamentals and Applications*. 2 ed.; John Wiley & Sons, INC: 2001.
3. Costentin, J.-M. S. C., *Elements of Molecular and Biomolecular Electrochemistry: An Electrochemical Approach to Electron Transfer Chemistry*. 2006; Vol. 2.
4. Eliaz, E. G. N., *Physical Electrochemistry: Fundamentals, Techniques, and Applications*. 2019; Vol. 2.
5. Elgrishi, N.; Rountree, K. J.; McCarthy, B. D.; Rountree, E. S.; Eisenhart, T. T.; Dempsey, J. L., A Practical Beginner's Guide to Cyclic Voltammetry. *Journal of Chemical Education* **2018,** *95* (2), 197-206.
6. Homer, M. K.; Kuo, D.-Y.; Dou, F. Y.; Cossairt, B. M., Photoinduced Charge Transfer from Quantum Dots Measured by Cyclic Voltammetry. *Journal of the American Chemical Society* **2022,** *144* (31), 14226-14234.
7. Hogan, C. F.; Bond, A. M.; Myland, J. C.; Oldham, K. B., Facile Analysis of EC Cyclic Voltammograms. *Analytical Chemistry* **2004,** *76* (8), 2256-2260.
8. Elgrishi, N.; Kurtz, D. A.; Dempsey, J. L., Reaction Parameters Influencing Cobalt Hydride Formation Kinetics: Implications for Benchmarking H2-Evolution Catalysts. *Journal of the American Chemical Society* **2017,** *139* (1), 239-244.
9. Savéant, J. M.; Xu, F., First- and second-order chemical-electrochemical mechanisms: Extraction of standard potential, equilibrium and rate constants from linear sweep

voltammetric curves. *Journal of Electroanalytical Chemistry and Interfacial Electrochemistry* **1986,** *208* (2), 197-217.

10. Kennedy, G. F.; Zhang, J.; Bond, A. M., Automatically Identifying Electrode Reaction Mechanisms Using Deep Neural Networks. *Analytical Chemistry* **2019,** *91* (19), 12220-12227.
11. Hoar, B. B.; Zhang, W.; Xu, S.; Deeba, R.; Costentin, C.; Gu, Q.; Liu, C., Electrochemical Mechanistic Analysis from Cyclic Voltammograms Based on Deep Learning. *ACS Measurement Science Au* **2022,** *2* (6), 595-604.
12. Wuttig, A.; Yaguchi, M.; Motobayashi, K.; Osawa, M.; Surendranath, Y., Inhibited proton transfer enhances Au-catalyzed CO2-to-fuels selectivity. *Proceedings of the National Academy of Sciences* **2016,** *113* (32), E4585-E4593.
13. de Ruiter, J.; An, H.; Wu, L.; Gijsberg, Z.; Yang, S.; Hartman, T.; Weckhuysen, B. M.; van der Stam, W., Probing the Dynamics of Low-Overpotential CO2-to-CO Activation on Copper Electrodes with Time-Resolved Raman Spectroscopy. *Journal of the American Chemical Society* **2022,** *144* (33), 15047-15058.
14. Ren, S.; He, K.; Girshick, R.; Sun, J., Faster R–CNN: Towards Real–Time Object Detection with Region Proposal Networks. *arXiv:1506.01497* **2015**.
15. Lin, T. Y.; Dollár, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S., Feature Pyramid Networks for Object Detection. *arXiv:1612.03144* **2016**.

Supplementary Discussion

Generation of Complex Cyclic Voltammograms: For each of the eight possible electrochemical mechanisms, datasets of ~5000 cyclic voltammetry (CV) traces were generated. These were used as the basis of the training set used in this work, where these CVs were combined together in a semi-arbitrary fashion to give a wide array of complex CV current potential responses. When combining CV traces, a few rules were established. First, the *T* and *EC'* mechanisms were constrained to the edges of the voltammograms. Beyond that, considerations primarily focused on the establishment of rules ensuring reasonable merging of voltammograms. For a CV to be merged from a side, say from the right (more positive potential), the current must reach an approximate steady state such that the current changes by no more than a relative value of 0.05 over the last 50 mV of the voltage window. In addition, in order to maintain the most defining characteristics of a CV trace, the ¼ peak heights of the current responses were calculated and the relevant vol tages were used to define minimum proximity to the host event that another redox event could occur. Visually, this is represented in the inset figure, where if the two events were to be merged, the closest the two upper peaks could be is at the alignment of the far right blue tick of the left trace and the far left blue tick of the right trace. This rule was intended to preserve the identifiability of each mechanism in a complex CV while also providing a reasonable level of redox event proximity so as to not make identification of sub-events trivial.



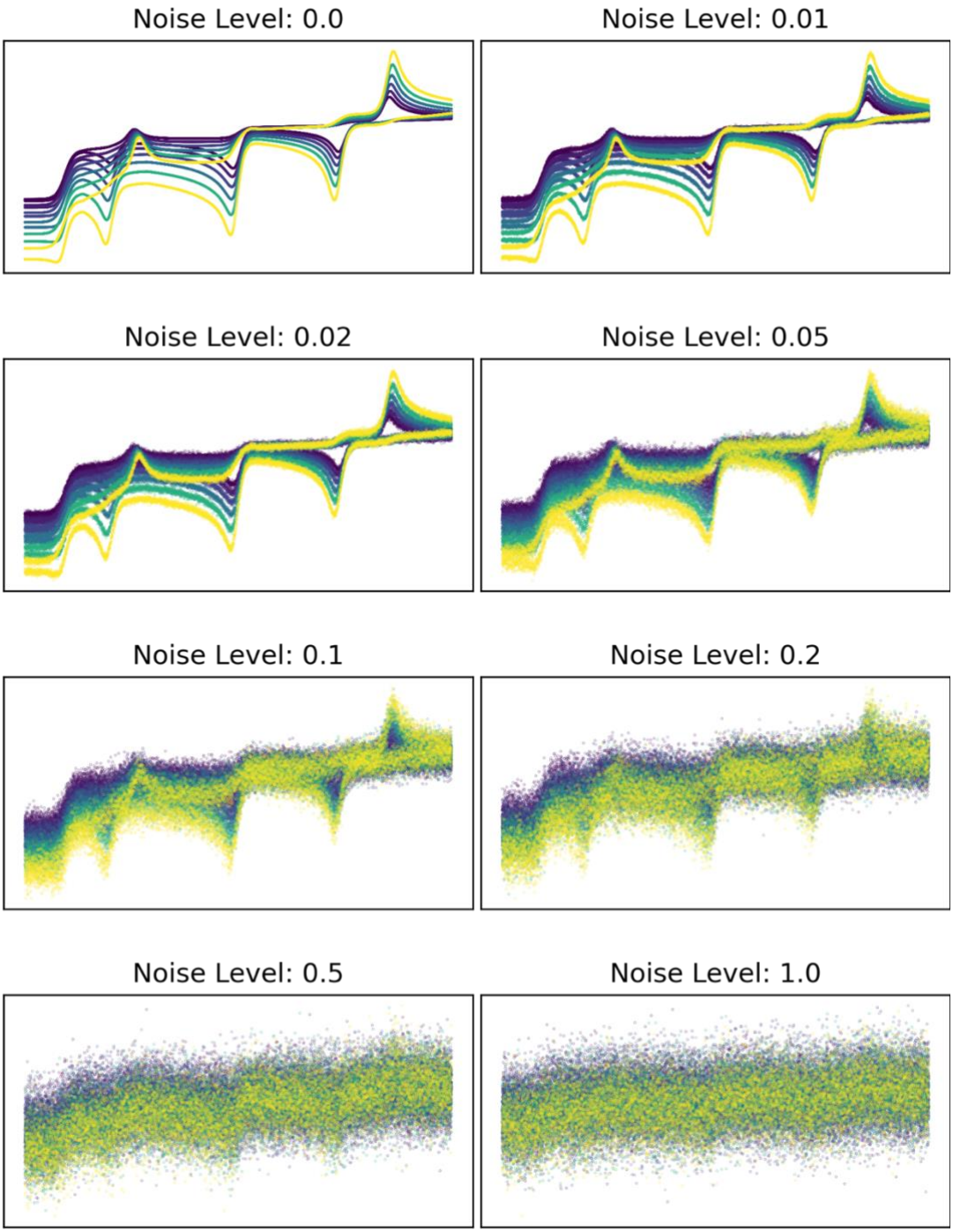


Figure S1. Examples of noise levels applied to CVs, value in title of each subplot is standard deviation of normal distribution from which noise was sampled

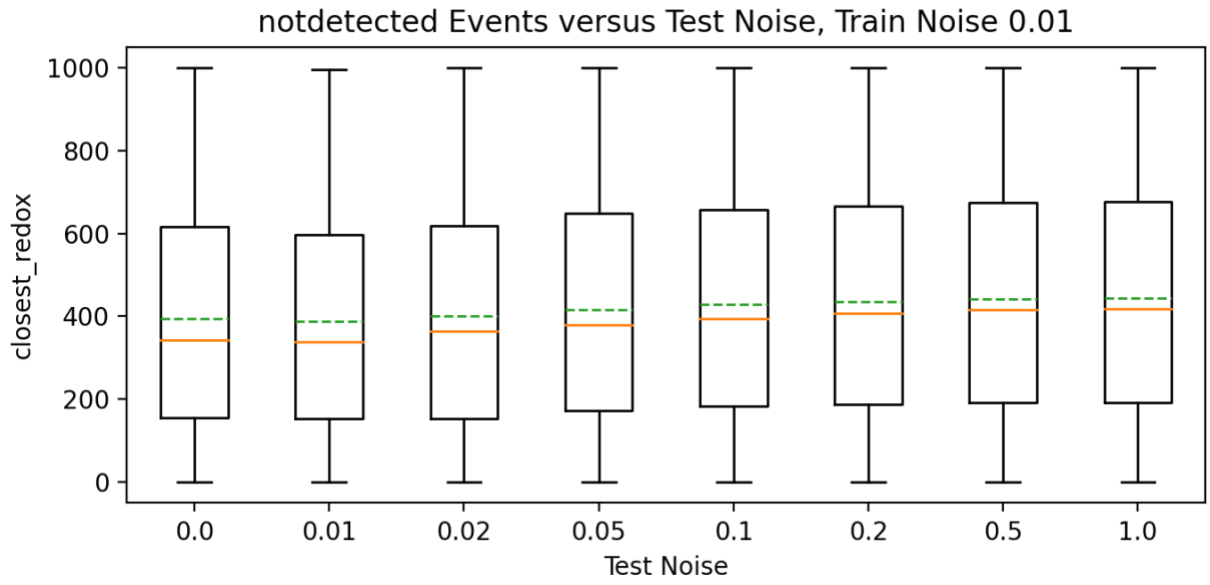
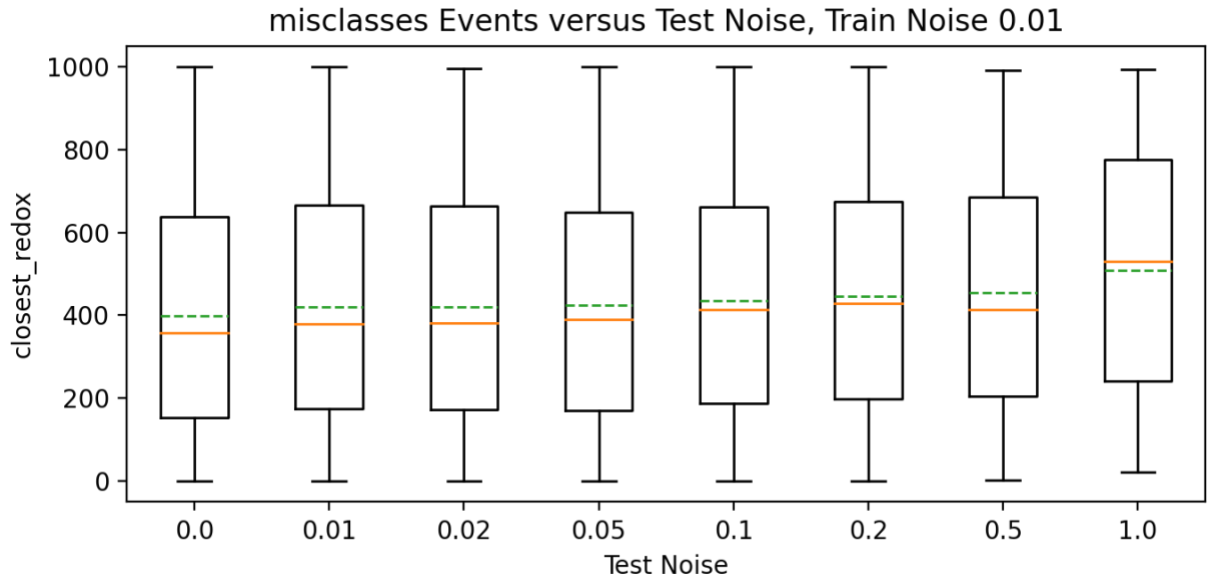


Figure S2. Box plot demonstrating influence of nearest redox event distance on CVNet performance versus test noise, average nearest event (green) and median nearest event (orange), global average 443 mV

A



B

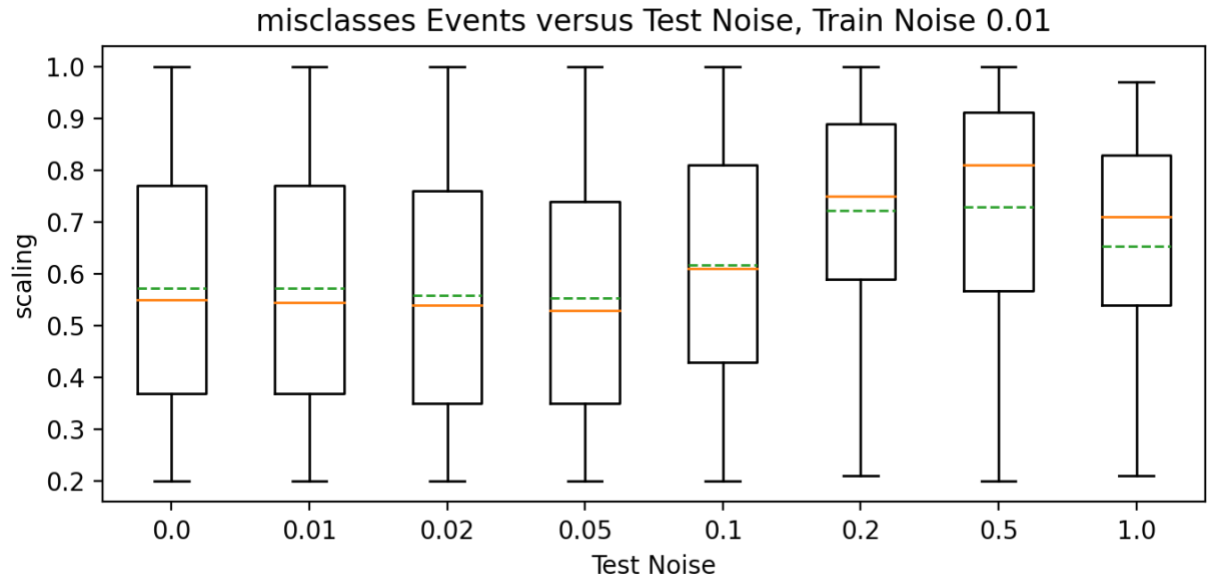


Figure S3. Impact of nearest redox event distance (A) and scaling (B) on classification, increases of average (green) and median (orange) at high test noise in panel (B) are a function of the inability of the network to detect events with weak current response due to their obfuscation by noise, therefore increasing the average scaling of misclassified events indirectly

Chapter 6. Sentiment analysis for student course reviews, an exploration of non-electrochemical data

This chapter is a version of submitted work of Hoar, B. B; Ramachandran, R.; Levis-Fitzgerald, M.; Sparck, E. M.; Wu, K.; Liu, C. "Enhancing the Value of Large-Enrollment Course Evaluation Data Using Sentiment Analsysis" *Journal of Chemical Education (in review)*

Abstract

In education, space exists for a tool that valorizes generic student course evaluation formats by organizing and recapitulating students' views on the pedagogical practices to which they are exposed. Often, student opinions about a course are gathered using a general comment section that does not solicit feedback concerning specific course components. Herein, we show a novel approach to summarizing and organizing students' opinions as a function of the language used in their course evaluations, specifically focusing on developing software that outputs actionable, specific feedback about course components in large-enrollment STEM contexts. Our approach augments existing course review formats, which rely heavily on unstructured text data, with a tool built from Python, LaTeX, and Google's Natural Language API. The result is quantitative, summative sentiment analysis reports that have general and component-specific sections, aiming to address some of the challenges faced by educators when teaching large physical science courses.

Introduction

Student course evaluations are a fundamental way in which students participate in their own education.[1, 2] By providing feedback at the end of a course, students allow instructors to reflect on their strengths and weaknesses, hopefully enhancing course learning experiences for both educators and students over time. One limitation of this approach, however, is that students are

often asked to provide feedback in an unstructured, open-ended manner.[3, 4] This approach is limiting in that it does not organize data into an easily digestible format that can be quickly analyzed by educators. Furthermore, faculty cannot gain quick insight on specific components of their course, such as lectures, homework assignments, exams, activities, discussions, etc., without personally parsing all comments. This issue is exacerbated in large-enrollment courses that cater to several hundred students at a time. In fact, it is typical for instructional faculty teaching undergraduate physical science courses at large research universities that receive upwards of a thousand course evaluations per year.[5] When this feedback data is presented as unstructured text, patterns in student sentiment may be missed, ignored, or falsely derived from biases in the mind of the reader.[6, 7] Because of this, a tool to organize and present student reviews in a way that provides both general and component-specific insight would be useful, especially in large-enrollment contexts. A combination of qualitative and quantitative data regarding course reviews would provide structure for iterative and trackable updates to pedagogical practices, facilitating positive course evolution.

One such way to accomplish this goal is via the development of software that leverages machine learning. Software can provide the means to algorithmically parse and clean unstructured data and machine learning can be utilized to score data on a "sentiment scale" that ranks phrases and sentences on a numerical one to five, negative to positive sentiment scale.[8, 9] Such evaluation of broad sentiment can help educators address long-standing issues in large-enrollment contexts, where evaluations of professors tend to decrease with class size[10, 11] while students report lower satisfaction.[12] In the following text, discussion will focus on how Python programming is combined with the Google Cloud Sentiment Analysis API (GCSA), as well as how text data was

gathered and used to train the GCSA machine learning model. Finally, educator feedback will be presented to gauge this tool's efficacy in its current state.

Background and Relevance

Machine learning is defined as "the use and development of computer systems that are able to learn and adapt without following explicit instructions, by using algorithms and statistical models to analyze and draw inferences from patterns in data."[13] In essence, machine learning algorithms can be trained on known data to predict a desired metric on arbitrary future data. A subset of machine learning, sentiment analysis, aims to learn how the vocabulary and language employed by an author correlates to positive, negative, or neutral opinions.[8] This technique is ubiquitous in companies aiming to gauge public opinion on their products or marketing strategies.[14] Often, these opinions are gathered from public platforms such as Twitter or Instagram and subsequently analyzed using sentiment analysis algorithms. Once opinions are gathered and scored, easy to interpret graphs and tables can be used to succinctly demonstrate how well a company's product, advertisements, etc. are being received by the public. One well known case within this context is the analysis of IMDb movie reviews, a source of highly opinionated text data that is ubiquitous in algorithm benchmarking and student learning.[15-17] Recently, sentiment analysis has aided with the analysis of politicians' communications about the COVID-19 pandemic and how their tone impacts the efficacy of messaging and information spread.[18] Finally, sentiment analysis has seen application in education contexts as well, having been used in the analysis of student feedback emotionality[19] and in classifying reactions to E-learning practices as positive or negative.[20] Given the success of sentiment analysis techniques in these contexts, the investigation of sentiment analysis for large-enrollment courses seems prudent as they are historically challenging to teach.[10] Further, existing issues are exacerbated by modern challenges with the

increasing diversity of educational backgrounds in course enrollees[21] and growing reliance on virtual communication of course material.[22, 23] Therefore, the demand for rapid, succinct, actionable feedback for educators in order to ensure that their classrooms provide effective and inclusive learning experiences is increasing.

To implement these techniques, various paths can be taken. These options include training and engineering a machine learning algorithm from scratch, using a pre-trained sentiment analysis algorithm, or using a service such as GCSA. In this paper, both pre-trained algorithms and service options are employed as they are the most flexible and "easy" to implement for researchers wishing to employ these techniques with only a minor background in machine learning – making the techniques reported here highly accessible and adaptable.

With the accessibility of both pre-trained algorithms and GCSA in mind, the first step in implementing these techniques was to gather and score data representative of student course evaluations, initially with an emphasis on UCLA large-enrollment STEM courses.[24] By gathering and labelling a representative dataset, a GCSA model can learn to predict student sentiment scores and enable the production of sentiment analysis reports[25, 26] (Figure 1).

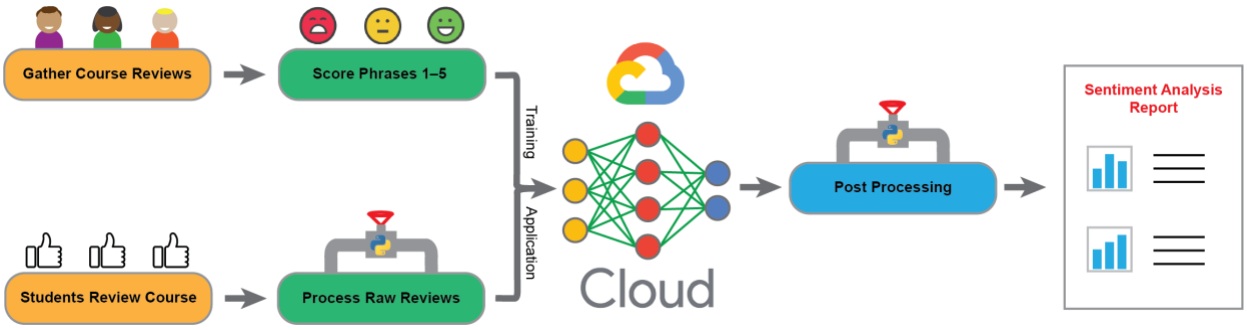


Figure 1. Overview of training and general application pipeline. In the algorithm training phase, course reviews are gathered, scored, and used to train a GCSA algorithm which can then be subsequently used to generate a sentiment analysis report. After training, student course reviews can be processed and scored by the GCSA algorithm alone.

In the case of this research, the input was student course reviews in the form of statements or sentences and the output was the relative positivity of these statements on a scale of 1–5. To gather this data, nine UCLA professors and five teaching assistants were solicited to provide student reviews of their courses from previous quarters. Initially, student course reviews were pulled from lower-divisional undergraduate courses in the physical sciences, with class sizes on the order of one hundred to three hundred students. Data sourced from professors were a valuable source of lecture and general course structure feedback, while teaching assistants' feedback highlighted student opinions on highly interactive environments, such as lab classes and course discussions. Typically, UCLA lab and discussion sections contain around twenty to thirty students each. While feedback provided to teaching assistants reflects smaller class environments, discussions and labs are affiliate courses to the full lectures and the feedback is indicative of the large-enrollment context. Furthermore, data related to student sentiment in one context is not independent of feedback sentiment at large, and more general training data provides a more robustly trained machine learning algorithm.[27] Due to the generally positive nature of UCLA student course reviews, *ratemyprofessor.com* was also consulted to extract more negative and neutral reviews to balance our training data. In acquiring data from this public source, the primary emphasis was the collection of data that directly discussed course contents and educator quality, while remaining largely agnostic to the course subject and thus pulling data evaluating college courses in general. In the end, this resulted in a dataset of 1,603 phrases (Figure S1). To minimize bias in the scoring of the dataset (initial data collection and approximate scoring was performed by Benjamin Hoar), a total of 20 UCLA undergraduate students were gathered to perform additional scoring, a process that required approximately four and a half hours of scoring time per student assuming a rate of one statement scored per ten seconds. Directions on scoring stated that

scoring should be conducted as objectively as possible[28] – meaning that the language should be used as the basis for scoring and not personal biases, as best as possible. In the end, the median of the total of 20 scores was counted as the *true* value of the phrase (Supplementary Methods). After scoring, this data was split into sets containing 1,282 training, 161 validation, and 160 testing instances. This split in data was necessary to select the best algorithm structure, with the training data used to extract correlations in the data and output, the validation set used to test the accuracy of the model *during* training, and the testing set used to test the accuracy of the model *after* training.[14] Using this split, the GCSA tool provided a trained sentiment analysis algorithm with an overall accuracy of 73.1% (Figure S2), in line with accuracies for modern fine-grained (i.e. non-binary) sentiment analysis approaches.[9, 29]

Following training, the sentiment analysis algorithm could be called from custom python scripts developed specifically for student course review supplementation. These python scripts provided the means to convert any given raw text source into the format required for input into the scoring algorithm and subsequently its organization into a summative report (Supplementary Report) for instructor review. The first stage of this report generation involved the splitting of raw text data into phrases. This was accomplished by splitting on standard sentence-ending punctuation and the word "but." Following this, each statement was converted to a standard form and then scored by the GCSA algorithm and saved for subsequent report generation. The last step prior to report generation was selection of terms to display and calculation of summary statistics to be presented. Scored statements returned from the GCSA algorithm were filtered for terms that were of common interest (selected by the authors) and those that were of interest to each specific report (selected by a word-cloud inspired algorithm). The word-cloud algorithm provided the most common terms mentioned by students minus a large list of meaningless words (often called stop

words) such as “and,” “the,” etc. Once all scores were provided and filtered, they were organized into summative and component-specific figures and tables to provide instructors an overview of student sentiment related to the class in general and to specific class components.

General Results Section

The “general” results section provided an overview of the entire text corpus and reflected student sentiment as a function of all student review statements. Here, four graphs were presented (Figure 2).

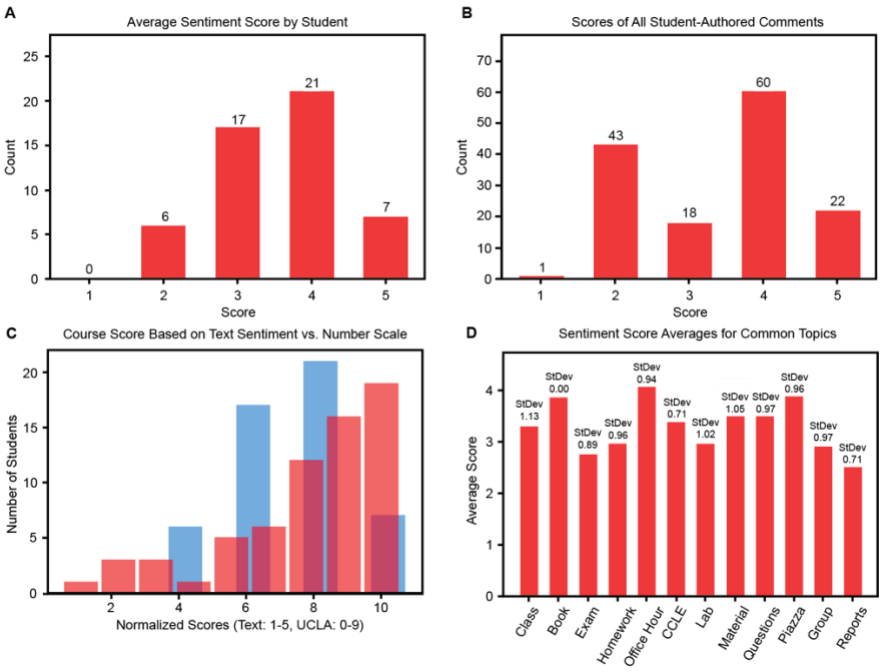


Figure 2. Overall course sentiment and average sentiment for selected course concepts. **A** Distribution of average sentiment of students {average: 3.57, std. dev.: 0.87}, in the figure 17 students had a neutral average sentiment (score 3). **B** All scores of all phrases independent of author in text corpus {average: 3.41, std. dev.: 1.09}. **C** Comparison of student sentiment as a function of sentiment analysis performed on their written feedback (blue, same distribution as A) versus their numerical rating of the course on a 0-9 scale (red). **D** Average and standard deviation of student sentiment derived from phrases related to the terms shown along the x-axis.

The first two of the four graphs presented a summary of student sentiment both as a function of student sentiment normalized *by author* (Figure 2A) and a second showing *all* statements' sentiment regardless of author (Figure 2B). The key purpose of figure 2A was to remove bias resulting from a highly verbose author (e.g., an outlier student writing many negative reviews would skew the 2B graph, but not 2A). Succeeding these histograms were two more summative graphs. Students in UCLA physical science courses were asked to rate the course on a scale of 0-9, giving an overall final score for their experience in their respective courses. In Figure 2C, the distribution of these scores was plotted versus the same distribution as presented in 2A, giving two ways of viewing student sentiment – one derived from a simple numerical scale and one derived from their written opinions. Finally, figure 2D encompassed the component-specific results section. In figure 2D, the average and standard deviation of scores related to a course component were presented as a prelude to the component-specific section, which would present a more direct analysis of each of these terms.

Component Specific Results Section

In addition to these summative reports, component-specific reports were also generated to provide insight into how specific course practices were received. Two methods of generating these reports were considered. First, via a discussion between UCLA faculty and advisors from the UCLA Center for the Advancement of Teaching, a list of globally relevant terms that encompassed course aspects common to most, if not all, university level physical science courses was established. The terms covered the instructor, lecture, textbook, exams, homework, CCLE (an announcement, grade, and document student portal), and office hours. In addition to these terms, six auto-selected terms were also used to generate component-specific reports. These terms were selected via a word-cloud inspired algorithm (Figure 3A).[30] In a word cloud, the most common

words in a body of text are presented in a graphic, with the most common words appearing larger. To select for these terms, the counts of all words of the entire student-review text corpus were calculated. From this list of word-counts, words already accounted for in the aforementioned component reports and meaningless "stop-words"[31] (words such as the, if, and, etc.) were also removed from this text body. What was left over was a list of the most referenced terms. Common auto-selected terms included "lab," "chemistry," "material," and "organic." Niche terms such as "Piazza" (a non-ubiquitous student Q&A forum), "recorded," (referring to recorded lectures) and "clicker" (referring to a remote students used to answer questions live in class) were also selected for their appropriate instructors. These auto-selected terms highlighted the value of this algorithm and general utility of the tool in general, as they can capture both subject-level (e.g., "chemistry") and class-level (e.g., "Piazza") terms to tailor the reports to their recipients. Using these terms, reports were generated for components most relevant to the specific course as dictated by the components of greatest interest to the students as implied by their reviews.

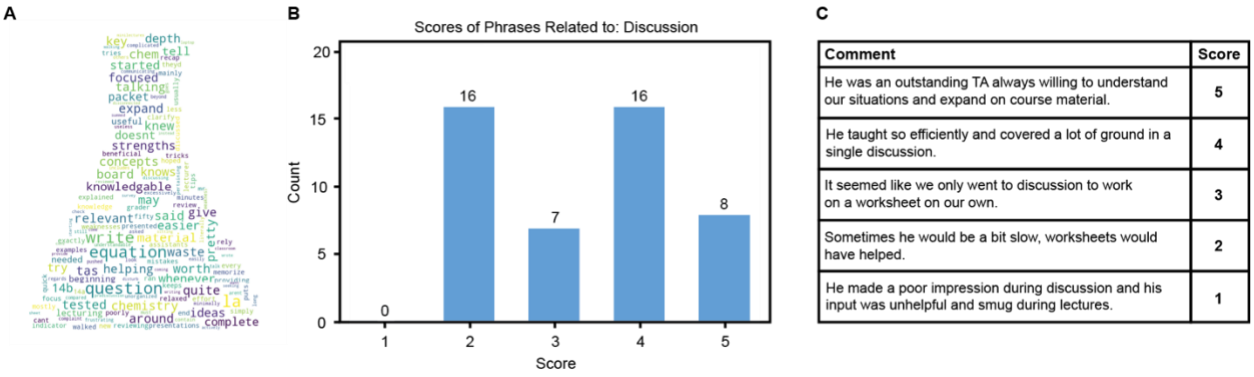


Figure 3. Visualization of techniques applied to the component-specific section of the report. **A** A word cloud visually representing the selection of prominent terms for automated selection of course feedback terms that were not included in the predetermined list of words. **B** Representative distribution of scores related to the report component "discussion." **C** Representative table of phrases derived from student feedback that represent the commentary of students pertaining to the "discussion" component – an exemplar for score 1 was sourced from a different report than **B** for completeness.

Considering these terms of interest, each component was given a page to demonstrate the overall sentiment of the students as it pertained to that component, as well as gave a subset of actual phrases from the students. On each of these pages a score distribution (e.g., Figure 3B) was presented atop a table (Figure 3C) of exemplary phrases – providing samples from each of the five scores of the ranking system. The phrases selected for presentation were further fed through a pre-trained, binary sentiment analysis algorithm, VADER (Valence Aware Dictionary and sEntiment Reasoner),[32] that provided secondary assurance that the presented scores were accurate representations of their classes, but was not involved in providing the ultimate sentiment score to any phrases. The desire for this was to mitigate the effect of errors stemming from the initial algorithm and provide quality control for the phrases presented in these component reports. Once phrases were selected, they offered the instructor insight into overall sentiment regarding a course component (Figure 3B) and showed specific statements that indicated strengths and weaknesses related to the execution of a course component (Figure 3C). Once all the data processing, scoring, and organization was accomplished, the data was automatically converted into a standalone report with the aid of pyLaTeX,[33] a python library that can be implemented automate the creation of LaTeX documents (Supplementary Report). Within this document, directions on how to interpret the data and additional insight into report generation were also included.

Effectiveness and Feedback

To gauge the utility of this proof-of-concept supplementary report, four UCLA professors provided voluntary feedback on the generated report via a survey (Supplementary Methods). This survey was developed by the UCLA Center for the Advancement of Teaching to garner anonymous feedback about the efficacy of this tool. In total, we received four responses to the opt-in survey, which precluded any statistical analysis of the tool's reception, but did provide insight

into initial opinion. In addition to their survey responses, two instructors provided longform commentary. Instructors unanimously agreed that the general format of this report was satisfactory, that the 5-point scale was sufficiently informative (a prior three-point scale was rejected and is not reported here), and that the report was a valuable complement to their analysis of student feedback. While no aspect of the report was unanimously unsatisfactory, some limitations were noted. Respondents noted that the selected keywords and report were not exhaustive enough, and both longform commenters wished to see a component report on "workload," and each provided additional terms such as "chemistry," "grading," and "collaboration" that they wished appeared in their reports. It is worth noting here, that had those concepts been popular among student feedback, the auto-selection algorithm would likely have captured them; however, there is room for flexibility in our practices that would preclude issues regarding instructor reliance on algorithm auto-selection and improve user trust in the tool. For instance, the auto-selection algorithm could be used as a complement to terms requested by instructors, rather than the only source of extra component-specific reports. Another suggestion provided was to provide more context for the scored phrases. In this iteration, student scores were parsed and isolated, removing them from their initial context. As was suggested, a solution in future iterations would be to present the scored phrases in context and highlight them. This would allow instructors to quickly scan their reviews due to the highlighting, and see the sentiment analysis in context, increasing the interpretability of the scored comments.

One significant practical update to this tool – one that could address all the concerns voiced about this iteration of the tool – would be to emancipate it from the standalone pdf output format (as with the Supplementary Report). A web-based app, for example, would enable rapid deployment of updates to the tool's function while increasing user autonomy by allowing easier

implementation of custom user queries. While a change in format may be valuable, this current approach does achieve our goals as indicated by survey feedback and overall sentiment prediction performance (Figure S2), and thus provides a worthy launching point to further augmentation to STEM based large-enrollment course evaluation practices with the aid of machine learning.

Conclusion

In short, a proof-of-concept tool has been developed to aid educators in their efforts to better understand and adapt to the sentiments of their students. By combining programming and machine learning principles, educators can now visualize how the language used by students in their course reviews relates to their feelings about course components in a dual quantitative and qualitative way. Furthermore, in current UCLA physical science course review practices, students only provide feedback in a "general comments" section, creating a barrier for educators aiming to analyze how well specific aspects of their course are being received. With this tool, educators can analyze graphical and tabular data on the opinions of their students with respect to the most common course components of interest. This format is especially valuable for large-enrollment STEM courses as pedagogical approaches need to be tuned to broad audiences while maintaining effectiveness. Further, the auto-selection of popular terms in feedback data allows for the presentation of course-specific reports that may not apply to every instructor but are invaluable to STEM instructors who employ a wide array of teaching aids to convey challenging subject matter. For example, the noted auto-selection of terms for reports related to both "chemistry" and class tools (e.g., clickers for live participation) can bolster the rapid evolution of teaching practices by highlighting the effectiveness of current approaches, guiding necessary changes. In recent times, rapidly changing demographics and teaching formats have provided an unprecedented challenge to educators who have faced diverse educational backgrounds, uncertainty related to the

effectiveness of new teaching practices (both locally and remotely), and huge numbers of students. Thankfully, machine learning approaches such as ours provide the opportunity for rapid course evolution and will hopefully improve education for both instructor and student.

References

1. Spooren, P.; Brockx, B.; Mortelmans, D., On the Validity of Student Evaluation of Teaching:The State of the Art. Review of Educational Research 2013, 83 (4), 598-642.
2. Zabaleta, F., The use and misuse of student evaluations of teaching. Teaching in Higher Education 2007, 12 (1), 55-76.
3. Mirończuk, M. M., Information Extraction System for Transforming Unstructured Text Data in Fire Reports into Structured Forms: A Polish Case Study. Fire Technology 2020, 56 (2), 545-581.
4. Li, Z.; Fan, Y.; Jiang, B.; Lei, T.; Liu, W., A survey on sentiment analysis and opinion mining for social multimedia. Multimedia Tools and Applications 2019, 78 (6), 6939-6967.
5. Lavelle, L. Laurence Lavelle https://lavelle.chem.ucla.edu/ (03 2023).
6. Humphreys, A.; Wang, R. J.-H., Automated Text Analysis for Consumer Research. Journal of Consumer Research 2018, 44 (6), 1274-1306.
7. Kuang, S. Y.; Kamel-ElSayed, S.; Pitts, D., How to Receive Criticism: Theory and Practice from Cognitive and Cultural Approaches. Medical Science Educator 2019, 29 (4), 1109-1115.
8. Serrano-Guerrero, J.; Olivas, J. A.; Romero, F. P.; Herrera-Viedma, E., Sentiment analysis: A review and comparative analysis of web services. Information Sciences 2015, 311, 18-38.
9. Tang, F.; Fu, L.; Yao, B.; Xu, W., Aspect based fine-grained sentiment analysis for online reviews. Information Sciences 2019, 488, 190-204.
10. Toby, S., Class size and teaching evaluation: Or, the "general chemistry effect" revisited. Journal of Chemical Education 1993, 70 (6), 465.

11. Toby, S., The relationship between class size and students, ratings of faculty: Or why some good teachers should not teach general chemistry. Journal of Chemical Education 1988, 65 (9), 788.
12. Fortes, P. C.; Tchantchane, A., Dealing with Large Classes: A Real Challenge. Procedia - Social and Behavioral Sciences 2010, 8, 272-280.
13. Dictionary, O. E., "art, n.1". Oxford University Press.
14. Géron, A., Hands-On Machine Learning with Scikit-Lean, Keras & TensorFlow. 2 ed.; O'REILLY: 2019.
15. Shaukat, Z.; Zulfiqar, A. A.; Xiao, C.; Azeem, M.; Mahmood, T., Sentiment analysis on IMDB using lexicon and neural networks. SN Applied Sciences 2020, 2 (2), 148.
16. Rehman, A. U.; Malik, A. K.; Raza, B.; Ali, W., A Hybrid CNN-LSTM Model for Improving Accuracy of Movie Reviews Sentiment Analysis. Multimedia Tools and Applications 2019, 78 (18), 26597-26613.
17. Qaisar, S. M. In Sentiment Analysis of IMDb Movie Reviews Using Long Short-Term Memory, 2020 2nd International Conference on Computer and Information Sciences (ICCIS), 13-15 Oct. 2020; 2020; pp 1-4.
18. Box-Steffensmeier Janet, M.; Moses, L., Meaningful messaging: Sentiment in elite social media communication with the public on the COVID-19 pandemic. Science Advances 7 (29), eabg2898.
19. Rani, S.; Kumar, P., A Sentiment Analysis System to Improve Teaching and Learning. Computer 2017, 50 (5), 36-43.

20. Kechaou, Z.; Ammar, M. B.; Alimi, A. M. In Improving e-learning with sentiment analysis of users' opinions, 2011 IEEE Global Engineering Education Conference (EDUCON), 4-6 April 2011; 2011; pp 1032-1038.
21. Clark, T. M., Narrowing Achievement Gaps in General Chemistry Courses with and without In-Class Active Learning. Journal of Chemical Education 2023.
22. Kollalpitiya, K. Y.; Partigianoni, C. M.; Adsmond, D. A., The Role of Communication in the Success/Failure of Remote Learning of Chemistry during COVID-19. Journal of Chemical Education 2020, 97 (9), 3386-3390.
23. Schmidt, S.; Wright, Z. M.; Eckhart, K. E.; Starvaggi, F.; Vickery, W.; Wolf, M. E.; Pitts, M.; Warner, T.; Taofik, T.; Ng, M.; Colliver, C.; Sydlik, S. A., Hands-On Laboratory Experience Using Adhesives for Remote Learning of Polymer Chemistry. Journal of Chemical Education 2021, 98 (10), 3153-3162.
24. Jordan, M. I.; Mitchell, T. M., Machine learning: Trends, perspectives, and prospects. Science 2015, 349 (6245), 255-260.
25. Mansouri Tehrani, A.; Oliynyk, A. O.; Parry, M.; Rizvi, Z.; Couper, S.; Lin, F.; Miyagi, L.; Sparks, T. D.; Brgoch, J., Machine Learning Directed Search for Ultraincompressible, Superhard Materials. Journal of the American Chemical Society 2018, 140 (31), 9844-9853.
26. Chandrasekaran, A.; Kamal, D.; Batra, R.; Kim, C.; Chen, L.; Ramprasad, R., Solving the electronic structure problem with machine learning. npj Computational Materials 2019, 5 (1), 22.
27. Alex Tamkin, M. B., Jack Clark, Deep Ganguli, Understanding the Capabilities, Limitations, and Societal Impacts of Large Language Models. arxiv 2021.

28. Denecke, K.; Deng, Y., Sentiment analysis in medical settings: New opportunities and challenges. Artificial Intelligence in Medicine 2015, 64 (1), 17-27.
29. Angelidis, S.; Lapata, M., Multiple Instance Learning Networks for Fine-Grained Sentiment Analysis. Transactions of the Association for Computational Linguistics 2018, 6, 17-31.
30. Mueller, A. WordCloud for Python, GitHub, 2020.
31. Jashanjot Kaur, P. K. B., A Systematic Review on Stopword Removal Algorithms. International Journal on Future Revolution in Computer Science & Communication Engineering 2018, 4 (4).
32. Hutto, C. J., Gilbert, E.E., VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. Eight International Conference on Weblogs and Social Media (ICWSM-14) 2014.
33. Fennema, J. PyLaTeX, 1.3.2; GitHub, 2015.
34. Kadam, P.; Bhalerao, S., Sample size calculation. International journal of Ayurveda research 2010, 1 (1), 55-57.

Supporting Information

Methods

Sample size determination is critical in any empirical studies in which the goal is to make inferences about population parameters from sample estimators.[1] In this study, the goal is to select an appropriate sample size which accurately represents the target population, the total number of students enrolled in life science courses at UCLA within an academic year. Given a 95% confidence level, a 5% margin of error, and an approximate population size of 1500 students, the minimum sample size to estimate the true population proportion with the required margin of error and confidence level is 306 students. The calculation of sample size (n) used the following formula: $n = \frac{N*X}{(X+N-1)}$, where $X = \frac{{Z_{\propto/2}}^2*p*(1-p)}{E^2}$. N is the population size (=1500), $Z_{\propto/2}$ is the critical value (=1.96) of the normal distribution at $\propto = 0.05$, p is the estimated sample proportion (=0.5), E is the margin of error (=0.05). However, affected by the factors of cost, time, and convenience of data collection, we randomly selected 20 student participants from life science courses, and their scoring of the database containing 1603 phrases served as the preliminary data used to train sentiment analysis model in GCSA.

The discrepancy between the calculated sample size (=306) and the used sample size (=20) can be counterbalanced by counting the median of the total 20 scores as the true value for each teaching evaluation phrase. As long as the lower bound of the confidence interval of the proportion of 20 participants scoring the same number is above 50%, we are 95% confident that the median of the total 20 scores can truly represent the median of the entire population. To calculate the minimum sample proportion (p) required to satisfy this condition, the formula $n = \frac{{Z_{\propto/2}}^2*p*(1-p)}{E^2}$ was used, where n is the sample size (=20), $Z_{\propto/2}$ is the critical value (=1.96) of the

normal distribution at $\propto = 0.05$, E is the margin of error (unknown). The result showed that when the proportion of 20 scores being the same number is above 70%, the resulting margin of error is 20% and thus the 95% confidence interval must be larger or equal to 50%, making the median score of our selected sample size valid and unbiased.

Figures

A

Score	Count
1	138
2	433
3	279
4	447
5	306

B

Dataset	Count
Training	1282
Validation	161
Testing	160

Figure S1. Class populations of the various sentiment analysis scores (A) and sizes of the various datasets used in the training and testing of the model (B).

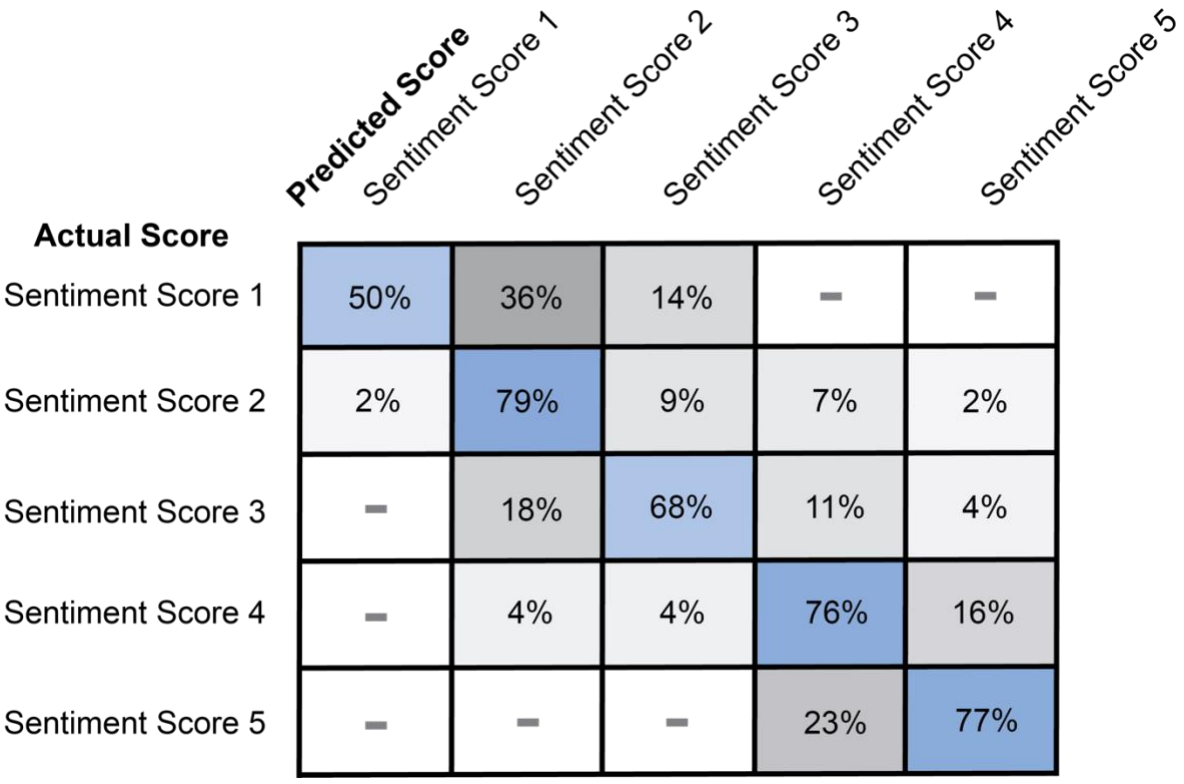


Figure S2. Confusion matrix of testing data after sentiment analysis model training. Percentages along the diagonal are correct classifications and other cells provide insight into types of errors made by the model.

Supplementary References

1. Kadam, P.; Bhalerao, S., Sample size calculation. *Int J Ayurveda Res* **2010,** *1* (1), 55-57.

Supplementary Report

Student Evaluations Sentiment Analysis

Reviews for: [redacted]

March 31, 2021

1 Overview

The following is the sentiment analysis report of your student reviews from:

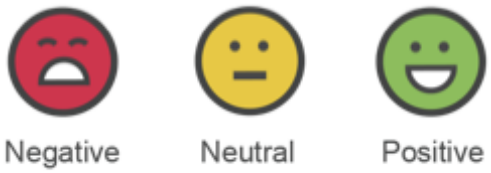
19S_CHEM_14B_DIS_[redacted].csv 20W_CHEM_14B_DIS_[redacted].csv
20W_CHEM_14B_DIS_[redacted].csv 18F_CHEM_14A_DIS_[redacted].csv

Scores were calculated using a predictive model trained via the Google Natural Language Sentiment Analysis API. The model was trained on real student course reviews sourced from UCLA faculty and RateMyProfessor.com. The data used for training was scored on a scale of 1 to 5 (negative to positive) by twenty-one UCLA students, the median score of the students was subsequently taken as the "real" score for each review. In this report, your raw review data has been processed, scored, and organized into summative and categorical sections. This review processing is similar to the example shown below:

Original Review:
The professor was very knowledgable. I thought she/he was very capable, but often they did not provide enough direction. Overall, though, I enjoyed the class!
Parsed Reviews:
The professor was very knowledgable.
I thought she/he was very capable
Often they did not provide enough direction.
Overall though, I enjoyed the class!

These parsed reviews, along with their scores, can be filtered and used to generate this report – aiming to enhance the existing review process by highlighting student sentiment towards various course aspects. Scores presented herein are predicted by the Google Natural Language API Trained model. The model has a 73% accuracy, however errored predictions typically fall within one point of the correct score (~92% of instances are either correct or within 1 point).

1: More Negative – 2: Negative – 3: Neutral – 4: Positive – 5: More Positive



2 Scoring System

The following is a more detailed discussion of the scoring scheme – note that the comments presented on this page are NOT from your course evaluations and are intended only to help introduce the the scoring system. As mentioned on page 1, the scale was initially determined by human scorers in order to train the algorithm that subsequently provides the scores presented in this and other reports. Please, note that the algorithm makes mistakes, but provides general guidance and insight into student opinion on course aspects – globally and specifically – providing strong aggregate results based on algorithm training results. Errors tend to be more pronounceed for the most negative reviews (false 1 predictions) when the true sentiment is biased highly positive. For example, over the course of a full body of reviews, some scores of 1 are bound to be given to statements as a function of misclassification. Due to the algorithms behind this report, those misclassifications are especially likely to be shown if the comments are overall heavily biased towards the positive end. We wanted to show examples of each, but if no students said anything negative, then the comments scored in error will appear as the representatives of those scores. A brief explanation and visual of the scoring system is given below. Phrases presented in the visual were scored by the algorithm:

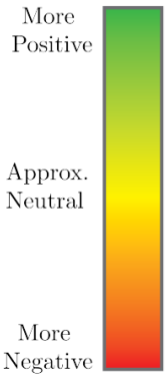
Score 5: These scores are considered to be exceptionally laudatory or positive.

Score 4: These scores are considered to be generally positive or favorable.

Score 3: These scores are considered to be neutral, or highly *subjective or ambiguous* in terms of favorability. For example, a comment claiming a course was "very fast-paced" may be considered a criticism to some, or neutral to positive to others.

Score 2: These scores are considered to be generally negative or critical.

Score 1: These scores are considered to be more unfavorable or negative.



Sample Reviews – De-Identified	Score
I believe that she is honestly the best chemistry TA out there.	5
The instructor made sure to ask and help others and was very easy to talk to in regards to the course material.	4
The tests were only focused on being able to complete math/ chemistry related problems.	3
He did not seem like he wanted to be there and sometimes sounded that he was just explaining the material to himself.	2
He made a poor impression during section and his input was unhelpful and smug during lectures when asked questions.	1

3 General Results

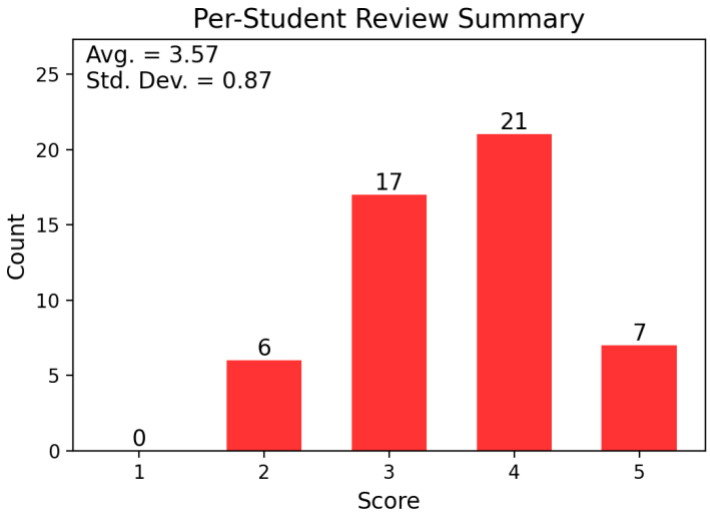


Figure 1: Scores of phrases averaged by their student authors. For example, one student may have made five comments with an average score of 4. That average is represented here.

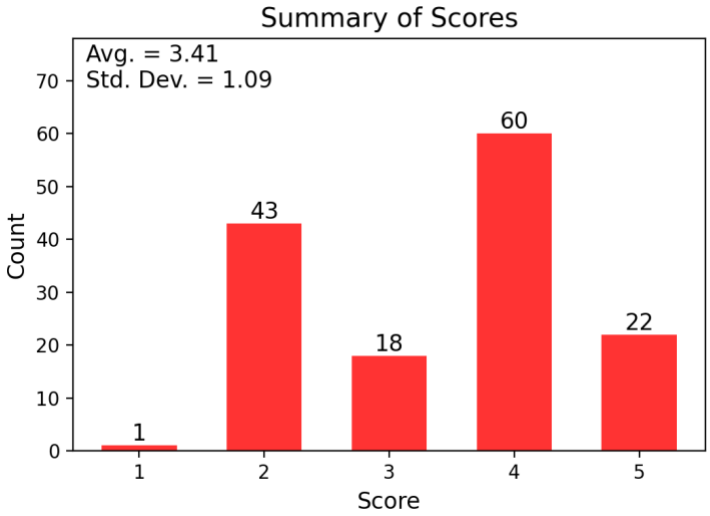


Figure 2: Raw scores of all phrases from all students. These are the scores for all parsed comments from students regardless of author. For example, if all students made 25 comments that were scored as 5, then the x-axis value of 5 would show a count of 25.

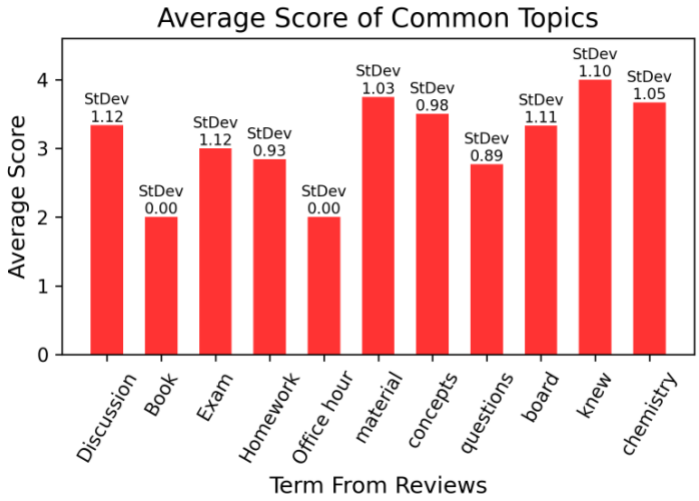


Figure 3: Average scores of phrases referencing common topics. The first 6 terms on the x-axis were manually chosen for their relevance to all courses, and the rest were chosen by an algorithm discussed on the next page.

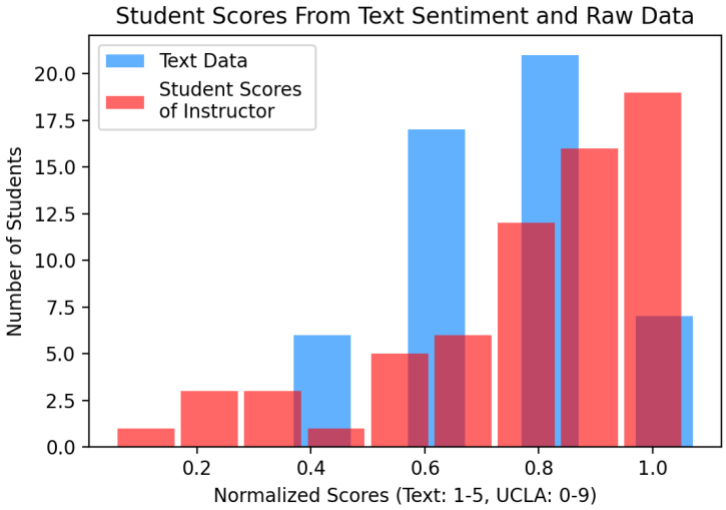


Figure 4: Overall course evaluation score distributions based on sentiment analysis (blue, see Figure 1) versus values provided by students in their course evaluations (red). Note – both scales were adjusted to a 0-1 scale for comparison purposes.

4 Word Specific Summaries

The following pages, until the end of the report, are word specific reports. These terms were selected in two ways as shown in the subsections below.

4.1 Pre-Selected Terms

The terms shown here represent aspects of the course that are considered of general interest; these terms were pre-selected by committee. Words that are grouped together are considered synonymous.

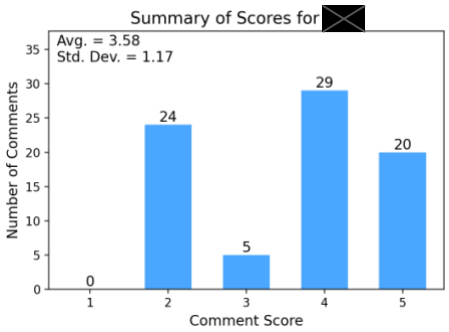
[' [redacted] ', 'They', ' He ', 'She ', 'instructor', 'lecturer', 'teaching assistant']
['Discussion', 'class', 'course']
['Book']
['Exam', 'midterm', 'final']
['Homework', 'hw ', 'worksheet', 'practice problem', 'assignment']
['Office hour', ' OH ']

4.2 Auto-Selected Terms

The figure shown here is a word-cloud of common terms (besides the ones already listed above) and was used as a foundation for selecting terms not encompassed by the above lists. The top 6 terms are presented in "word-reports" below.

The auto-selected terms are: ['material', 'concepts', 'questions', 'board', 'knew', 'chemistry']

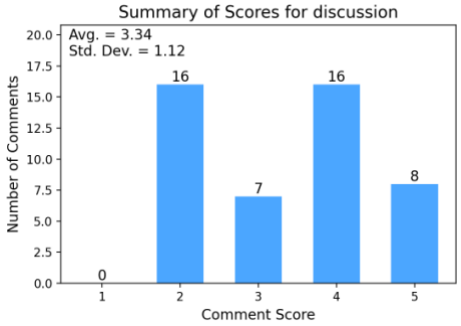
5 Subreport for terms associated with: [redacted]



5.1 Table

Review	Score
[redacted] was an outstanding ta always willing to understand our situations and expand on course material.	5
the ta was super helpful.	5
[redacted] you are by far one of my favorite ta's!	5
the ta was extremely knowledgeable in the course material, and made sure to always write on the board what from the lectures we would need to complete the discussion packet.	5
his discussion sections were very useful to me as they covered confusing topics from lecture.	4
the ta really focused on going into greater depth of what was being taught in class	4
[redacted] taught so efficiently and covered a lot of ground in a single discussion.	4
[redacted] humor also brings light to the grogginess of our 8am discussion...	4
but he did his best to try to reinforce what we learn in lecture.	4
[redacted] did his best to clarify material from the lectures.	4
the only thing i found helpful from discussion sections was what [redacted] would write on the board before class started.	3
the packet itself is the only complaint in that they were very difficult compared to the exam.	2
the course was confusing at times, so the ta helped clarify the content very well.	2
sometimes he would be a bit slow, worksheets would have helped.	2
always said they'd go over problems with us in the end	2
not too relaxed in which he was not efficient	2

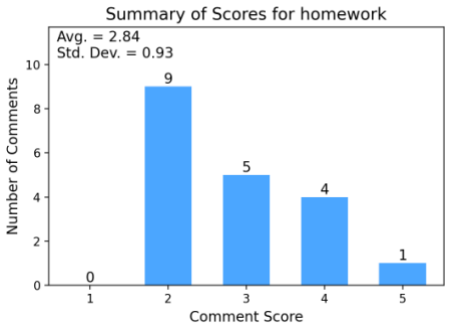
6 Subreport for terms associated with: Discussion



6.1 Table

Review	Score
the ta was extremely knowledgeable in the course material, and made sure to always write on the board what from the lectures we would need to complete the discussion packet.	5
i felt that he genuinely cared that we knew and understood the course material, and he took the time to cover what he believed were the most challenging concepts.	5
[redacted] was an outstanding ta always willing to understand our situations and expand on course material.	5
he made sure that during discussions we were aware of what was most important especially for upcoming exams.	4
the ta was very knowledgeable about the course and advocated for student's learning via discussion sections.	4
he expanded on the course material in a way that made it much easier to understand what we were taught.	4
his discussion sections were very useful to me as they covered confusing topics from lecture.	4
the ta really focused on going into greater depth of what was being taught in class	4
you never graded unfairly like some ta's that i have met and you explained concepts to me that i found very confusing during discussion section!	3
it seemed like we only went to discussion to work on a worksheet on our own.	3
i wish we would have focused less on the discussion worksheets and more on reviewing content, since we can learn the worksheet concepts on our own once the key was posted.	2
i think he could improve on walking around the classroom and actively seeking to help students instead of students coming to him.	2
but would result in not very effective discussions as sometimes no one had real questions or just started learning a new topic.	2
the discussion worksheets should not be so difficult and should contain relevant information pertaining to exams.	2
[redacted] keeps to himself most of the time during discussions.	2

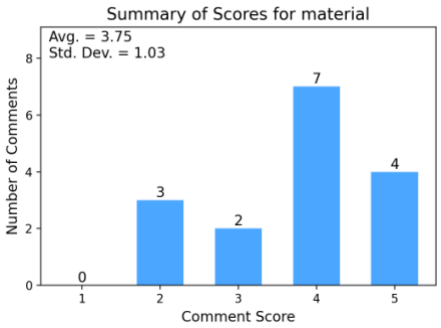
7 Subreport for terms associated with: Homework



7.1 Table

Review	Score
[redacted] does well in explaining concepts and ideas taught in lecture to us in more detail, providing examples and doing practice problems.	5
the discussions for this class aren't worth my time since i can easily try the worksheets on my own time and still do well in the course.	4
his pre-discussion lecture was helpful for the worksheet that day.	4
very relaxed with menial things such as turning in homework late.	4
his homework policy was highly appreciated.	4
i just think it would be more helpful to focus mainly on that rather than the worksheets, especially since i dont think the worksheets are a good indicator of what the professor tested us on.	3
i do wish he continued lecturing and explaining certain questions and concepts rather than allowing the majority of the discussion to be individual work time on the worksheets	3
the ta would expand on course ideas which was helpful i think the difficulty of the worksheets was too much	3
it seemed like we only went to discussion to work on a worksheet on our own.	3
i hoped he would go over more worksheet questions during discussion	3
the discussion worksheets were excessively long and complicated, and with a ta that never really helped, discussion sections for me were quite literally a frustrating, useless waste of my time.	2
i wish we would have focused less on the discussion worksheets and more on reviewing content, since we can learn the worksheet concepts on our own once the key was posted.	2
there wasn't enough las to help everyone so mostly discussion seemed like a waste since we could have done the worksheet on our own as homework or on our own time.	2
additionally, he was quite lazy, simply doing homework problems in class rather than explaining course material further or expanding on the professors lectures.	2
all this ta did was write equations on the board, go over said equations, and then tell us to work on the worksheet ourselves.	2
the ta could work on the interest of helping students and going through discussion worksheet problems with the class.	2
the discussion worksheets should not be so difficult and should contain relevant information pertaining to exams.	2
however, he made many mistakes during practice problems, leading to confusion.	2
sometimes he would be a bit slow, worksheets would have helped.	2

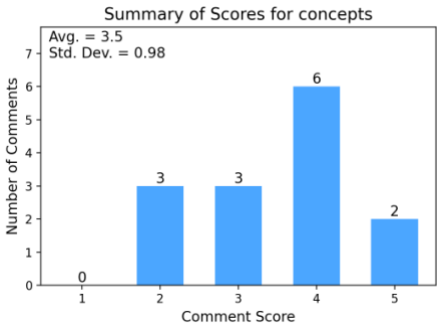
8 Subreport for terms associated with: material



8.1 Table

Review	Score
strengths: gave very useful and organized mini-lectures at the beginning of discussions that summed up what we had just learned in class and he really knew the course and its material well.	5
the ta was extremely knowledgeable in the course material, and made sure to always write on the board what from the lectures we would need to complete the discussion packet.	5
i felt that he genuinely cared that we knew and understood the course material, and he took the time to cover what he believed were the most challenging concepts.	5
[redacted] was an outstanding ta always willing to understand our situations and expand on course material.	5
i also really enjoy how he does a brief complete summary of all the notes we've learned the previous week because it sums up the new material.	4
he simplifies the complex material and teaches us exactly what we need to know, and definitely makes discussion worth attending!	4
other than that, the ta made sure to ask and help others and was very easy to talk to in regards to the course material.	4
he expanded on the course material in a way that made it much easier to understand what we were taught.	4
he was always open to taking students' questions and clearing any misconceptions about the material.	4
[redacted] did his best to clarify material from the lectures.	4
good knowledge of course material.	4
sometimes went to in depth with material not tested on/easy concepts, and other times did not go in depth enough for relevant material/more difficult concepts.	3
the presentations of material from class was mostly just writing down an equation sheet.	3
it is understandable that students must be tested and pushed to truly understand the material however, sometimes it seemed like the questions were beyond what we need to understand.	2
additionally, he was quite lazy, simply doing homework problems in class rather than explaining course material further or expanding on the professors lectures.	2
while the material presented in the lecture was difficult, the ta was able to reintroduce the material to students so that we can fully understand it.	2

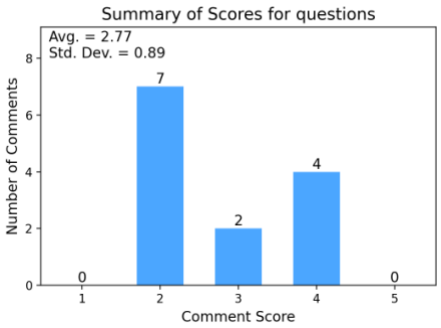
9 Subreport for terms associated with: concepts



9.1 Table

Review	Score
i felt that he genuinely cared that we knew and understood the course material, and he took the time to cover what he believed were the most challenging concepts.	5
[redacted] does well in explaining concepts and ideas taught in lecture to us in more detail, providing examples and doing practice problems.	5
[redacted] knows what he is talking about and provides his students with a lot of tips and tricks to memorize important concepts.	4
[redacted] is really skilled in breaking down difficult concepts and getting the most important points across!	4
he writes and go over the weekly concepts on the board during discussion which are very helpful.	4
he accommodates our level of understanding to help us work through difficult concepts.	4
[redacted] clarified many of the concepts that i was concerned about.	4
but, [redacted] did review concepts very well which helped.	4
i do wish he continued lecturing and explaining certain questions and concepts rather than allowing the majority of the discussion to be individual work time on the worksheets	3
sometimes went to in depth with material not tested on/ easy concepts, and other times did not go in depth enough for relevant material/ more difficult concepts.	3
you never graded unfairly like some ta's that i have met and you explained concepts to me that i found very confusing during discussion section!	3
i wish we would have focused less on the discussion worksheets and more on reviewing content, since we can learn the worksheet concepts on our own once the key was posted.	2
i think that it was hard for him to explain concepts as 14b is a very weird class with not that much interconnection	2
but i wish he went over harder concepts sometimes rather than the easier ones	2

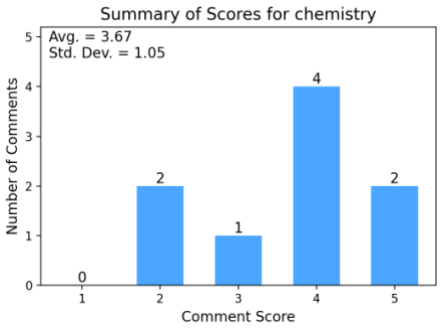
10 Subreport for terms associated with: questions



10.1 Table

Review	Score
he made sure to give a quick recap of the topics discussed in lecture and walked around for questions students may have.	4
he is also extremely knowledgeable about chemistry and tries to answer as many questions as possible.	4
he was always open to taking students' questions and clearing any misconceptions about the material.	4
he is quite knowledgeable and is able to answer my questions.	4
i do wish he continued lecturing and explaining certain questions and concepts rather than allowing the majority of the discussion to be individual work time on the worksheets	3
i hoped he would go over more worksheet questions during discussion	3
i think sometimes the questions we went over were not as beneficial as some of the other times so perhaps spending more time looking for questions or asking for student suggestions may be beneficial.	2
it is understandable that students must be tested and pushed to truly understand the material however, sometimes it seemed like the questions were beyond what we need to understand.	2
but would result in not very effective discussions as sometimes no one had real questions or just started learning a new topic.	2
he doesn't really do much in discussion and doesn't expand on any topics or goes over many questions.	2
he was dry and frankly quite rude, making it difficult to ask questions or attend his office hours.	2
whenever we have questions, we usually have to rely on the learning assistants for help.	2
weakness includes not discussing many of our questions on exams.	2

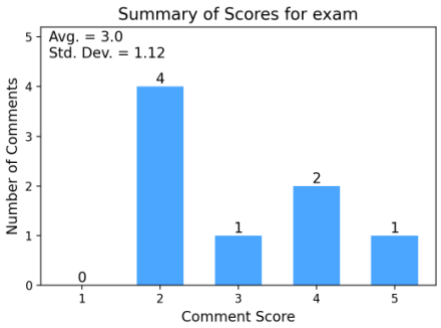
11 Subreport for terms associated with: chemistry



11.1 Table

Review	Score
i also had him for chem 14a last quarter, and the amount of effort he puts into helping the students in his section is amazing.	5
i believe that [redacted] is honestly the best chemistry ta out there.	5
he is also extremely knowledgeable about chemistry and tries to answer as many questions as possible.	4
strengths: [redacted] really cared about the students and know a lot about chemistry.	4
he's cool and definitely knows chemistry well and explains it well	4
[redacted] is a good ta for chem 14b.	4
it is difficult to teach a week's worth of chemistry in fifty minutes	3
my teaching assistant for chem 14b, [redacted], was extremely unapproachable and unfriendly as a ta.	2
i wish i went to his office hours more often in chem 14a	2

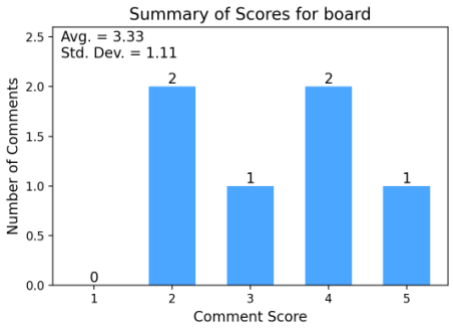
12 Subreport for terms associated with: Exam



12.1 Table

Review	Score
[redacted] does well in explaining concepts and ideas taught in lecture to us in more detail, providing examples and doing practice problems.	5
he made sure that during discussions we were aware of what was most important especially for upcoming exams.	4
very good overview of topics presented in lectures and relevant examples.	4
he acknowledges what will be on the exam and how difficult each subject is.	3
the discussion worksheets should not be so difficult and should contain relevant information pertaining to exams.	2
the packet itself is the only complaint in that they were very difficult compared to the exam.	2
did not waste time going over things that we already knew or would not be on exams.	2
weakness includes not discussing many of our questions on exams.	2

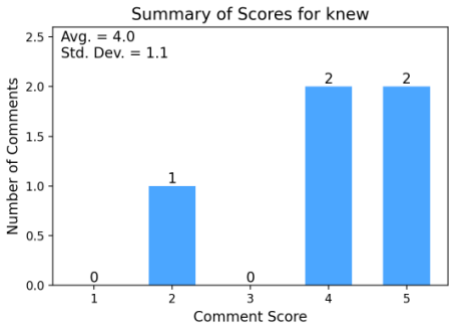
13 Subreport for terms associated with: board



13.1 Table

Review	Score
the ta was extremely knowledgeable in the course material, and made sure to always write on the board what from the lectures we would need to complete the discussion packet.	5
i really enjoyed how he reviewed key topics before starting the survey and wrote important equations on the board.	4
he writes and go over the weekly concepts on the board during discussion which are very helpful.	4
the only thing i found helpful from discussion sections was what [redacted] would write on the board before class started.	3
i stopped going because he was simply just doing problems from the book on the board for fifty minutes, a huge waste of my time.	2
all this ta did was write equations on the board, go over said equations, and then tell us to work on the worksheet ourselves.	2

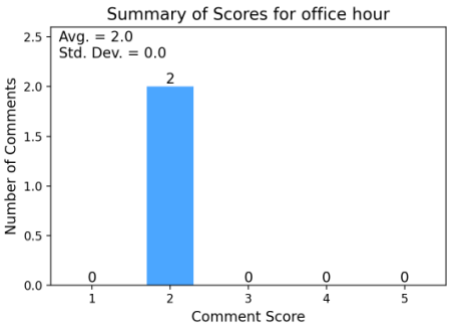
14 Subreport for terms associated with: knew



14.1 Table

Review	Score
strengths: gave very useful and organized mini-lectures at the beginning of discussions that summed up what we had just learned in class and he really knew the course and its material well.	5
i felt that he genuinely cared that we knew and understood the course material, and he took the time to cover what he believed were the most challenging concepts.	5
knew exactly what we needed help with at any time.	4
he really knew what he was talking about.	4
did not waste time going over things that we already knew or would not be on exams.	2

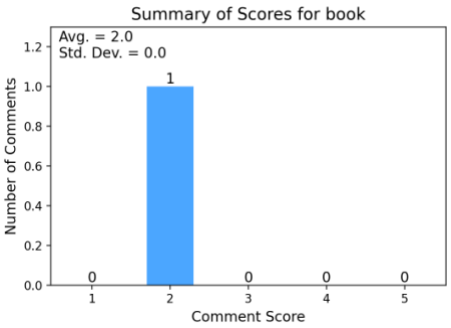
15 Subreport for terms associated with: Office hour



15.1 Table

Review	Score
he was dry and frankly quite rude, making it difficult to ask questions or attend his office hours.	2
i wish i went to his office hours more often in chem 14a	2

16 Subreport for terms associated with: Book



16.1 Table

Review	Score
i stopped going because he was simply just doing problems from the book on the board for fifty minutes, a huge waste of my time.	2

Chapter 7. Concluding Remarks

The work I have presented in this dissertation highlights my work throughout my graduate student career in integrating deep-learning with data-intensive electrochemistry problems, developing new tools, techniques, and ways of looking at systems familiar to electrochemists. I, along with my collaborators, have showed that deep learning can be applied to a range of systems and data types related to chemistry (and education).

I showed that nitrogen reduction electrodes with a constrained morphology can be simulated and evaluated in bulk – highlighting putative limits to NRR electrodes assuming a wire morphology, an associative NRR mechanism pathway, and a Volmer–Heyrovsky HER pathway in acidic ambient conditions. Further, this work showed the viability of the general procedure through which I would conduct my research throughout my graduate career. Not only were design insights gained, but insights into likely kinetic bottlenecks were elucidated via importance analyses of the trained network. In addition, the importance of N_2 availability, rather than proton availability, was highlighted as the limiting factor in electrode performance – providing evidence for the value of carefully selected morphologies whenever reactant depletion as a function of electrode morphology is probable.

Following this work, I demonstrated how a similar research path could be taken to rationally design reactant gradients. This demonstrated the value of the workflow I had established by showing its applicability to research outside of the obvious desire to optimize electrodes for difficult reactions. The on-demand generation of oxygen and reactive oxygen species (e.g., H_2O_2) gradients opens the door to the controlled study of extracellular microenvironments relevant to

microbial colonies – a heterogeneity that is fundamental to understanding the conditions that give rise to changes in microbial protein expression and survivability.

In a pivot of focus, but not general approach, I again demonstrated the applicability of this workflow by applying it to a fundamental problem in electrochemistry – mechanism assignment in cyclic voltammetry. Without proper mechanism assignment, study of a system becomes much more difficult as the underlying reaction scheme that defines the system will not be understood. Fortunately, my experience with electrochemical systems and the bulk generation of training data was well suited to the investigation of approaches to augment mechanism assignment practices, enabling the more confident use of the technique for non-domain experts while also providing a starting point for the automation of electrochemistry experiments where one mechanism, indicating a reaction pathway, is preferred over another.

The success of that project, and the follow-up into object detection approaches for detecting constituent mechanisms in CV data, carved a new path of research in cyclic voltammetry – a field that was rife with classical analyses but had a dearth of modern data science tools applied to it. Going forward, the research into deep learning and artificial intelligence tools that my approach enabled will likely bring exciting new chapters into the world of applied cyclic voltammetry. Projects are already underway for the use of ResNet classification to guide automated experimentation to steer chemical systems towards desired pathways. Further, the development of experimental automation alongside these tools allows for bulk screening of systems, perhaps to extract kinetic parameters (or other values) of more fundamental interest to chemists.

While the developments alongside the advancements in automated analyses are exciting, there is still room for yet more powerful approaches to mechanisms assignment and extraction of information from CVs directly from deep-learning tools. The development of more complex

COMSOL models where pathways are allowed to explicitly compete could result in greater numbers of mechanism classes and greater ambiguity in mechanism assignment which could be leveraged by post-analyses that aim to the establish relative contributions of underlying mechanisms to CV responses. The addition of regression models as a post-processing step could aid in providing more concrete estimations of reaction competition – rather than the current reliance on SoftMax probability distributions which provide a more or less happenstance handle from which to estimate electrochemical mechanism contributions to CV traces.

While the advent of general tools for mechanism assignment is powerful and necessary, more specific deep learning tools that leverage cyclic voltammetry could provide powerful niche uses (niche not indicative of a value judgement). For instance, models trained to establish contributions from Faradaic and charging currents could be used to prioritize different modes of charge transfer, permitting the eventual selection of system parameters that result in purely capacitive or Faradaic processes. Additionally, arguments could be made for the same workflow in any system where a degradation step, perhaps following the *EC* mechanism, is resulting in material degradation or catalyst deactivation. These proposals imply the need for mechanism assignment, but also the need for more niche tools that achieve specific engineering goals. In a final approach inspired by findings from the ResNet classification work – the noted noise tolerance of the model is indicative of possible applications in fields where sensitive detection is of paramount importance, perhaps in Fast Scan CV or detection of toxic trace elements. Sensitive detection of changes from a noisy baseline CV could enable the use of deep learning for the advancement of detection in challenging systems. In essence, the development of deep learning tools for CV has lots of room to grow, especially as applied techniques expand beyond mechanism assignment and into other types of detection or regression for specific CV related tasks.