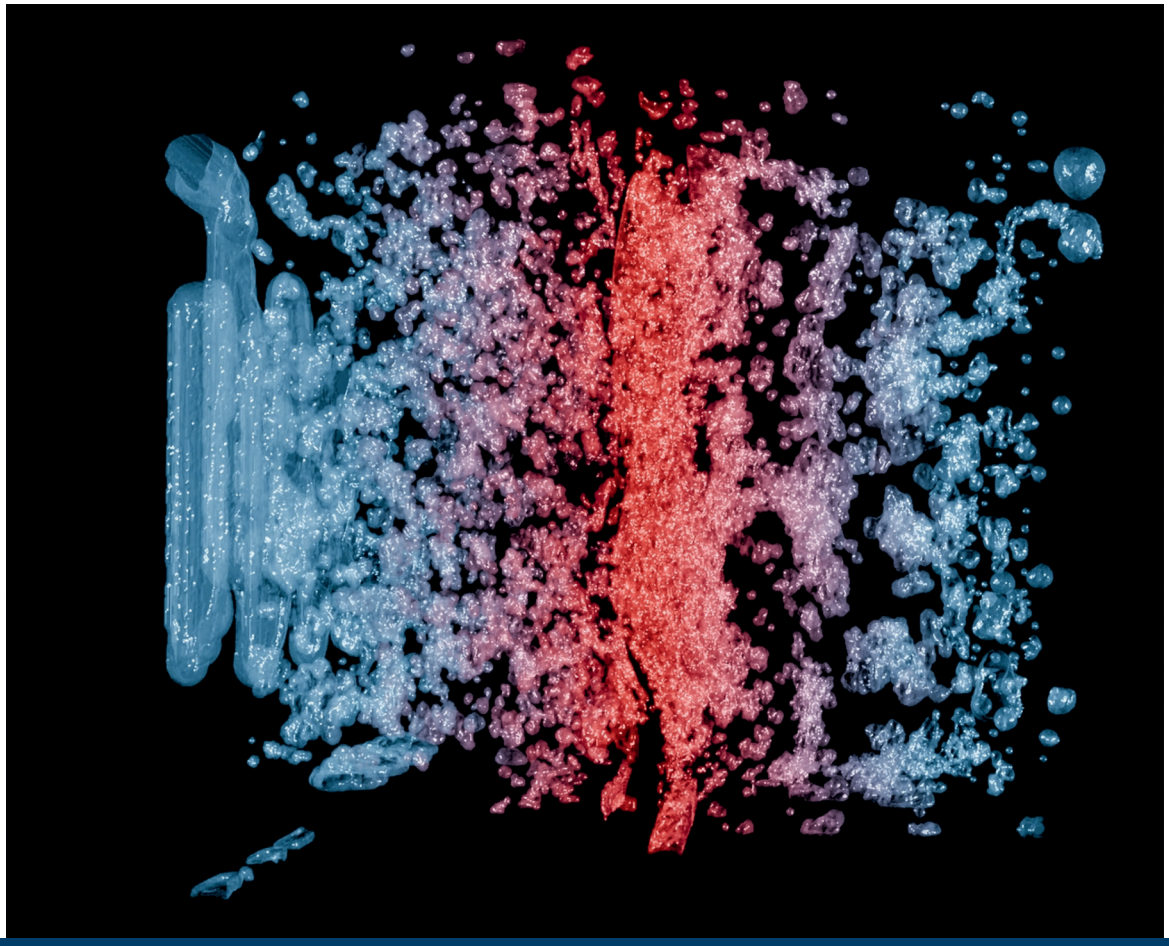

Autonomous Image Analysis to Accelerate the Discovery and Integration of Energy Materials

André Colliard Granero

Energie & Umwelt / Energy & Environment
Band / Volume 707
ISBN 978-3-95806-911-4

Mitglied der Helmholtz-Gemeinschaft

JÜLICH
Forschungszentrum

Schriften des Forschungszentrums Jülich
Reihe Energie & Umwelt / Energy & Environment Band / Volume 707

Forschungszentrum Jülich GmbH
Institute of Energy Technologies (IET)
Theorie und computergestützte Modellierung von Materialien
in der Energietechnik (IET-3)

Autonomous Image Analysis to Accelerate the Discovery and Integration of Energy Materials

André Colliard Granero

Schriften des Forschungszentrums Jülich
Reihe Energie & Umwelt / Energy & Environment Band / Volume 707

ISSN 1866-1793 ISBN 978-3-95806-911-4

Bibliografische Information der Deutschen Nationalbibliothek.
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der
Deutschen Nationalbibliografie; detaillierte Bibliografische Daten
sind im Internet über http://dnb.d-nb.de abrufbar.

Herausgeber und Vertrieb:	Forschungszentrum Jülich GmbH Zentralbibliothek, Verlag 52425 Jülich Tel.: +49 2461 61-5368 Fax: +49 2461 61-6103 zb-publikation@fz-juelich.de www.fz-juelich.de/zb
Umschlaggestaltung:	Grafische Medien, Forschungszentrum Jülich GmbH
Druck:	Grafische Medien, Forschungszentrum Jülich GmbH
Copyright:	Forschungszentrum Jülich 2026

Schriften des Forschungszentrums Jülich
Reihe Energie & Umwelt / Energy & Environment, Band / Volume 707

D 82 (Diss. RWTH Aachen University, 2026)

ISSN 1866-1793
ISBN 978-3-95806-910-7 (Print)
ISBN 978-3-95806-911-4 (E-Book)

Vollständig frei verfügbar über das Publikationsportal des Forschungszentrums Jülich (JuSER)
unter www.fz-juelich.de/zb/openaccess.

Abstract

Accelerating the discovery and development of advanced energy materials is critical to transitioning to a sustainable clean energy future. Data-driven methods, especially those using artificial intelligence (AI) and deep learning (DL), offer unprecedented exploration opportunities to significantly increase the speed at which new energy materials can be explored and optimized. Central to materials innovation and understanding of materials-related phenomena are advanced characterization techniques, in which imaging plays a vital role. Interpreting these datasets manually remains time-consuming and human-biased. To fully realize the potential of data-driven techniques, it is essential to accelerate the analysis via artificial intelligence and a standardized and efficient system for managing imaging data.

This work demonstrates the research acceleration capabilities of using DL to automate image data analysis. For this purpose, highly relevant use case scenarios were investigated with a focus on energy-related electrochemical systems. First, in the context of polymer electrolyte water electrolyzers (PEWEs), a 2D DL-based framework was developed that is tailored for high-throughput analysis of optical video recordings of the oxygen bubble evolution of a transparent cell. After binary bubble segmentation, the automated software provides quantitative insights into bubble dynamics, including time-resolved bubble coverage evolution, size distributions, bubble density maps, and morphological analysis. These results significantly improve the experimental understanding of oxygen gas dynamics in PEWE cells under different conditions, providing faster data extractions previously inaccessible due to inefficient traditional methods.

Building on these 2D bubble-analysis results, the challenges of 3D bubble analysis in vanadium redox flow batteries (VRFBs) are addressed through automated DL-based analysis of 3D synchrotron X-ray tomograms. A multi-class semantic segmentation approach was developed using a comprehensive dataset of 2294 annotated images from three different battery configurations to distinguish between bubble, electrolyte, membrane, and gasket classes with excellent performance. The developed tool addressed the challenging task of accelerating high-throughput volume analysis and streamlining bubble quantification

processes. The software enables automatic feature extraction of bubble volumes, shapes, and membrane blockage. An interactive 3D visualization tool was developed to improve the visual inspection of the analyzed volumes and their properties on the fly.

Having demonstrated 3D bubble segmentation, 3D architectures were employed to tackle complex porous-material segmentation in PEM technologies. The complexity of accurately characterizing gas diffusion layers (GDLs), microporous layers (MPLs), and catalyst layers (CLs) typically renders traditional manual and algorithmic methods inefficient. Therefore, a DL framework for segmenting micro-CT and FIB-SEM characterized porous volumes was developed and validated with physical porosity measurements. Using state-of-the-art 3D neural networks, the software successfully achieved robust binary pore/material segmentation and multi-class pore/GDL/MPL segmentation. This segmentation tool provides case-specific quantification capabilities for pore size distributions, porosity, MPL crack analysis, and MPL intrusion assessment. In addition, the 3D segmentation volumes can be visualized *in situ* employing the developed visualization capabilities.

Based on the experience gained in these contributions, developing an infrastructure designed for the standardized storage and retrieval of imaging and characterization data from energy materials experiments was crucial. Using an adapted version of the elementary multiperspective materials ontology (EMMO), a standardized metadata schema has been developed to capture the essence of different imaging modalities related to energy materials. This was achieved by combining the expertise of seven experts in various fields. This approach ensures that imaging datasets are universally stored in a semantically consistent format, as FAIR guidelines require. In addition, the work implements the metadata schema into a native graph database integrated with an ontology alignment system based on large language models (LLMs). As a result, the developed methods bridge the gap between FAIR principles and a practical implementation for experimentalists.

Together, these milestones form an end-to-end DL pipeline for automated, high-throughput analysis of electrochemical systems and the subsequent data storage. The open-access tools developed through research included in this thesis represent a significant advancement in data-driven image analysis and standardized data management. Furthermore, the expertise gained from this work has allowed the development of standardized metadata protocols to facilitate data interoperability and discoverability, accelerating the further training of DL models. Finally, combining both approaches opens a path for conceptualizing a unified platform that benefits from the developed metadata approach for data storage to create an accelerated materials research environment.

Kurzfassung

Die Beschleunigung der Entdeckung und Entwicklung fortschrittlicher Energiematerialien ist entscheidend für den Übergang zu einer nachhaltigen, sauberen Energiewirtschaft. Datengetriebene Methoden, insbesondere solche, die künstliche Intelligenz (KI) und Deep Learning (DL) einsetzen, eröffnen beispiellose Möglichkeiten, die Geschwindigkeit, mit der neue Materialien erforscht und optimiert werden können, zu steigern. Im Zentrum der Materialinnovation stehen fortschrittliche Charakterisierungstechniken, bei denen die Bildgebung eine Schlüsselrolle spielt. Traditionell bleibt die Auswertung dieser Datensätze jedoch hochgradig manuell, zeitaufwändig und durch menschliche Voreingenommenheit geprägt. Um das Potenzial datengetriebener Verfahren vollständig auszuschöpfen, ist es daher unerlässlich, die Analyse mithilfe künstlicher Intelligenz zu beschleunigen und dafür ein standardisiertes System für das Management bildgebender Daten zu etablieren.

In dieser Arbeit wird gezeigt, wie DL die Bildanalyse automatisiert und damit Forschung beschleunigt. Zunächst wurde für Polymerelektrolytmembran-Wasserelektrolyseure (PEMWE) ein 2D-DL-Framework entwickelt, das auf die Hochdurchsatzanalyse optischer Videoaufnahmen der Sauerstoffblasenentwicklung in einer transparenten Zelle zugeschnitten ist. Nach der binären Segmentierung der Blasen liefert die Software quantitative Einblicke in die Blasendynamik, beispielsweise zeitaufgelöste Entwicklung der Blasen, Größenverteilungen, Blasendichtemaps und morphologische Analysen einzelner Blasen. Diese Ergebnisse verbessern das experimentelle Verständnis der Sauerstoffgasdynamik in PEMWE-Zellen unter verschiedenen Bedingungen erheblich und ermöglichen schnellere Datenauswertungen, die mit traditionellen Methoden bisher nicht zugänglich waren.

Aufbauend auf den 2D-Ergebnissen zur Blasenanalyse haben wir unseren Deep-Learning-Workflow auf 3D-Tomographiedaten von Vanadium-Redox-Flow-Batterien erweitert. Hier-für wurde eine automatisierte DL-gestützte Auswertung von 3D-synchrotronischen Röntgen-tomogrammen realisiert. Mit einem Datensatz von 2294 annotierten Bildern aus drei Batterie-Konfigurationen wurde ein Multi-Klassen-Segmentierungsmodell entwickelt, das Blasen, Elektrolyt, Membran und Dichtung mit hoher Genauigkeit unterscheidet. Die entwickelte Software beschleunigt die Hochdurchsatz-Volumenanalyse und optimiert

den Prozess der Blasenquantifizierung, indem sie automatisch Blasenvolumen, -formen und Membranblockaden extrahiert. Ein interaktives 3D-Visualisierungstool ermöglicht die sofortige Inspektion der analysierten Volumina und ihrer Eigenschaften.

Nach der erfolgreichen 3D-Blasensegmentierung setzten wir 3D Architekturen zur Segmentierung komplexer poröser Materialien in PEM-Technologien ein. Die komplizierte Charakterisierung von Gasdiffusionsschichten (GDL), mikroporösen Schichten (MPL) und Katalysatorschichten (CL) macht traditionelle manuelle und algorithmische Verfahren ineffizient. Daher wurde ein DL-Framework zur Segmentierung von Mikro-CT- und FIB-SEM-Volumina entwickelt und anhand physikalischer Porositätsmessungen validiert. Mit modernen 3D Modellen erzielte die Software sowohl robuste binäre Segmentierungen (Pore/Material) als auch Mehrklassen-Segmentierungen (Pore/GDL/MPL) mit hoher Genauigkeit. Das Tool liefert Kennzahlen wie Porengrößenverteilungen, Porosität, Rissanalysen in der MPL und Eindringtiefe in die MPL. Darüber hinaus können die 3D-Segmentierungsergebnisse mithilfe unserer Visualisierungstools *in situ* dargestellt werden.

Aufbauend auf diesen Ergebnissen war es entscheidend, eine Infrastruktur für die standardisierte Speicherung sowie den Abruf bildgebender und charakterisierender Daten aus Energiematerial-Experimenten zu entwickeln. Unter Verwendung einer angepassten Version der Elementary Multiperspective Materials Ontology (EMMO) wurde ein standardisiertes Metadatenschema erstellt, das verschiedene Bildgebungsmodalitäten in der Materialcharakterisierung abdeckt und das Fachwissen von sieben Experten integriert. Dieser Ansatz gewährleistet, dass Bilddatensätze in einem semantisch konsistenten Format gespeichert werden, wie es die FAIR-Prinzipien verlangen. Das Metadatenschema wurde in eine native Graphdatenbank implementiert und mit einem auf großen Sprachmodellen basierenden Ontologie-Abgleichssystem verknüpft. So überbrücken die entwickelten Methoden die Lücke zwischen theoretischen FAIR-Anforderungen und einer praxisnahen Umsetzung für Experimentatoren.

Zusammen bilden diese Meilensteine eine End-to-End-DL-Pipeline für automatisierte, hochdurchsatzfähige Analysen und standardisierten Datenverwaltung elektrochemischer Systeme. Durch die Adressierung 2D- und 3D-Bildanalyseaufgaben wird das Potenzial von DL-Ansätzen zur Beschleunigung der Materialforschung deutlich. Zusätzlich ermöglichte die enge Zusammenarbeit mit Experimentpartnern die Entwicklung standardisierter Metadatenprotokolle, die die Interoperabilität und Auffindbarkeit weiter verbessern und so das Training von DL-Modellen beschleunigen. Die Kombination dieser Ansätze ebnet den Weg für eine einheitliche Plattform, die von der entwickelten Metadateninfrastruktur für die Datenspeicherung profitiert und gleichzeitig die Bereitstellung trainierter DL-Modelle für eine beschleunigte Materialforschung ermöglicht.

Acknowledgement

I am deeply thankful that, of all the choices one has to make in life with infinite unexpected outputs, I ended up doing my PhD here. Curiosity-driven decisions started my academic journey in experimental chemistry and drove me all the way to artificial intelligence-accelerated investigation, enriching my experience from many perspectives and culminating in what I can now be sure is my passion. Nevertheless, the thin boundary between heaven and hell in an academic experience is not topic-related; it is totally environment-dependent. Therefore, in this section, I want to acknowledge the perpetrators of the most remarkable and enjoyable chapter of my life so far.

First of all, I want to express my most sincere gratitude to my PhD supervisor, Prof. Dr. Michael Eikerling, director of the IET-3, for accepting and trusting me as his PhD student for the past three years. Our story starts in 2021 with a brief internship in his group. After completing it, I was able to grasp the excellent working conditions offered by the guidance of a wise professor. His critical thinking could overcome all technical details and provide outstanding logical guidance in topics like mine, often far from the core theoretical electrochemistry, characteristic of the group. Throughout my PhD, Prof. Eikerling not only provided academic supervision but also carried out interesting and inspiring conversations about career, future, and life, making him a true example to follow.

I also want to thank my direct supervisor, Dr. Kourosh Malek, who supported me since day one and trusted me throughout the process. Kourosh offered research opportunities and opened paths that not all supervisors would be capable of, and therefore, I want to express my gratitude. Beyond academic guidance, Kourosh showed me another type of critical thinking and guided me through a process of understanding how the system works.

In addition to Kourosh, I want to thank Dr. Mehrtoos Eslamibidgoli, my day-to-day supervisor, for his assistance during my early academic development. I also greatly appreciate his constant availability and support, especially for the detailed and fast revision

of all scientific outputs from my side, which significantly enhanced my scientific communication skills over the years.

I want to express my gratitude to my fellow PhD students, Fabian Tipp, Max Dreger, and Nils Bruch. Their day-to-day support and the interesting conversations made the difference during this time. I can not imagine better colleagues to share this adventure from the beginning of the Bachelor's to the end of the PhD, supporting each other as a family.

Furthermore, I want to thank all the members of the IET-3 who made this experience unforgettable. I am delighted that we had the opportunity to spend time together this time, both inside and outside the institute. A special thanks to everyone who took the institute's cake rules seriously, creating valuable breaks during busy days.

My PhD work would not have been possible without the experimental collaborators' support. Thank you to Prof. Aimy Bazylak from the University of Toronto, Canada, for hosting me for a three-month research stay. Her guidance, combined with the TEAM group, was exceptional for that short time, turning the research stay into one of the best experiences in my life.

Additionally, I want to sincerely thank all the experimental collaborators from my PhD project, Prof. Jasna Jankovic, Dr. Christian Rodenbücher, Dr. Tobias Morawietz, and Prof. Roswitha Zeis, for their support and patience during this period. I also want to express my gratitude for the financial support I received on various occasions from Helmholtz Imaging, MITACS, StoRIES, HITEC, GCMAC, and the Ministry for Education and Research (BMBF). Without their support, it would not have been possible to connect with researchers worldwide and spread my investigation at international conferences.

Outside the academic world, I want to mention a special thank you to my parents, Enriqueta, Gustavo, and Fernando, for being the reason that I made it here. Also, I thank my best friend Hamza for his unconditional support over 17 years and 13 years from the distance. Thank you, Julian and Pepe, for inspiring me during my academic career. And I also want to extend my gratitude to my closest family, Facu, Lala, Jan, Mashita, Joaqui, Mili, and abuelita Nelly.

Thanks to my beloved girlfriend, Dani, for being the central pillar of my life in the past few years. Your unconditional support in the good and bad times was the fuel I needed to accomplish this. Therefore, part of this work also belongs to you.

Last but not least, I want to thank myself for never giving up against the adversities and continuously fighting for a better future.

List of Acronyms

Acronym	Definition
2D	Two–dimensional
3D	Three–dimensional
AI	Artificial Intelligence
AE	Acoustic Emission
API	Application Programming Interface
BMIT-ID	Biomedical Imaging and Therapy Beamline Identifier
CAS	Chemical Abstracts Service
CHADA	Characterization Data
CL	Catalyst Layer
CMOS	Complementary Metal–Oxide–Semiconductor
CNN	Convolutional Neural Network
CSV	Comma–Separated Values
CT	Computed Tomography
CUDA	Compute Unified Device Architecture
CV	Computer Vision
DL	Deep Learning
EIS	Electrochemical Impedance Spectroscopy
EMMO	Elementary Multiperspective Materials Ontology
FAIR	Findable, Accessible, Interoperable, Reusable
FIB-SEM	Focused Ion Beam Scanning Electron Microscopy
FF	Flow Field
FN	False Negative
FP	False Positive
GIF	Graphic Interchange Format
GDL	Gas Diffusion Layer
GPU	Graphics Processing Unit

HDF5	Hierarchical Data Format, version 5
HER	Hydrogen Evolution Reaction
IoU	Intersection over Union
JPEG	Joint Photographic Experts Group
JSON	JavaScript Object Notation
LLM	Large Language Model
LVM	Large Vision Model
MAP	Materials Acceleration Platform
MPL	Microporous Layer
PEM	Polymer Electrolyte Membrane
PEMFC	Polymer Electrolyte Membrane Fuel Cell
PEMWE	Polymer Electrolyte Membrane Water Electrolyzer
PFIB-SEM	Plasma Focused Ion Beam Scanning Electron Microscopy
PNG	Portable Network Graphics
POT	Poly(o-toluidine)
RHE	Reversible Hydrogen Electrode
ROI(s)	Region(s) of Interest
SHE	Standard Hydrogen Electrode
STXM	Scanning Transmission X-ray Microscopy
SEM	Scanning Electron Microscopy
TEM	Transmission Electron Microscopy
TIFF	Tag Image File Format
TN	True Negative
TP	True Positive
VRFB	Vanadium Redox Flow Battery
VTK	Visual Toolkit
YAG	Yttrium Aluminum Garnet
YAML	YAML Ain't Markup Language

Contents

Chapter 1

Introduction

1.1 Future energy conversion and storage technologies

The transition from a fossil fuel-based economy to a green energy economy is pivotal for minimizing the consequences of climate change and meeting the ever-increasing global energy demand.[1,2] In the current situation, alternatives for producing green energy, such as solar, wind, geothermal, and hydroelectric power are rapidly growing.[3,4] However, a crucial drawback of renewable energy sources lies in the intermittency caused by fluctuations in the energy generated due to unpredictable natural influences.[5] Therefore, the scientific community is challenged to find alternative ways to store excess energy when supply exceeds demand and release when demand exceeds supply.[6]

1.1.1 Lithium vs. hydrogen: quantifying the storage-capacity gap

Currently, the focus has been shifting towards employing lithium to solve the energy storage problem. Lithium battery technology has been developed for decades, and batteries have reached a high commercial maturity in recent years with the electric car revolution.[7] However, there is one major problem with battery technologies that will be difficult to overcome for large-scale deployment in energy storage: the limited amount of battery precursor materials on Earth.[8] To better understand the issue, a simple question should be posed: how much storage capacity is needed to meet the global energy demand?[9]

Recent reports estimate that the global annual energy consumption is on the order of 600 EJ.[10] To ensure grid stability in systems with a high share of intermittent renewable

technologies, it is generally estimated that storage capacity should cover about 10% of the total annual energy consumption.[11] This equates to about 60 EJ that needs to be stored at all times. A follow-up question is how many electrons must be stored if 60 EJ are required? To convert this energy into the total number of electrons needed at an average lithium-ion battery cell voltage (about 3.7 V), first, the total electric charge Q is given by,

$$Q = \frac{E_{\text{storage}}}{V} = \frac{60 \times 10^{18}\,\text{J}}{3.7\,\text{V}} \approx 1.62 \times 10^{19}\,\text{C}. \tag{1.1}$$

Since the elementary charge of an electron is approximately $1.602 \times 10^{-19}\,\text{C}$, the total number of electrons N_e needed is given by

$$N_e = \frac{1.62 \times 10^{19}\,\text{C}}{1.602 \times 10^{-19}\,\text{C/electron}} \approx 1.01 \times 10^{38}\,\text{electrons}. \tag{1.2}$$

With the total number of electrons stored, it is possible to calculate the potential number of electrons that could be stored if all the mineable lithium in the Earth's crust were used. According to the U.S. Geological Survey, global lithium reserves are about 21 million tons. Converted to kilograms, it is $2.1 \times 10^{10}\,\text{kg}$.[12] Considering that the molar mass of lithium is $7 \times 10^{-3}\,\text{kg/mol}$, the available number of moles is equal to

$$n_{\text{Li}} = \frac{2.1 \times 10^{10}\,\text{kg}}{7 \times 10^{-3}\,\text{kg/mol}} \approx 3.0 \times 10^{12}\,\text{mol}. \tag{1.3}$$

To obtain the total number of Li atoms, it is multiplied n_{Li} by Avogadro's number ($6.022 \times 10^{23}\,\text{mol}^{-1}$), which equals a total of 1.81×10^{36} atoms, considering that each lithium atom participates in the transfer of one electron during battery operation. Now it is possible to compare the potential number of electrons stored by lithium with the total global demand,

$$\frac{N_{\text{Li}}}{N_e} \approx \frac{1.81 \times 10^{36}}{1.01 \times 10^{38}} \approx 0.018\ (\sim 1.8\%). \tag{1.4}$$

With this knowledge, it is observed that if theoretically all lithium atoms were used to store energy, it could only cover about 1.8% of the total world demand, which is insufficient, especially considering that the energy demand is increasing year by year.

But what about hydrogen technologies that use water as a precursor for storage? Again, estimating the number of electrons that could be stored using the available water in the Earth's crust is possible. It is known that the Earth's oceans contain about $1.4 \times 10^{21}\,\text{kg}$ of water.[13] With a molar mass of 18 g/mol, the total number of water molecules is

$$N_{H_2O} \approx \frac{1.4 \times 10^{21}\,\mathrm{kg}}{18 \times 10^{-3}\,\mathrm{kg/mol}} \times 6.022 \times 10^{23}\,\mathrm{mol}^{-1} \approx 4.69 \times 10^{46}\,\mathrm{molecules}. \tag{1.5}$$

Since each water molecule contains two hydrogen atoms, each atom can transfer one electron in a fuel cell reaction, resulting in a potential number of storeable 9.38×10^{46} electrons. Comparing the number of electrons to the total storage requirement of 60 EJ, obtaining

$$\frac{N_{e,\mathrm{water}}}{N_e} \approx \frac{9.38 \times 10^{46}}{1.01 \times 10^{38}} \approx 9.3 \times 10^{8}. \tag{1.6}$$

These data indicate that water's electron storage capacity is approximately 930 million times greater than current energy-storage demands, vastly surpassing the limits of lithium-based batteries. The result highlights the considerable promise of hydrogen-based technologies, yet their broad implementation remains constrained by several technical and economic barriers. Although alternative storage solutions are often preferred, these obstacles must be overcome before commercial deployment can proceed, presenting clear opportunities for innovation. Therefore, a thorough understanding of the underlying causes behind these challenges is essential to guide the development of effective solutions. For this reason, this work focuses on two of the most critical technologies for a sustainable green-hydrogen economy: electrochemical technologies using polymer electrolyte membranes.

1.1.2 PEM fuel cells and water electrolyzers

Polymer electrolyte membrane technologies are promising candidates for sustaining an economy based on hydrogen conversion.[14] These technologies allow the production of hydrogen via PEM water electrolyzers (PEMWEs) and the subsequent conversion back to energy via PEM fuel cells (PEMFCs), enabling the storage and distribution of energy in a gaseous carrier and the conversion back to energy at the point of use. The advantages of PEMWEs over other technologies are highlighted by the high efficiencies achieved of around 80%, the high purity of the hydrogen produced, and the compact design by the use of thin films.[15] PEMFCs are also a solid fuel cell candidate against other technologies, achieving efficiencies of 65% at low temperatures, with high power densities, and all consistent with the compact design characteristic of PEM technologies.[16,17]

Both technologies work in a similar way, one being the inverse reaction of the other. Figure 1.1 shows a simplified PEM stack. To explain the cycle, first, the water must be converted into hydrogen using the PEMWEs. Here, the water is introduced into the

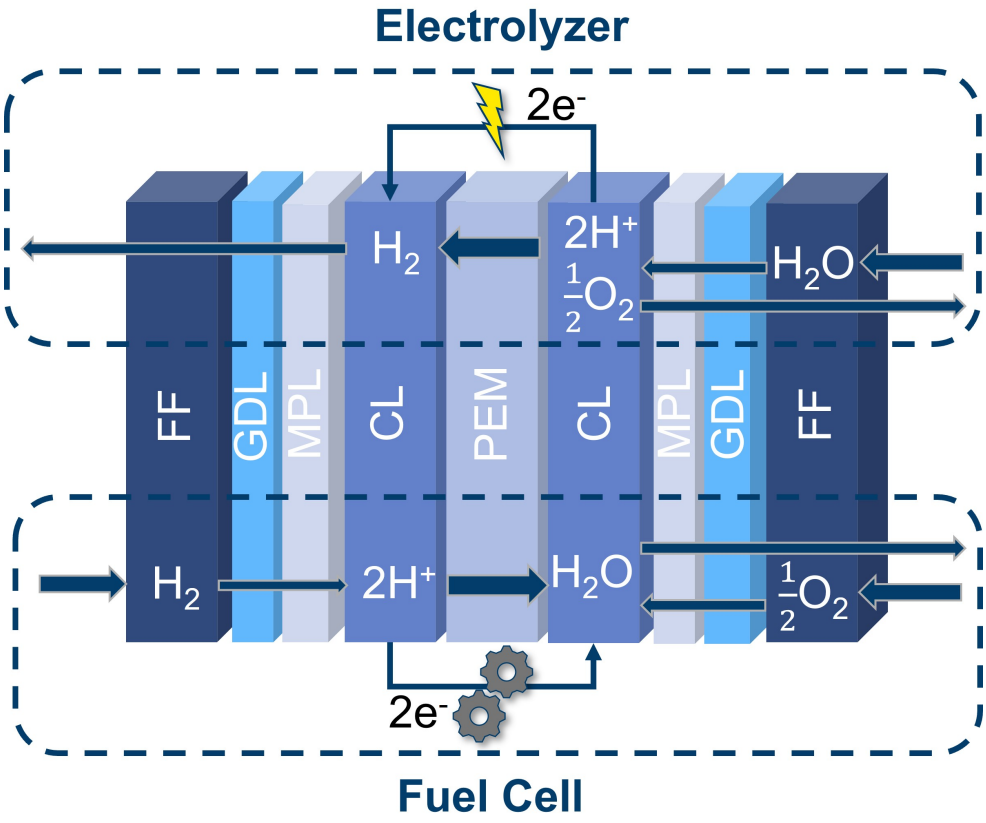

Figure 1.1: Schematic illustration of the operational principle of a PEM cell, accompanied by the constituent components. The upper part of the figure is associated with the water electrolyzer function, while the lower part corresponds to the fuel cell function.[14]

cell's anode through the flow field (FF), evenly distributed throughout the cell thanks to the gas diffusion layer (GDL) and the microporous layer (MPL).[14] When the water reaches the anode catalyst layer, it is oxidized and split into oxygen molecules, protons, and electrons;

$$\mathrm{H_2O} \rightarrow 2\mathrm{H^+} + \frac{1}{2}\mathrm{O_2} + 2\mathrm{e^-}. \tag{1.7}$$

The generated protons pass through the electron-isolating polymer electrolyte membrane (PEM) and recombine in the cathode catalyst layer (CL), where they are reduced to hydrogen molecules by electrons provided by an external source,

$$2H^+ + 2e^- \rightarrow H_2. \tag{1.8}$$

This gives the overall reaction to produce hydrogen from water using a PEMWE,

$$2\,H_2O \rightarrow 2\,H_2 + O_2. \tag{1.9}$$

Once the hydrogen is generated and stored, it can be converted back to energy using the PEMFC. To do this, molecular hydrogen is supplied to the anode CL, where it splits into two protons and two electrons,

$$H_2 \rightarrow 2H^+ + 2e^-. \tag{1.10}$$

The protons diffuse through the PEM to the cathode CL, while the electrons are transported through an external circuit to the cathode. Finally, at the cathode, the protons and electrons react with the oxygen from the air to form water,

$$H_2 \rightarrow 2H^+ + 2e^-. \tag{1.11}$$

Thus, the overall reaction becomes

$$2\,H_2 + O_2 \rightarrow 2\,H_2O. \tag{1.12}$$

PEM technologies producing and consuming hydrogen as an energy carrier can power several industries, particularly the transportation industry, due to the compactness of the cells, the fast start-up, and the low-temperature and low-pressure operating conditions coupled with high volumetric power densities from 3.1 kW/L in the Toyota Mirai[18] up to 9.8 kW/L in experimental stacks.[19,20] However, these technologies still face difficulties reaching a broader market and competing with other alternatives. It is noticeable that the CL, where the reaction occurs, deeply influences the cell's performance. The CL still faces the problem of reliance on expensive transition metal catalysts such as platinum, palladium, and iridium, which, with their high cost and scarce availability, limiting the widespread adoption of PEM technologies.[21]

Furthermore, PEM technologies highly depend on appropriate low-temperature (50--100°C) operating conditions for efficient cell performance.[22] Therefore, the GDLs, MPLs, and FFs are still on the research focus as well, since these layers are responsible for the

water management within the cell and the mass and heat transport needed to ensure the efficient performance of the reactions at the CL level.[23]

In addition, the PEM layer plays an important role in the efficiency of the cell. Today, Nafion is the most common material used to form the membrane, an expensive and environmentally hazardous perfluorinated polymer that allows only protons to pass through, as required. Despite great efforts to replace this material with a more affordable and environmentally friendly alternative, this layer still has room for improvement.[24]

1.1.3 Vanadium redox flow batteries

While PEM devices excel in mobile applications and hydrogen conversion, storing hydrogen for long-term periods is a complex challenge that requires heavy, high-pressure vessels. Therefore, another complementary technology for the long-term storage of surplus energy generated by renewable energies is investigated in this work, the vanadium redox flow batteries (VRFBs).

VRFBs excel in performance for long-term, large-scale stationary grid energy storage.[25] This is because VRFBs have a decoupled energy capacity, meaning that their energy storage capacity is determined solely by the amount of liquid electrolyte used to store the energy. This has the advantage that VRFBs can increase the size of the electrolyte tanks to increase energy storage capacity without changing the cell stack. Furthermore, this technology also shows a high cycle life of over 14000 charge-discharge cycles and lifetimes of 10 to 20 years, achieving 100% depth-of-discharge without any degradation, while maintaining a high-coulombic efficiency of over 98 %.[26] Additionally, VRFBs employ aqueous electrolytes, ensuring operational safety and making them a strong candidate for large-scale energy storage.[25,26]

The working principle of VRFBs is shown in Figure 1.2. It is based on the principle of two separate electrolytes. As the name implies, each tank contains vanadium ions in different oxidation states. The two electrolytes are kept in separate tanks and pumped through the cell stacks where the electrochemical reactions occur. The cell stack itself consists of two individual cells similar to the PEM cell, each of which containing a porous carbon felt electrode for the redox reactions to take place, and an ion exchange membrane that keeps the electrolytes separated in two electrolyte streams, but allows the permeation of protons.[28] In the negative half-cell, the vanadium is oxidized during the discharge event and protons acts as mobile charge carriers to maintain charge neutrality. This phenomenon is given by,

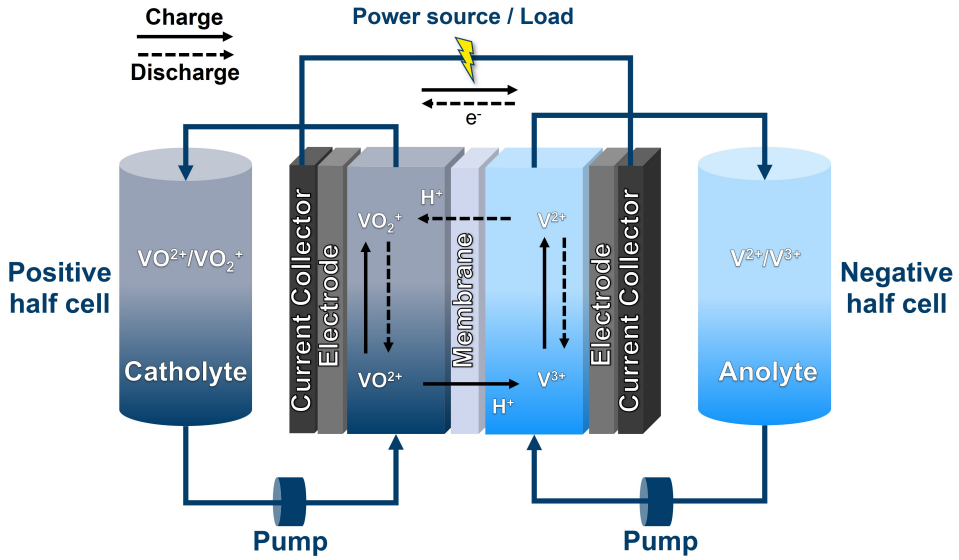

Figure 1.2: Schematic illustration of the operational principle of a VRFB. Solid arrows represent the processes of charge, while dashed arrows represent the discharge processes.[27]

$$\mathrm{V^{2+} \rightarrow V^{3+} + e^{-}}. \tag{1.13}$$

During charging, the reverse reaction takes place,

$$\mathrm{V^{3+} + e^{-} \rightarrow V^{2+}}. \tag{1.14}$$

On the other side, the positive half-cell involves the following redox reaction, during discharging,

$$\mathrm{VO_2^{+} + 2\,H^{+} + e^{-} \rightarrow VO^{2+} + H_2O}. \tag{1.15}$$

During charging, the reverse reaction occurs,

$$\mathrm{VO^{2+} + H_2O \rightarrow VO_2^{+} + 2\,H^{+} + e^{-}}. \tag{1.16}$$

These complementary reactions control the entire charging and discharging process. Separating the two redox couples in different electrolyte tanks (using the same vanadium

species in different oxidation states) minimizes cross-contamination and simplifies the chemistry.[25]

However, several challenges need to be addressed to commercialize this technology further. The main issues revolve around the low energy density obtained from the vanadium species, whose cost is subject to price volatility. In addition, recent studies have revealed performance losses associated with hydrogen evolution side reactions in the electrolyte, leading to active site blockage and reduced cell performance.[29]

1.2 AI in energy science: the fourth industrial revolution and its application in materials science

To accelerate the exploration of alternative energy conversion and storage technologies, it is critical to understand their bottlenecks and how they are slowing the development of these technologies. To put the problem in perspective, developing new materials from initial concept in the laboratory to actual implementation at a device level for commercialization often takes 10 to 20 years. In addition, it costs tens to hundreds of millions of dollars, hindering rapid responses to the pressing technology challenges listed in the previous sections.[30] Traditionally, materials discovery pipelines rely on sequential experimentation and human analysis, which face the challenge of exploring vast chemical composition and mechanical spaces with low success rates. This problem makes the process resource-intensive and slow.[31]

Artificial intelligence (AI), particularly deep learning (DL), offers powerful tools to accelerate research into green energy technologies. Over the past decade, convolutional neural networks (CNNs) have driven breakthroughs in image-based materials analysis, beginning with the CNN-based AlexNet's victory in 2012 in on the the most relevant image classification challenge called ImageNet. The success of CNNs was rapidly extended to deeper variants such as the visual geometry group neural network (VGGNet)[32,33] and residual neural network (ResNet)'s residual architectures.[34] For pixel-precise segmentation tasks, the U-shaped neural network (U-Net) demonstrated how encoder-decoder designs with skip connections can extract detailed structures from microscopy volumes.[35] Meanwhile, recurrent models such as long-short term memory networks (LSTMs)[36] and deep recurrent neural networks (RNNs)[37] have enabled time-series and sequential data analysis, from spectroscopic signals to experimental protocols. These methods have been made possible by GPUs, accelerated by the common GPU programming language called compute unified device architecture (CUDA) from NVIDIA, allowing massively parallel training and inference.[38] Together, these advances reflect a technological shift that experts liken to

a fourth industrial revolution,[39,40] fundamentally changing how complex, heterogeneous data is processed.

The main advantage of employing AI for scientific and industrial purposes is the ability to automate and accelerate complex tasks that require understanding of a variable context that cannot be summarized in a set of hard-coded rules. This fact replaces, accelerates, and increases the performance in activities that in the past had to be done by humans.[41]

This paradigm shift is emerging simultaneously with the global energy transformation.[42–45] Although AI is widely adopted at the user level, its integration into science, especially energy science, has been slow and limited.[46,47] For this reason, bridging the gap between AI and energy science becomes crucial to accelerate the research in the green energy field.

Based on the previous statement, a logical follow-up question is how AI can be practically applied to help address challenges of energy transformation. There is no simple answer to this question, as AI can be used in various ways to accelerate and enhance energy research.[48] For the AI-accelerated discovery of new energy materials, the possibilities range from using AI to automate experimentation by using robotic labs or materials acceleration platforms (MAPs) that autonomously fabricate new samples and further optimize specific properties, to direct the AI-based exploration of the vast chemical space.[49–51] AI approaches are also capable of accelerating the understanding of existing materials under use through forecasting of degradation processes.[52] Moreover, AI can be employed to accelerate density functional theory and molecular dynamics simulations,[53,54] and discover correlations between materials structures and performance metrics.[55]

However, not all fields of AI and energy materials science have been developed to the same extent. The field of AI-based autonomous image analysis still faces a large gap for the scientific community to address.[56] The bottleneck lies in extracting specific properties from imaging data, due to the complexity, abstractness, and variety of information that can be of interest from a material sample. More progress has been made in other scientific fields, such as biology and medicine, where a heavy reliance on image analysis as a primary characterization method has enabled a higher level of development in automated image analysis.[57] Similarities in the type of analysis performed facilitate the transfer of medical and biological-inspired methods to materials science (e.g., cell analysis vs. nanoparticle analysis).[58,59] Therefore, closing the gap found in image analysis for accelerating the development of for energy technologies proved to be essential to optimize and improve the current, outdated manual workflows for extracting quantities and properties from imaging data.[56]

1.3 Artificial intelligence-based imaging data characterization

Based on the previous sections, it is clear that certain aspects of energy storage and conversion technologies need to be improved to reach a level mature enough for commercialization. These improvements are closely related to understanding the phenomena that occur inside the cells and the interplay between the different layers in the devices and their structure.

Characterization of experimental samples often relies on visual assessment of the properties of materials. The acquisition of sample visualizations varies greatly depending on the scale range of interest (e.g., nanoscale to macroscale), the dimensionality of the required data (e.g., images, videos, or tomographies), and the type of interaction with the materials to create the visual information (e.g., photons, X-rays, electrons, or repulsion). For this reason, various methods are available to image material samples, such as electron microscopy, X-ray radiography, and tomography.[60]

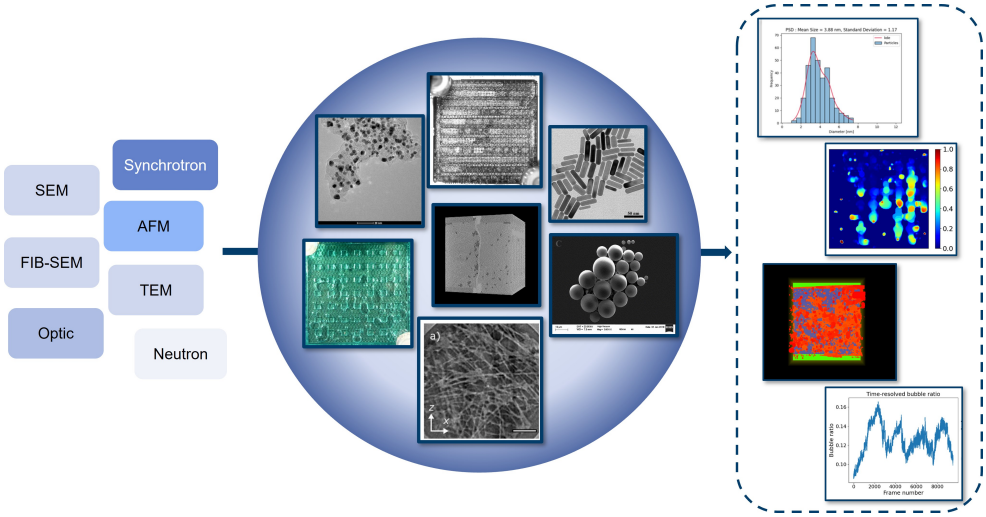

Figure 1.3: Overview of the process of different characterisation methods with the corresponding imaging data and the subsequent quantifications extracted from the data.[59,61,62]

As shown in Figure 1.3, there are many different imaging-based materials characterization methods, each outputting a diverse range of 2D, 3D, and even 4D image data. Collecting those imaging datasets from materials samples is an extremely tedious and expensive process, often requiring the use of diverse equipment ranging in price from hundreds of thousands of dollars for electron microscopy to millions of dollars for synchrotron X-ray facilities, and additionally requiring highly trained personnel to operate them. Addition-

ally, the characterization of materials is not limited to the complicated collection of data; additional data analysis is often required to extract the quantities or properties of interest from the imaging data. This extra layer of complexity involves manual analysis of generated data by experts, which increases analysis time and cost. Furthermore, the results obtained are often biased by the operator's judgment, compromising trustworthiness and repeatability of results.

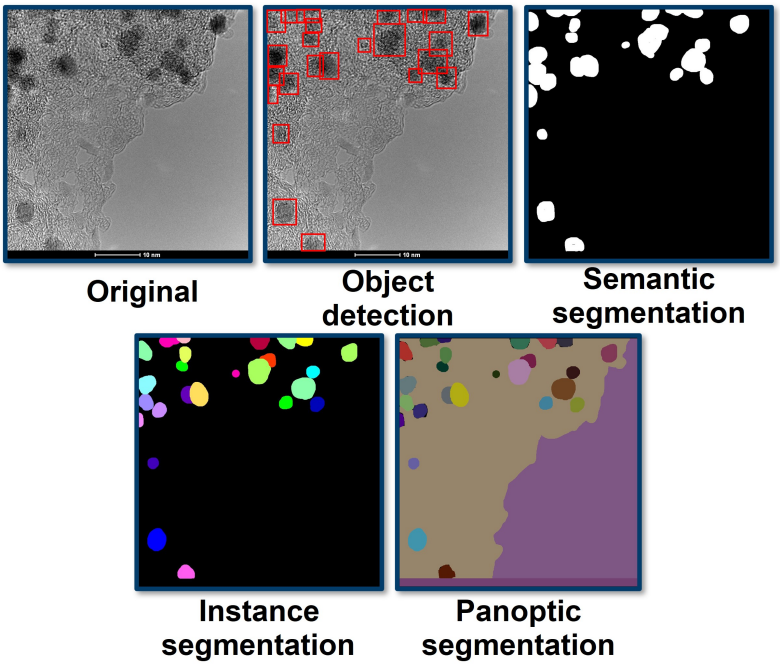

Figure 1.4: Overview of different segmentation methods.

In the pre-AI era, the complexity of collected imaging data was often too high to develop automated approaches based on hard-coded algorithms, leaving manual analysis as the only option to analyze the images. Recent data-driven developments in DL for image analysis using CNNs and visual transformers (ViTs) for pattern recognition are revolutionizing the field, as this technology can learn from human annotations and repeat the characterization procedure on unseen examples. Segmentation algorithms are of particular interest in the imaging domain. As shown in Figure 1.4, these models can identify regions of interest in images, and depending on the complexity of the analysis, it is possible to identify four types of methodologies:

- Object detection: The simplest and fastest algorithm is often used in real-time applications. It involves identifying and localizing objects in an image by drawing bounding boxes around them and assigning class labels.
- Semantic segmentation: A pixel-by-pixel classification of an image, where a label is assigned to each pixel that belongs to a particular class.
- Instance segmentation: A more complex version of semantic segmentation, where only one class of interest is considered. In addition to classifying each pixel, the model can distinguish between individual objects of the same class.
- Panoptic segmentation: This method combines the advantages of instance and semantic segmentation by providing a complete interpretation of the image. Each pixel is classified into a semantic level, with the addition of differentiating the individuals from the same class.

The speed and reliability of analysis provided by deep learning (DL)-based approaches are a promising alternative to the inefficient standard manual or hard-coded algorithmic analysis employed to date by experimentalists. Therefore, in recent years, a major focus has been placed on the automation of image analysis and the subsequent development of computer vision approaches to streamline the extraction of quantities and properties from imaging data via DL. Inside image analysis, diverse types of datasets require different DL approaches. For example, two-dimensional CNNs were employed at the level of the CL, where U-Net architectures have been used for three-dimensional nano-imaging of the catalyst layer, segmenting Pt/C nanoparticles in TEM volumes with high fidelity and enabling quantitative analysis.[63] To analyze GDLs, DL algorithms were applied to elucidate water clusters and transport pathways from computer tomography (CT) data, revealing wetting mechanisms in dry and flooded regions.[64] Even full-scale membrane electrode assembly segmentation algorithms were developed for the analysis of X-ray micro-CT scans by first increasing the resolution from 2.8 microns to 700 nm cubic pixels, or also called voxels, via DL super-resolution algorithms and then classifying the voxels into membrane, catalyst layer, microporous layer, GDL, pore space, and gas channel classes.[65]

DL has already been used to track and analyze dynamic events from video recordings automatically for *in situ* and *operando* studies. Faraz *et al.* developed a DL workflow combining registration, YOLO (You Only Look Once) architecture-based nanoparticle detection, and iterative tracking to extract trajectories and statistical descriptors of palladium nanoparticle evolution under reactive conditions.[66] Shen *et al.* introduced an end-to-end system for TEM video analysis of defect loops using YOLO with geometry and motion modules, achieving high performance on clean and structured datasets.[67] In

addition, Fu *et al.* applied LSTM models with feature disentanglement to predict future TEM images of Au nanoparticle transformations, demonstrating early-stage morphological prediction in dynamic catalytic environments.[68]

In 3D materials tomography, 3D DL architectures have been shown to exploit spatial context to improve automatic segmentation of volumetric materials datasets. For example, 3D U-Net models, adapted initially from medical imaging, have been trained on both experimental and synthetically generated X-ray tomographic microscopy images of graphite-silicon battery electrodes, accurately distinguishing active particles, binder, and pore phases where traditional thresholding fails.[69] Machine learning pipelines applied to PFIB-SEM tomography of thick cathode stacks use 3D U-Nets to extract individual particle morphologies and statistical descriptors of local heterogeneity and degradation pathways.[70] Emerging architectures such as the 3D Swin U-Net, which integrates Swin transformer blocks into a transformer-only encoder-decoder architecture, show promise for capturing long-range dependencies in porous media segmentation and outperform baseline U-Nets in complex microstructures.[71]

In summary, behind all of these advances are core architectures such as augmented U-nets,[35] volumetric V-nets,[72] hybrid transformer CNNs such as Swin U-Net,[73] and high-capacity backbones such as ResNeXt-101[74] that together form a versatile toolkit for tackling heterogeneous, high-dimensional imaging challenges in energy materials. These architectures were initially developed for other fields such as medicine and biology. Still, the overlap in visual features shared across fields allows for the transfer of these discoveries to materials science.

However, developing image analysis tools based on DL in energy materials has limitations. As the name of data-driven approaches implies, the available data plays a crucial role in developing AI solutions. This becomes the main issue in the energy materials field, where the data type obtained is field-specific, the amount of data is limited, and often comes with certain sharing restrictions, such as industry contracts and non-disclosure agreements (NDAs). The combination of these two barriers results in a slowed development of energy materials-related image analysis approaches and the obligation of direct collaboration with experimental partners, which limits the independent development of tools for the community. In addition, inefficient data storage in the experimental community further complicates the discovery and sharing of large amounts of data, even when it is potentially available.[75]

1.4 Databases and repositories for imaging data

As discussed in the previous section, data are central to accelerating the investigation of energy technologies using AI-based approaches. The heterogeneous, large, and complex nature of imaging data from material samples increases the difficulty of storing all data in specific databases or repositories in a standardized way for shared access.[76,77] In addition, the value of imaging data is closely tied to the context, in which they were acquired. For this reason, rich, standardized metadata protocols covering data precedence, fabrication procedures, and measurement conditions are fundamental for properly interpreting stored data.[78,79]

In reality, the majority of research groups still use outdated, inefficient, and disorganized folder-based data structures to store their data, which often results in the loss of a large part of the knowledge after the student or staff member responsible leaves the group, with crucial consequences for the development of data-driven models, correlative data-based analysis, and costs.

To address these issues, specialized databases are needed to store raw imaging data and ensure that data and metadata are captured in standardized and accessible formats. Several entities have attempted to address this issue by creating specific repositories for scientific materials imaging data. For example, the Materials Data Facility,[80] is a community-driven data ecosystem that allows publication of materials science data with special emphasis on interoperability, assigning unique identifiers to each dataset to track and find the corresponding metadata. Still, it lacks widespread standard protocols for storing data from different sources. TomoBank is another interesting initiative to create a repository specialized only in tomography datasets,[81] especially those obtained from synchrotron X-ray sources for high-quality 3D data. This example remarkably addresses metadata standardization for 3D synchrotron datasets, at the cost of applying it to only one characterization method. EMPIAR is the most prominent database used by the community that exploits electron microscopy data.[82] It began as a database for biological electron microscopy, but its widespread use has opened it up to other disciplines such as materials science. This initiative also supports standardized metadata protocols and file formats, at the cost of collecting only measurement metadata and only for electron microscopy data. Finally, Materials Cloud is a public database created by ETH Zurich that hosts a wide variety of materials data, from experimental imaging data to computational simulation data, focusing on bridging these experimental and computational workflows in one place for seamless interoperability and visualization of the stored data.[83]

As mentioned above, these databases constitute a significant step towards accessible and standardized image data storage. Still, the self-imposed restrictions limit the wider us-

ability of the different approaches in a general way. Therefore, the areas requiring further development to achieve the standardization needs are:

- Standardized file formats such as OME-TIFF for microscopy data or HDF5 for multi-dimensional datasets to ensure interoperability and consistency across datasets.
- Metadata harmonization across all characterization methods ensures that the imaging data's original context is preserved. This will require community consensus on what information is essential to record along with the image data to ensure its reusability.
- Seamless integration into AI pipelines for accelerated discovery and training of machine learning approaches. The key factor in achieving these points is the implementation of programmatically accessible interfaces or API protocols to accelerate data acquisition and subsequent storage in a machine-readable format.

Therefore, to fully realize the potential of standard and AI-enhanced systems, further efforts are needed to bridge all the different components under a unified paradigm and provide the community with a more seamless, collaborative, and innovative approach to accelerate tomorrow's energy storage challenges.

1.5 Scope and goal of this work

As mentioned in the introduction to this thesis, there is an urge for society to move towards green energy conversion and storage. Still, the actual challenges are not being addressed as quickly as needed. This is where AI-accelerated analysis comes in to help researchers perform more reliable and faster image analysis. Nevertheless, it should be pointed out that there is still a gap to be bridged in developing specific DL models to analyze imaging data from energy materials. At the same time, there is an urgent need to introduce a harmonized and standardized platform to store experimental data in an efficient and accessible way.

Even though solving all these problems looks like a gigantic task, during this work, it was tried to tackle these issues and advance in the research towards creating a data-driven environment for accelerated energy materials development. This dissertation results from involvement in the project entitled UTILE: aUTonomous Image anaLysis of Energy materials. The UTILE project provided the perfect conditions to combine the expertise of several experimental researchers specialized in different energy technologies and diverse imaging methodologies with specific automation needs and the corresponding data. The

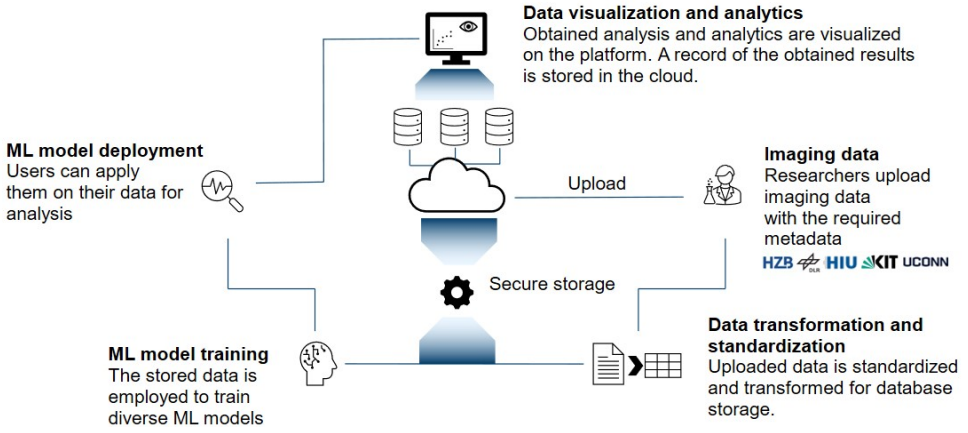

Figure 1.5: Schematic of the UTILE platform and its capabilities, from image data storage in a standardized environment to machine learning model training and inference directly from the cloud platform.

scope of this work is illustrated in Figure 1.5, where the goal of the UTILE platform is to develop automated DL solutions for challenging cases and finally, with the acquired knowledge, close the loop by improving the imaging data storage for future AI model development.

The concept of the final cloud-based platform is to facilitate streamlining the data collection process for experimentalists by providing standardized metadata protocols to capture all essential information around imaging data and ensure that the context of data is captured. Then, data will be stored in a secure cloud-based server to allow collaborators to access the data and models remotely. Once the data are stored according to the platform standards and in a machine-readable form, it will be easier to train AI models to predict properties and find correlations. Afterwards, deploying the trained models in the platform and visualizing the results on the fly should be possible.

After a detailed explanation of the methodology used during this work in **Chapter 2**, the various efforts to automate scientific image analyses and standardize the imaging data collection will be described. The research work was carried out in four subprojects, corresponding to the four main chapters of this thesis, as shown in Figure 1.6.

- **Chapter 3: UTILE-Oxy** deals with automated optical video analysis of oxygen bubble dynamics in PEMWE flow fields. Here, DL is used to perform semantic segmentation of video recordings of a transparent cell to elucidate the exact influence of voltage changes on bubble formation and to assert correlations with parameters and flow field pattern choice. In addition to the DL bubble detection algorithm,

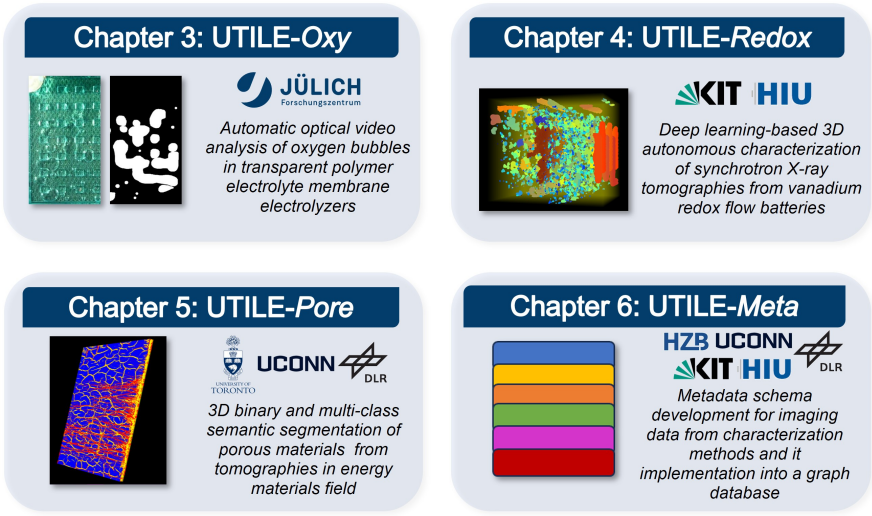

Figure 1.6: Presentation of the different subprojects treated in this thesis, forming the UTILE project.

several computer vision tools have been developed to quantify certain characteristics, such as bubble cell coverage, bubble density maps, and even individual bubble shape analysis.

- **Chapter 4: UTILE-Redox** entails a multi-class semantic segmentation approach to analyze the 3D structure of hydrogen bubbles inside the negative half cells of VRFBs from synchrotron X-ray tomographies. For this purpose, DL architectures were benchmarked to classify the cell into four classes: bubble, membrane, electrolyte, and gasket classes. Then, a set of computer vision functions was developed to perform a deeper analysis of the bubble distribution in 2D and 3D and individual 3D bubble shape analysis. The software's performance was tested by analyzing the correlation between parasitic hydrogen evolution reactions and variations in electrode material.

- **Chapter 5: UTILE-Pore** is an improved DL-based methodology, developed for the analysis of 3D porous materials in CLs, GDLs with MPLs from micro-computed tomography (micro-CT) and focused ion beam scanning electron microscopy (FIB-SEM). A series of 3D DL models were benchmarked on a mixed data set with physical porosity calibrated masks. Therefore, a binary and a multi-class approach were developed to analyze CL and GDLs and allow the differentiation of MPLs from GDLs. For further analysis of the obtained segmentations, several computer vision

tools have been designed to extract porosity, tortuosity, and pore size distribution from the CLs and GDLs, and functions for understanding the interplay between the MPLs and GDLs.

- **Chapter 6: UTILE-Meta** synergies knowledge acquired by working closely with the experimental partners into a metadata strategy to capture all the details of the different image characterization methods in a standardized schema based on the collective expertise of the project collaborators. This approach is then implemented in a graph database with AI-guided context recognition and information organization to ensure correct classification of potential user input.

Lastly, the main findings and future outlook of the projects presented for developing DL tools for image analysis and the subsequent standardization of the imaging data storage for improved AI model training are summarized in **Chapter 7**.

Chapter 2

Theory and Methodology

2.1 Data, metadata, and databases

As stated in the introduction, this work focuses on data-driven acceleration in the investigation of energy materials. Central to this topic is the word data, which is widely employed across all fields, but usually no explicit definition is given. Therefore, in this section, the definition of data with the needed connotations to understand this work is introduced and explained with a special focus on imaging data and metadata.

2.1.1 A definition of data and its types

Data are the central pillar in this work and, therefore, must be adequately introduced for a general understanding of the upcoming sections. A definition of data that fits the use in this dissertation might be the one following the statement given by William Kent: "Data refers to the raw facts and figures that are collected, stored, and processed to extract meaningful information". This definition was coined in the groundbreaking, philosophical book "Data and reality: a timeless perspective on perceiving and managing information".[84] More focused on the scientific context applied here, data can come from various sources such as experiments, observations, or simulations of real-world phenomena. The manifestation of data depends on how they were generated, collected, and stored, and can even span multiple dimensions.

1D data are the most common type of data. They include information that is organized in a linear fashion, such as string data. String data are defined as a sequence of characters that, when organized, can represent text, labels, or identifiers. Spanning numbers or characters in 2D, they become tabular data typically organized in rows and columns,

similar to the familiar spreadsheet or comma-separated values (CSV) file. These forms of data are the basic principles for storing and processing simple, sequential information.[85]

2D data also include image data, representing information in a grid of values called pixels. In grayscale images, each pixel in the grid is represented by a value that describes a property of the image, such as an intensity. For colored images, each pixel spans into an extra dimension composed of an array of values to describe colors such as red, green, and blue (RGB) channels. Depending on how the image is encoded, the grayscale or color can vary, allowing for complex grayscale or color combinations as the bit size increases. For example, the minimal 8-bit pixel can store 256 different intensities with values ranging from 0 to 255, usually 0 being black and 255 being white. In the case of RGB images, where color depends on the combination of three pixel intensities, it is possible to represent 16.7 million colors at the most basic encoding level. However, by increasing the image bit depth from 8 bits to 16 bits, it is possible to obtain 65,536 different intensity levels at the cost of increasing the storage cost of the image. Digital images can be stored in various formats for reading, processing, and analysis, such as PNG, TIFF, JPEG, etc. However, some formats compress images, which can result in information loss due to the image format, such as JPEG. Other formats, such as TIFF files, offer lossless compression and are the most widely used format for storing scientific images.[86–88]

3D imaging data are conceived as an extended concept of stacked 2D images that expands the collected information into a third dimension. This creates volumetric representations that include depth as an additional dimension, but is not limited to the spatial domain, as the third dimension can also be a time-resolved representation of a system, such as video. Volumetric data are often generated by techniques such as computed tomography (CT) or 3D microscopy in three spatial dimensions, or by 2D techniques taken sequentially at time intervals, such as optical videography or radiography for the time dimension. 3D data are beneficial for elucidating complex 3D structures in materials or time-resolved phenomena such as gas evolution or degradation processes. For example, increasing the spatial dimensions of the images into volumes usually increases the richness of information. It expands the understanding of the spatial relationships extracted from a structure, but at the cost of a cubic increase in the volume of data that must be processed and stored.[89–91]

Furthermore, 4D volumetric data integrates two classes of 3D data into a single entity. It extends the 2D data in the third spatial dimension and captures the time evolution of the volume. This advanced dynamic data type is not widely used due to the limited number of methods capable of achieving this type of data acquisition. The trade-off between the time cost of measuring a single 3D structure and the minimum resolution quality required to capture necessary details for analysis severely hampers this approach. This technology

is a promising next step in developing imaging methods, as it can simultaneously elucidate spatial and temporal properties. This is particularly useful, for example, for gas volume evolution in electrochemical cells or for further investigation of degradation mechanisms in real time with 3D relations.[92,93]

2.1.2 Metadata, data about data for contextualization

Metadata are referred to as data about data. They aim to provide essential information about the datasets to improve their understanding and interpretation. They often include details about the data's origin, structure, and parameters, such as acquisition methods, calibration settings, and experimental conditions, as shown in Figure 2.1. This additional layer of description clarifies the meaning behind the raw data and ensures the reproducibility and interoperability of datasets over time.[94]

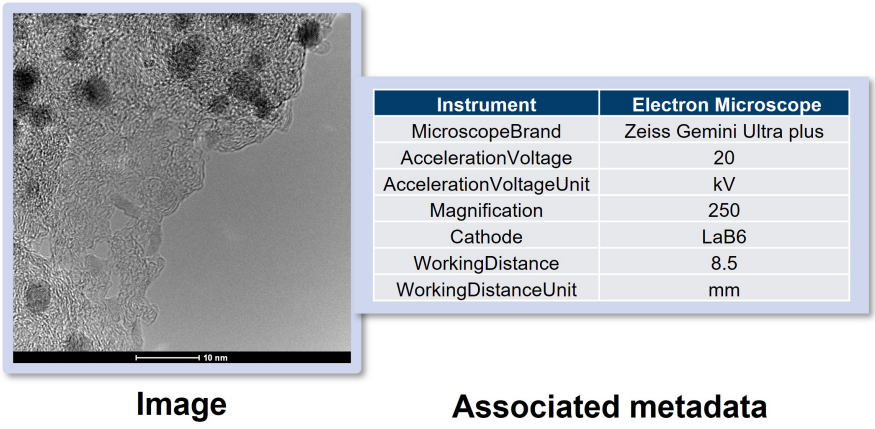

Figure 2.1: Showcase of a transmission electron microscopy image with the corresponding instrumental metadata.

Collecting robust metadata along with experiments is fundamental for data discovery, integration, and the development of AI solutions, particularly in the scientific domain. However, even though metadata collection is crucial, its methodology is highly dependent on the field of study, the purpose of the data, and the equipment used. For example, there is no specific way or guidelines to define and collect metadata, and these parameter dependencies raise important issues in the current information age. To address these problems, certain metadata schemas have been developed to facilitate the seamless exchange and long-term preservation of data, such as Dublin Core or ISO.[95,96] Metadata schemas can only be considered standards when the community accepts and integrates them widely.[97]

We can categorize different types of metadata based on the type of information they represent:

- Descriptive metadata typically provide general information about the identification and context of the data, such as sample analyzed, author, institution, etc.
- Administrative metadata capture technical aspects of the file itself, such as file format, file size, and creation details.
- Structural metadata explain the relationships between different parts of the dataset.
- Field-specific metadata provide the context, in which the data have been generated and help understand provenience and experimental details for future usage.

Furthermore, metadata are only helpful if they are continuously collected and stored between datasets in an effective data management strategy. Moreover, identifying the essential metadata to be collected is primordial and has to be guided by the well-known FAIR guidelines. The so-called FAIR principles are used to ensure the data strategy's usefulness.[77] FAIR stands for Findable, Accessible, Interoperable, and Reusable and FAIR principles aim to support researchers in creating metadata strategies. The different terms are defined as follows:

- The Findability principle focuses on the fact that data should be easy for humans and machines to find. This principle is often implemented by assigning traceable, unique identifiers to each record, such as a DOI, and by enriching metadata with clear descriptions of the data, allowing seamless data integration and retrieval from large databases.
- Accessibility means that data, or at least metadata, should be easily retrieved from the database once found. For this reason, standard extraction protocols are recommended in the various database paradigms, such as REST API protocols, which allow automated data extraction from databases using standard tools.
- Interoperability is one of the FAIR principles' most critical and complex aspects. It describes the ability to use a common language to define metadata, allowing datasets from different organizations to communicate and facilitate discovery. Standard vocabularies, ontologies, and formats are preferred for cross-disciplinary, inter-institutional, and automated data processing. Today, great efforts are being made in this field to standardize different practices effectively, with the added difficulty of substantial domain variation and reuse of terms by other fields and even different institutions.

- The notion of Reusability is the ultimate goal of a FAIR data structure. It requires rich metadata collection for replication, verification, and re-use of the datasets in future research. This principle is also linked to clear data licensing, detailed origin, and deep contextual information about how the data was generated.

In summary, by following the FAIR Principles, researchers and institutions can improve their data management practices' transparency, reliability, and sustainability, ultimately contributing to a more robust and connected scientific community. However, while the principles are straightforward to follow, they are complex to implement. On one hand, findability and accessibility are relatively easy to achieve by using suitable technologies and descriptors. On the other hand, achieving a standard protocol to enable interoperability and future reusability of results is more complicated. It requires high expertise in different areas related to the scientific and programmatic fields.[77]

2.1.3 Organizing knowledge in ontologies

The previous section raises a logical follow-up question: how can an interoperable, standard metadata schema be developed? One possible solution is to develop a domain-specific ontology, creating a common language that researchers in the same domain can adapt to ensure interoperability within a database. An ontology is defined as a formal representation of knowledge in a particular domain, with a graph-like structure in which domain-specific terms or concepts can be represented as classes. These classes can have specific instances and can be connected through relationships. Attributes define classes, and rules and constraints govern their interaction.[98]

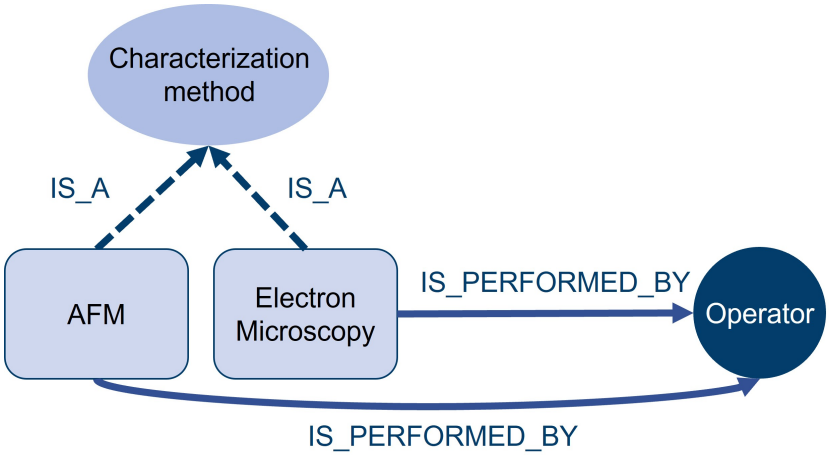

Figure 2.2: Simplified version of an ontology for explanation purposes.

As shown in Figure 2.2, hierarchical (parent-child) relationships allow detailed descriptions of concepts by adding levels of specificity. For example, the parent node "characterization method" might have a child node "electron microscopy". Typically, child and parent classes overlap in their attributes and are connected by IS_A relationships, while the child class often has additional attributes such as subclasses or characteristics. Cross-relationships, on the other hand, connect nodes that are not in the same hierarchical branch. For example, the "electron microscopy" node can be connected to an "operator" node via an "IS_PERFORMED_BY" relationship, even though "operator" is not a child of "characterization method". Using hierarchical and transversal relationships allows us to describe complex knowledge domains necessary to represent scientific experiments accurately.

The advantage of using an ontology lies in the resulting standardization of the terminology, which is fundamental to achieving interoperability as described by the FAIR principles. However, proper standardization can only be achieved if a large community agrees to use the same ontology.

For this reason, the ontology presented in this work is an extension of the EMMO, a European effort to standardize the interdisciplinary nature of materials science under a unified ontology.[99] EMMO has three levels of complexity that are visualized in Figure 2.3 and given by:

- Top level: An abstract layer that defines an object and introduces philosophical concepts such as time and space.
- Middle Level: Represents different perspectives (e.g., holistic, physicalistic) and describes how objects are characterized within each perspective. For example, the holistic perspective distinguishes between objects and processes to describe knowledge.
- Bottom level: The domain-specific layer that can be extended and developed by each discipline within materials science.

Relevant projects that are based on the top and middle levels of EMMO and have extended the bottom level include CHAMEO[100] for materials characterization, BattINFO[101] for battery knowledge description, and NOMAD,[102] a well-known repository for modeling data. EMMO is also being employed in other fields, expanding the standardization horizons towards other disciplines such as nuclear energy, circuit modelling, and chemical substances.[103] The widespread use of the EMMO ontology by different projects facilitates interoperability between different data repositories, since they share the same basic

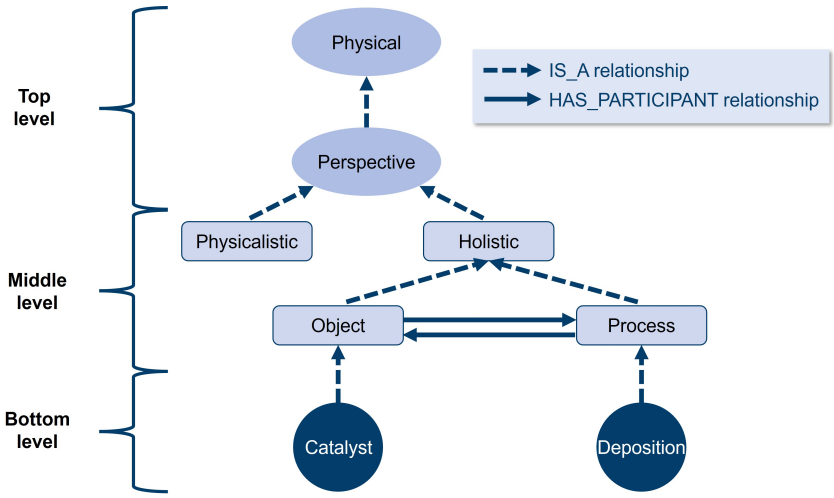

Figure 2.3: Minimalistic description of the different levels in the EMMO ontology. The visualization showcases the different instances from the philosophical top level with the definition of perspective, through the diverse instances of perspective in the middle level, to the domain instances in the bottom level.

structure for describing knowledge. It also allows future projects to reuse the extended branches where parts of the knowledge graph overlap.

2.1.4 Organizing data in databases

After developing a domain-specific ontology, the next logical step is to implement the ontology as the underlying structure of a database. In this way, it is possible to start storing data under a common language and to facilitate data retrieval according to the specified terminology. The choice of database greatly influences the flexibility and complexity of the information that can be stored. Databases can be broadly categorized into two main types based on how they store data: relational and non-relational.

As shown in Figure 2.4, relational databases, such as PostgreSQL[104] and MySQL,[105] are industry standards for storing data in tables. Each table represents a class of real-world objects, such as "characterization methods", and the rows within a table represent instances of those objects, such as "optical microscopy". Relationships between tables are established using foreign keys, which act as links between entries. While relational databases are advantageous for highly structured data with few links between tables, they become less efficient as data becomes more interconnected and dynamic. Increased interconnectivity and dynamism in stored data result in slower retrieval and more complex database structures.

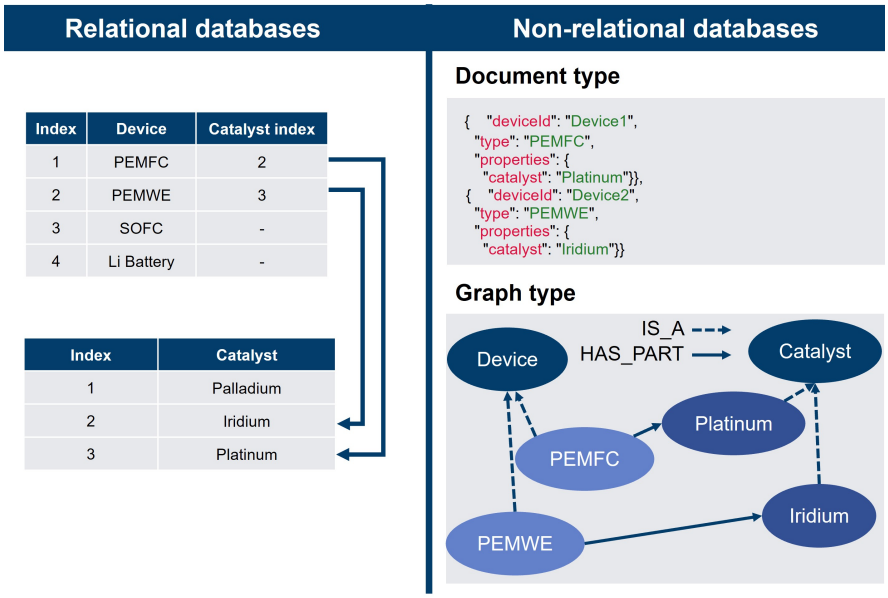

Figure 2.4: Schematic description of the different types of databases and their storage paradigms.

In contrast, non-relational databases excel at handling more dynamic data models. Two common types of non-relational databases are document-oriented databases and graph databases. Document-oriented databases, such as MongoDB,[106] use flexible document structures such as YAML and JSON. Each document is described by unique keys that can adapt to dynamic and customized data structures, as depicted in Figure 2.4. However, they still rely on the same principle as relational databases to connect different documents: using foreign keys. This reliance makes it less effective for highly related data points.

Graph databases, such as Neo4j,[107] are optimal for efficiently storing highly dynamic, flexible, and complex data. These databases use the principles of graph theory to store data as nodes connected by relationships. The nature of graphs allows us to represent real-world domains directly and intuitively, without additional layers of abstraction in the database implementation. Graph databases are specifically designed to efficiently query interconnected data, allowing data extraction with constant time complexity (O(1)), which means that the data extraction runtime required is not dependent on the size of the database, making finding information significantly faster across huge databases.

In addition, graph databases store data as key-value pairs within each node, and each relationship is directed. Directed relationships represent asymmetric relationships and provide more context, such as describing a process with directed relationships like ”pro-

cessedTo". In the context of this work, graph databases are ideal for describing experiments, manufacturing workflows, or characterization processes. These processes involve a wide range of different parameters and complex sub-processes that can be naturally represented as graphs, consisting mainly of key-value pairs stored in nodes or sequences of steps where a manufacturing technique transforms an input material into a material or property output. In addition, the flexible nature of graphs lends itself to the dynamic and ever-changing structure of processes, where properties, steps, or parameters can vary from experiment to experiment, without the need to predefine what the structure of a process should look like.

2.2 Deep learning on imaging data

Once the general concepts of data, knowledge, and storage are established, the basics of artificial intelligence (AI) are introduced. Starting from the definitions and differentiation of AI, machine learning (ML), and deep learning (DL), the working principle of DL is subsequently explained. Furthermore, the imaging-specific DL approaches of convolutional neural networks (CNNs) and visual transformers (ViTs) are introduced.

2.2.1 Artificial intelligence, machine learning, and deep learning

As shown in Figure 2.5, AI is a broad term used to describe various phenomena in computer science related to the ability of computers to "think". This term has evolved as newer techniques have been developed, extending the ability of computers to imitate human behaviors such as reasoning, learning, or even seeing. Today, this field has become deeply entrenched in society, achieving growing maturity and permeating every scientific and non-scientific field. For this reason, a comprehensive definition of AI in the current context would be the effort to automate intellectual tasks normally performed by humans. As such, AI is a broad term encompassing several different subfields, not all of which are directly related to trendy machine learning algorithms or neural networks. The popular connotation of AI related to "machine learning" is a relatively new concept not present in the early days of AI in the 1950s. As an example, hard-coded Canny edge detection algorithms are considered AI if the programmers could handcraft a sufficiently large set of explicit rules that imitate human-level artificial intelligence in a particular domain.[108]

Hard-coded AI, or symbolic AI, quickly reached its limits when faced with tasks of increasing complexity, more fuzzy problems, and many new exceptions that could not be predicted in advance. Tasks such as image analysis, speech recognition, or natural language translation would require an infinite number of exception handling statements in

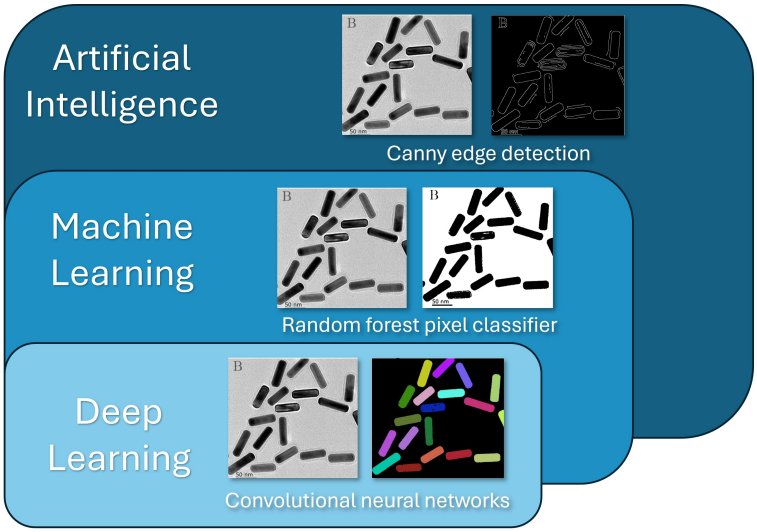

Figure 2.5: Schematic visualization of the relationship between artificial intelligence, machine learning, and deep learning, and the increasing complexity of tasks that can be solved. A corresponding algorithm represents each category. Canny edge detection is considered artificial intelligence, but it is not capable of learning features; rather, it performs edge detection based on adaptive thresholds. Random forest pixel classifiers learn from a dataset to classify colors in the image. However, they are unable to distinguish objects of interest from other aspects of the image that have similar colors. Finally, deep learning algorithms based on convolutional neural networks can detect only the objects of interest, as a human would.[62,109–111]

code to cover all possibilities, or simply the complexity of the data cannot be rationalized by a human to implement in a programming language. To solve this problem, a new branch of AI was created: machine learning (ML).[108]

However, to create it, the conceptual issues of symbolic AI should be understood. As shown in Figure 2.6, the usual method of solving a problem with classical programming is for a programmer to write rules to create a computer program that applies them to input data to obtain a particular answer. This approach assumes that the programmer is aware of the rules that need to be implemented and is humanly capable of coding all of those rules, which leads to the problems mentioned above. This is where machine learning comes in with a new paradigm. Instead of generating an answer based on input data and rules, the algorithm is left to figure out the rules to follow by itself by providing the input data and the expected answer. This is how machine learning is trained. The ML algorithm is provided with many examples with the expected answer, and it finds statistical structure in these examples that eventually allows the system to infer rules for automating the task.[108]

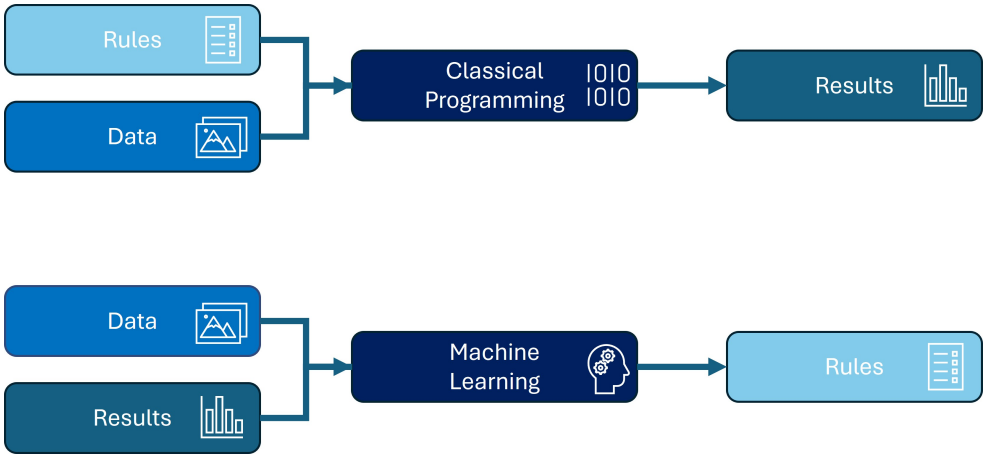

Figure 2.6: Visual description of the logical working principle of classical programming and how it differs from the working principle of machine learning.[108]

ML is related to mathematical statistics, but differs in several ways, particularly in the systems it works on. ML is applied to large and complex systems, ideally with millions of data points, for which classical statistical approaches such as Bayesian analysis would be impractical. Furthermore, ML has a small mathematical theory behind it compared to established statistical models, making the field an engineering discipline. In contrast to theoretical mathematics or physics, these aspects mark machine learning as heavily dependent on empirical discoveries as well as software and hardware advances to be further developed.[108]

Since the main focus in this work is on image data analysis, the following explanations are tailored to the application of ML to image analysis. To use ML to solve any imaging problem, it is necessary to connect four main parts:

- The input data, in this case images, is used to represent the use cases for teaching the model.
- The labels are examples of what is required to extract from the input data, such as the pixel-by-pixel location of objects of interest.
- The model architecture to train.
- The metrics to measure whether or not the trained ML model is predicting accurately on the input data. This is necessary to analyze the model's performance and correct it if necessary. Correcting an ML algorithm is what is called learning.[108]

With these four aspects, it is possible to develop an ML algorithm capable of solving a given task based on the input data, the expected output of the model, and a metric to assert the correctness of the predicted output by the model. Based on the type of labels used, there are three relevant learning procedures to train an ML model to predict on a task: supervised learning, unsupervised learning, and semi-supervised learning.

As shown in Figure 2.7, supervised learning is the most common learning method and the one used in this thesis. Supervised learning algorithms rely on the training of models based on a large human-labeled dataset. In the classification example, if the task is to differentiate TEM images of nanoparticles by particle shape, the images are labeled by the shape of the particle they contain, e.g., cubic-, round-, and rod-shaped. The images with the correct labels are then fed into the model, and it learns to discriminate between them and apply this knowledge to new examples.[112]

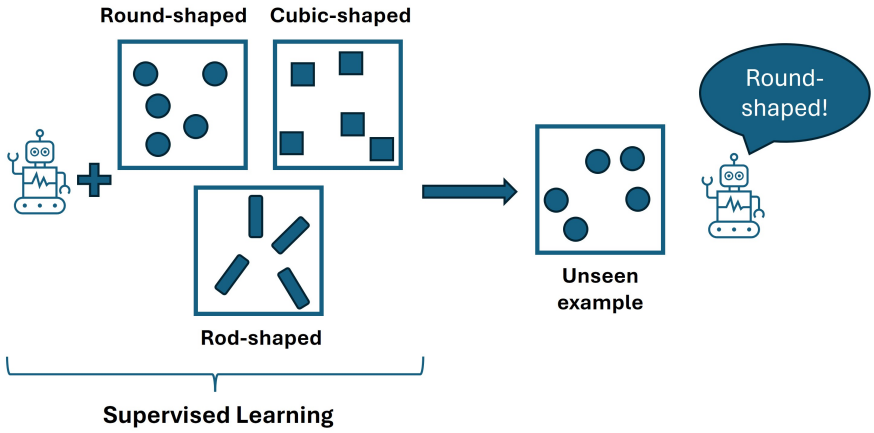

Figure 2.7: Simplified description of the working principle of supervised learning for the image classification of different particle shapes.

In unsupervised approaches, the models are fed with unlabeled examples by manually specifying certain constraints as expected output, and the algorithm tries to find correlations or similar patterns within the dataset, as shown in Figure 2.8. This approach is useful when the dataset is not labeled beforehand, or when the user wants to discover a new correlation within the data. A typical unsupervised learning technique is data clustering. In the example of images of TEM images of nanoparticles with different shapes, the unsupervised learning approach would be fed with the images of differently shaped nanoparticles and by a given number of clusters, the model would identify the different groups and categorize the images into the corresponding buckets.[112]

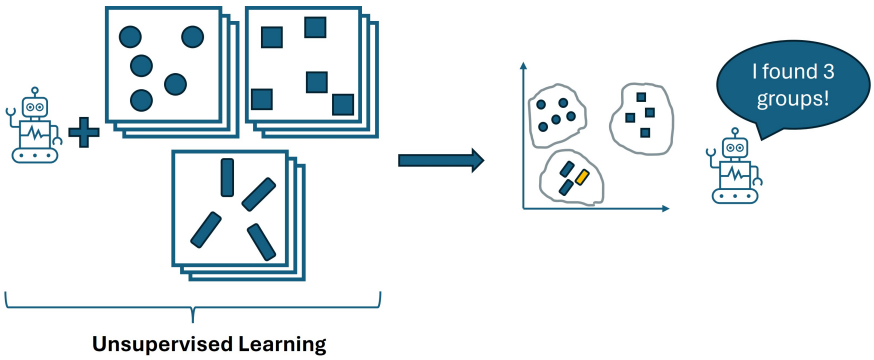

Figure 2.8: Simplified description of the working principle of unsupervised learning for the image clustering of different particle shapes.

Semi-supervised learning approaches combine the supervised and unsupervised procedures, as shown in Figure 2.9. In the above example of nanoparticle TEM images, manual labeling can be costly and time-consuming for large datasets, so a logical solution would be to annotate the dataset and then cluster the images partially. Depending on the labels of the annotated images in each cluster, the labels can be extrapolated to the unlabeled members of the same cluster.[112]

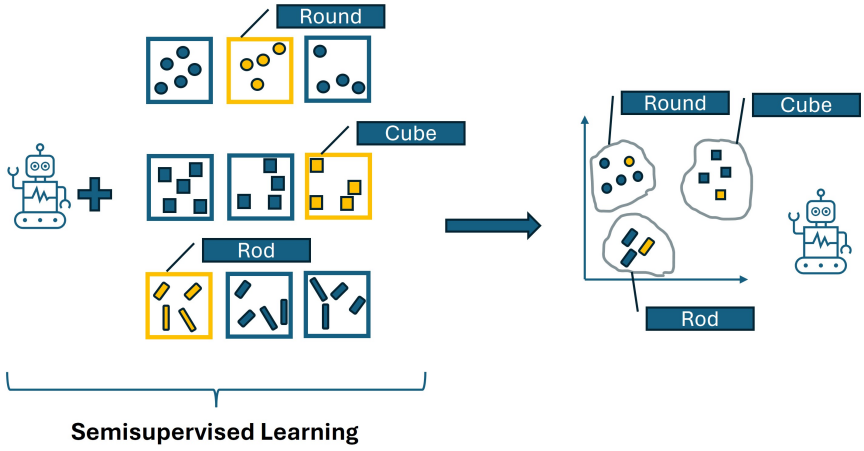

Figure 2.9: Simplified description of the working principle of semi-supervised learning for the image clustering and subsequent classification of different particle shapes.

ML has embraced a vast landscape of methods that fall under its definition, and deep learning has emerged as the most successful branch of it, often leading to an interchange of the two terms. By definition, DL is the branch of ML that uses artificial neural networks (ANNs) to learn hierarchical representations from raw data automatically. To untangle

this definition, it is first needed to explain what artificial neural networks are, how they work, and what distinguishes them from DL.[112]

2.2.2 Teaching a machine to learn by emulating a human brain

Artificial neural networks (ANNs) are algorithms that attempt to mimic how the human brain works by processing combinations of different stimuli into an output. As the name suggests, ANNs are interconnected neurons arranged in layers that transfer information from neuron to neuron, layer by layer, to produce an output. To understand how ANNs work in detail, it is possible to start with the simplest ANN, which consists of a single building block: the neuron. A single neuron is also called a perceptron. A perceptron is a computational unit that takes a set of inputs x_i, linearly combines them with learned weights w_i and a bias b, and then passes the result through a non-linear activation function f to produce an output y, as shown in Figure 2.10. Mathematically, this is expressed as

$$y(x_i) = f\left(\sum_i w_i x_i + b\right) \tag{2.1}$$

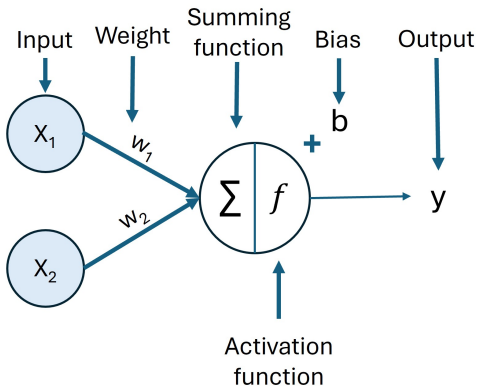

Figure 2.10: Depiction of the structure of a perceptron and its components.

The inputs x_i are the features or data points passed through the perceptron. At the same time, each weight w_i is associated with a particular input x_i and indicates the importance of the input in the decision step. A bias parameter b is often added to the multiplication of the weight w and the input x to allow the model to shift the decision boundary towards a better fit to the data. Finally, the activation function is applied. This nonlinear function takes the result of the previous calculations and determines the final output of the model, allowing the model to learn complex patterns. In the classic perceptron model, it is Boolean for a binary classification.

A simple perceptron model is limited in the complexity of information it can capture in its weights and biases to correctly predict an unseen input. Therefore, to handle more complex tasks, it is possible to have multiple neurons in parallel and organize them into layers stacked one on top of the other, as shown in Figure 2.11. The layers are organized according to a simple scheme: the first layer of neurons is called the input layer, where the raw input data is introduced into the model. Depending on the case, a variable number of middle layers can be added, called hidden layers, interconnected with the previous layers' output as input. Finally, the last layer is called the output layer and is responsible for outputting the final prediction of the model on the processed input data. This layer often has the same number of neurons as the possible expected outcomes. For example, in the case of binary classification, the output layer often has two neurons to predict whether the input data is 1 or 0.[112]

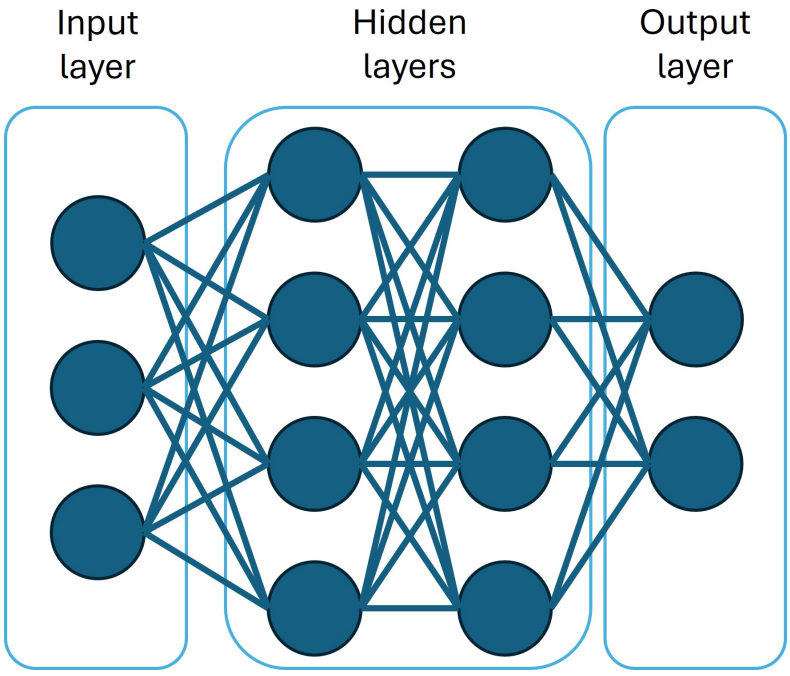

Figure 2.11: Schematic representation of the structure of an artificial neural network.

By definition, DL is the branch of machine learning that focuses on deep neural networks (DNNs), which are all ANNs with at least one hidden layer between the input and output layers. Stacking layers of interconnected neurons allows the model to extract "deeper" and more complex representations from raw data.

The mechanism of prediction on data with a DNN is called feed-forward, and can be easily explained using a four-layer network with one neuron per layer for simplicity, as shown in Figure 2.12.[112]

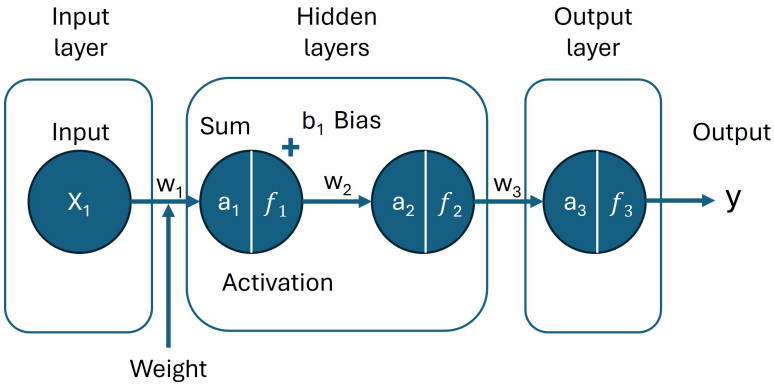

Figure 2.12: Schematic description of the building blocks of a simple multi-layer perceptron with four nodes.

The prediction mechanism of a DNN can be explained as simply as the nested prediction mechanisms of several individual perceptrons, one behind the other. Following Equation 2.1, it is possible to derive the nested equations as follows:

$$a_i^{(1)} = \sum_j w_{i,j}^{(1)} \cdot x_j + b_i^{(1)}, \tag{2.2}$$

$$z_i^{(1)} = f^{(1)}\left(a_i^{(1)}\right), \tag{2.3}$$

$$a_i^{(2)} = \sum_j w_{i,j}^{(2)} \cdot z_j^{(1)} + b_i^{(2)}, \tag{2.4}$$

$$z_i^{(2)} = f^{(2)}\left(a_i^{(2)}\right), \tag{2.5}$$

$$a^{(3)} = \sum_j w_j^{(3)} \cdot z_j^{(2)} + b^{(3)}, \tag{2.6}$$

$$y = f^{(3)}\left(a^{(3)}\right), \tag{2.7}$$

Where x_j are the input features, $w_{i,j}^{(l)}$ are the weights connecting neuron j in layer $l-1$ to neuron i in layer l, $b_i^{(l)}$ is the bias of neuron i in layer l, $a_i^{(l)}$ is the pre-activation (i.e., the weighted input sum before applying the activation function), $z_i^{(l)}$ is the post-activation (i.e., the output of the neuron after applying the activation function), and $f^{(l)}(\cdot)$ is the activation function applied elementwise in layer l.

As described by equations 2.2–2.3, the first hidden layer of the network operates similarly to a single perceptron. Each input feature x_j is multiplied by a corresponding weight $w_{i,j}^{(1)}$, and a bias term $b_i^{(1)}$ is added. The result is the pre-activation $a_i^{(1)}$ for each neuron i in the first hidden layer. This value is then passed through a nonlinear activation function $f^{(1)}$, yielding the post-activation output $z_i^{(1)}$, which serves as the input to the next layer. This process is repeated layer by layer according to equations 2.4–2.7, until the final output y is produced by the output layer. The value of y represents the model's prediction.

The feed-forward pass is only valid after a model has been trained on a specific task. Training a DNN model on a supervised task involves adjusting the weights and bias values to enable the model to interpret the input data and make accurate final predictions based on prior experience. The standard way to optimize the weights and bias is an iterative process consisting of two main steps repeated until a specific performance is achieved. These two steps are the already explained feed-forward pass and a subsequent backpropagation step.

In a general case, for the training of a model with a given set of examples x_i, their corresponding correct labels y_i, and a standard DNN, it can be described by

$$y = f_{\text{DNN}}(x; \theta), \tag{2.8}$$

with

$$\theta = \{f, w, b\}, \tag{2.9}$$

it is needed to estimate the best parameters of the network for the unknown weights w and biases b based on the known given data set x_i and y_i. The training loop is executed as follows:

1. Feed forward pass: The raw data x_i is fed through the model with randomly initialized weights and bias. After the data has been randomly processed, the output of each node is stored.

2. Loss Calculation: The output of the model y_i is compared to the correct label $\hat{y}_i$, and the error is calculated using a specified loss function. A variety of different loss calculation methods are used in the literature. A simple example is the mean square error (MSE), given by

$$L = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 . \tag{2.10}$$

A typical state-of-the-art function to calculate the loss, used in this work, is the cross-entropy loss calculation given by

$$L = -\frac{1}{N} \sum_{i=1}^{N} \left[y_i \lg\left(\hat{y}_i\right) + (1 - y_i) \log\left(1 - \hat{y}_i\right) \right] . \tag{2.11}$$

This loss function penalizes confident but incorrect model predictions more than less confident ones.[112]

3. Optimization Step: With the calculated error, it is necessary to adjust the model's parameters so that the predicted label obtained by the model in the next loop is closer to the true labels. Weight optimization is usually done by gradient descent. This calculates how changing the weights affects the loss function and adjusts the weights to minimize the loss.

As shown in Figure 2.13, if it were possible to compute all infinite points of the loss function as a function of weights, it would be possible to observe a 3D surface representing all minima and maxima of the function. Since it would be computationally costly to compute all the alternatives to the weights, it is required to move the weights toward the minima following the slope blindly based on the gradient slope. When data is passed through the model and the loss is calculated with random weights, one point of the surface is revealed while the rest remains unknown. This is where the iterative computation of the gradient descent and the stepwise shifting of the weights towards the closest minima opens an efficient way to minimize the loss function of the model in the specific task.[112]

The principle of gradient descent is the reverse of feed-forward, starting from the output layers to the input layer in a backpropagation step. As the name implies, all weights are interconnected and dependent, so modifying them must consider their effect on the following nodes to minimize the loss efficiently. To minimize the error on a simple perceptron using the gradient descent technique, the derivative of the loss function L for the weight w is calculated to determine the change in the loss function per unit change in y,

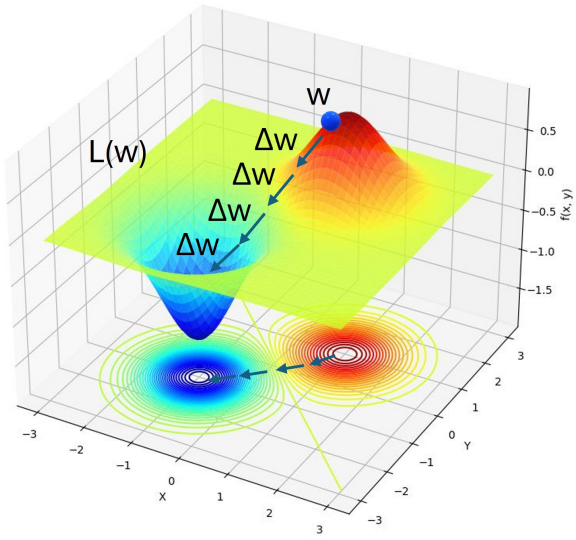

Figure 2.13: 3D representation of the surface of the loss function. The redder colors represent higher losses, while the colder colors represent lower losses. The optimization steps Δw shift the model weights to minimize the loss value.

$$\frac{dL}{dw} = \frac{d}{dw}\left\{\frac{1}{2}(y - \hat{y})^2\right\}. \tag{2.12}$$

Knowing that $y = wx$, the Equation 2.12 can be solved as

$$\frac{dL}{dw} = -(y - \hat{y})\frac{d(wx)}{dw}. \tag{2.13}$$

And Equation 2.13 is equivalent to

$$\frac{dL}{dw} = -(y - \hat{y})\, x = \Delta w, \tag{2.14}$$

where Δw is a shortened form to describe the gradient. Then, the weight is updated, moving it in the negative direction of the gradient, given by

$$w_{\text{new}} = w - \Delta w. \tag{2.15}$$

In this way, it is possible to compute the gradient and update the weights towards a minimum in the simplest case of using only one perceptron. As the number of neurons in-

creases, the complexity of the operations scales due to the interdependence of the weights. For DNNs, backpropagation starts by computing the gradient of the nodes from the last layer to the input layer. To take into account the interdependence of the neurons, it is necessary to apply the chaining rule to obtain the chained partial derivative, given by

$$\frac{\partial L}{\partial w_3} = \frac{\partial f_3}{\partial w_3} \cdot \frac{\partial y}{\partial f_3} \cdot \frac{\partial L}{\partial y}. \tag{2.16}$$

This is necessary to obtain the dependence of the loss function L for a slight change in the weight w_3. Applying the chain rule further, getting all the dependencies of the loss function based on the network weights in a further part of the network is possible. In the example given in Figure 2.12, the dependency of the loss function on the weight w_1 must be calculated by

$$\frac{\partial L}{\partial w_1} = \frac{\partial f_1}{\partial w_1} \cdot \frac{\partial a_1}{\partial f_1} \cdot \frac{\partial f_2}{\partial a_1} \cdot \frac{\partial a_2}{\partial f_2} \cdot \frac{\partial f_3}{\partial a_2} \cdot \frac{\partial y}{\partial f_3} \cdot \frac{\partial L}{\partial y}. \tag{2.17}$$

This scheme is applied to more complex architectures in the same way by extending the chain rule until all the weights are connected and adjusted one by one.[113]

A loop or epoch is considered complete, when the entire training dataset has been run through the model and the weights have been updated. However, usually the whole dataset cannot be run through in one go because it would require a huge amount of RAM to hold all the computed weights in memory. Therefore, the dataset is divided into batches that are passed at once. The batch size also affects the model's performance and how it is trained. The larger the batches, the more context the model gets at once to compute the weights, but it is more computationally expensive.

The number of epochs needed to achieve the desired performance is case-dependent and depends on several factors. However, the number of iterations required to maximize model performance must be balanced against the computational cost of training. This is achieved by minimizing the number of loops the model needs to be trained by controlling the size of the steps taken towards the computed gradient.

A manually selected hyperparameter, the learning rate (LR), controls this. The learning rate determines how large the steps towards the minimum are in each loop. Logically, larger steps would speed up the convergence of the model and reduce the computational cost, but on the other hand, the risk of overshooting the minimum increases if the steps are too large. Smaller steps have the opposite effect, increasing the computational cost at the cost of better precision in finding the minimum of the loss function. In practical cases, a combination of the two is used. Typically, the learning rate is dynamically adjusted

during learning, starting with larger learning rates to speed up finding the minimum loss, and reducing it the moment a plateau appears until a lower LR bound is reached.[113]

A sufficiently representative dataset with corresponding labels must be collected to train a model. An appropriate architecture must be chosen to solve the task in practical cases. A manual selection of hyperparameters must be determined: the number of epochs, the batch size, and the learning rate. The dataset is divided into three parts: training, validation, and test sets. The training set, usually the largest (about 70-80 % of the total available data), is used to tune the weights to minimize the loss function. The validation set, on the other hand, is used to check the model's performance on unseen data. This is crucial to determine that the model does not overlearn features of the training set and can recognize the correct representations of the data to extrapolate to newer examples. The test set is used at the end to ensure that the model is not just tuned to the validation set and can effectively predict on data not used in the training phase.

Once the model has achieved satisfactory performance, the weights can be saved, and the model is ready to be used to predict unseen data by the target users.[113]

2.2.3 Convolutional neural networks for image analysis

For image analysis, using standard fully connected feedforward networks becomes infeasible. This is due to the number of parameters needed in the training step, considering the number of data points a single image has. For example, a normal 1000×1000 pixels image has one million data points. If it is fed into a fully-connected layer with just 1000 nodes, the model would need an extreme amount of one billion parameters to process the data of one image into a model. Since this is not computationally feasible, researchers had to devise a solution to apply deep learning to image data. The solution was the invention of convolutional neural networks (CNNs).[112]

CNNs are neural networks based on convolutional layers inspired by how the human visual cortex works. The special feature of convolutional neural networks is that the first layer is not connected to specific pixels of an image, as with standard DNNs. Still, it only considers pixels in a predetermined "receptive field". The second hidden layer performs the same approach on the output of the first layer, allowing the network to extract simple features in the first hidden layer. Going deeper into the network allows for extracting more complex features from the image.[112]

As mentioned in Section 2.1.1, images are composed of pixels, and each pixel represents a single number (also called intensity) in grayscale images. Convolutional layers are grid filters applied to the image to extract features from it and reduce its size to reduce the level

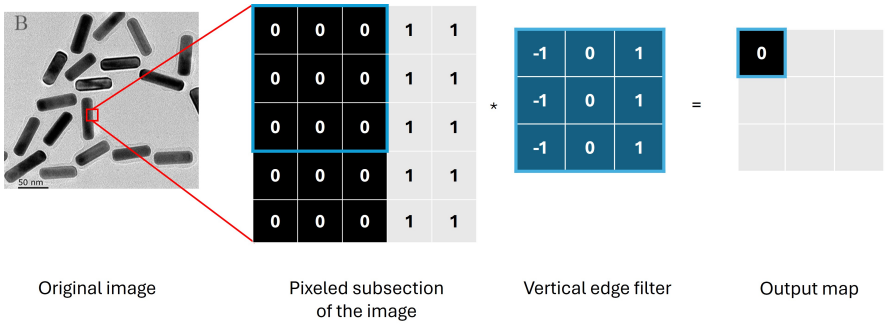

Figure 2.14: Schematic illustration of how a convolutional layer works. A filter to extract the vertical shapes is applied to the image by multiplying the filter values with the corresponding pixels. The results are summed and represented to create a new output map.

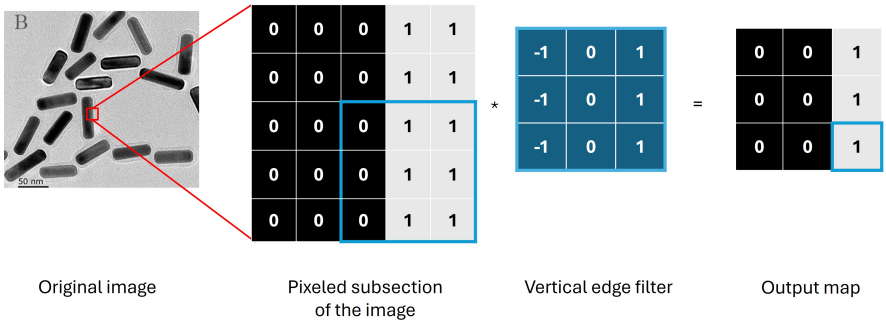

Figure 2.15: Schematic illustration of how a convolutional layer works. After dragging the filter over the whole image, the output map represents the detected patterns with higher values, while uninteresting zones remain deactivated.

of detail from the image and highlight more prominent features. Grid filters are quadratic grids of numbers that can recognize specific patterns in the image when they appear in their receptive field. The filter is applied to the image from left to right by multiplying each grid value by the corresponding overlapping image values within the receptive field. All multiplication results are then added to form a new pixel value of the output map, as shown in Figure 2.14. This example shows a 3×3 grid filter applied to an image to detect a vertical line in a simplified binarized image with 1s and 0s for visualization purposes. After the filter is used, a manually selected number of pixels moves to the right, called a stride, to reapply the filter and continue analyzing the image.[112]

Following the example of Figure 2.14, the filter is applied to the whole image, and the transformed image is the output where the detected pattern is more activated (has a higher value). In comparison, the zones that do not contain the interesting pattern are less activated (have a lower value).

This example was specified to detect vertical lines, but the values of the filters can vary to detect all other types of features in the images. For example, filters in the first few layers extract general features such as vertical and horizontal lines as shown in Figure 2.16.

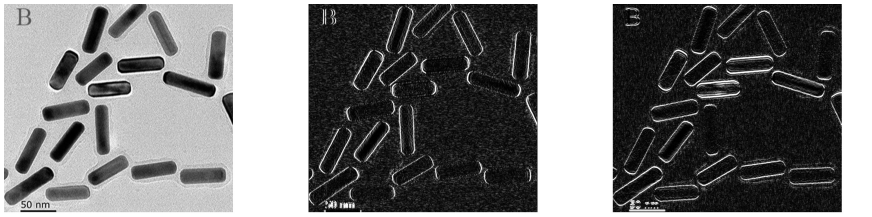

Figure 2.16: Visualization of vertical and horizontal filters applied over the whole image.

In contrast, filters in deeper layers can extract more complex shapes such as circles, particles, or entire microscopic systems. This phenomenon can be well observed in the Figure 2.17.[112]

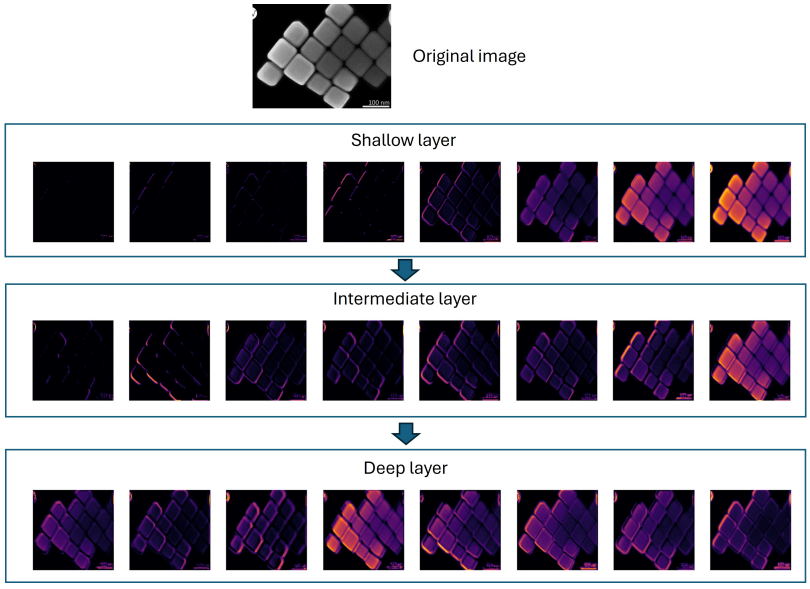

Figure 2.17: Representation of activation maps extracted from a neural network at different depths to visualize the detected patterns.Representation of activation maps extracted from a neural network at different depths to visualize the detected patterns.

An example of a CNN architecture for image analysis is shown in Figure 2.18. The image of interest is input to the CNN. The initial layers extract general features from the images via convolutional operations and apply subsampling to the images to reduce the

complexity of the sampling and to encode the extracted information in lower-dimensional spaces. Finally, the feature maps are converted into a 1D feature vector that contains all the detected information about the patterns. From the 1D feature vector, there are several ways to transform it into an output of interest. If the task is a classification of images, Figure 2.18 shows the connection of the feature vector to a fully connected layer to further analyze the data to an output layer to classify the image into the classes of interest.[112]

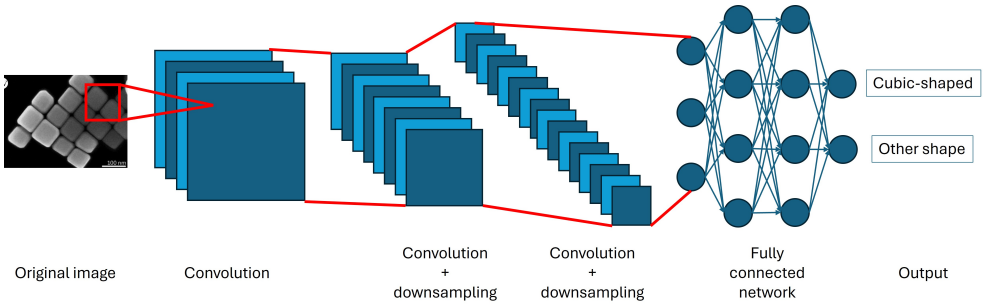

Figure 2.18: Schematic description of the working principle of a CNN for image classification tasks.

This dissertation mainly deals with segmentation tasks. To transform the 1D vector into a segmentation map with the initial resolution, it is necessary to apply inverse convolutions and upsampling operations to obtain the desired pixel-wise segmented output mask, as shown in Figure 2.19.

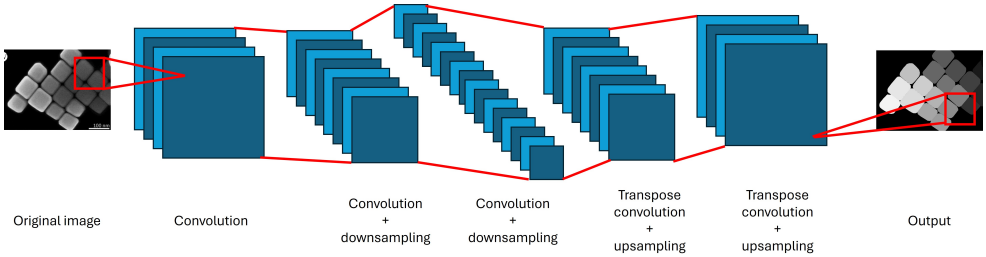

Figure 2.19: Schematic description of the working principle of a CNN for image segmentation tasks.[112]

2.2.4 Visual transformers as the emerging alternative

Recently, a new technology has made a breakthrough in deep learning based image analysis beyond CNNs. This technology is called visual transformers (ViTs), which are a variation of the standard natural language processing (NLP) transformers used in the well-known

large language models such as ChatGPT or DeepSeek.[114,115] ViTs apply the transformer architecture to image data, often achieving competitive performance on small datasets and even better performance than CNNs on large datasets.

The core idea behind ViTs is analogous to NPL approaches, where instead of partitioning text, e.g, into words, called tokens, they partition images into patches and then use self-attention mechanisms to capture relationships between the different patches of the images. This approach has the advantage of learning long-range relationships and global context between the patches, rather than being limited to the locality of receptive fields found in CNNs.[113]

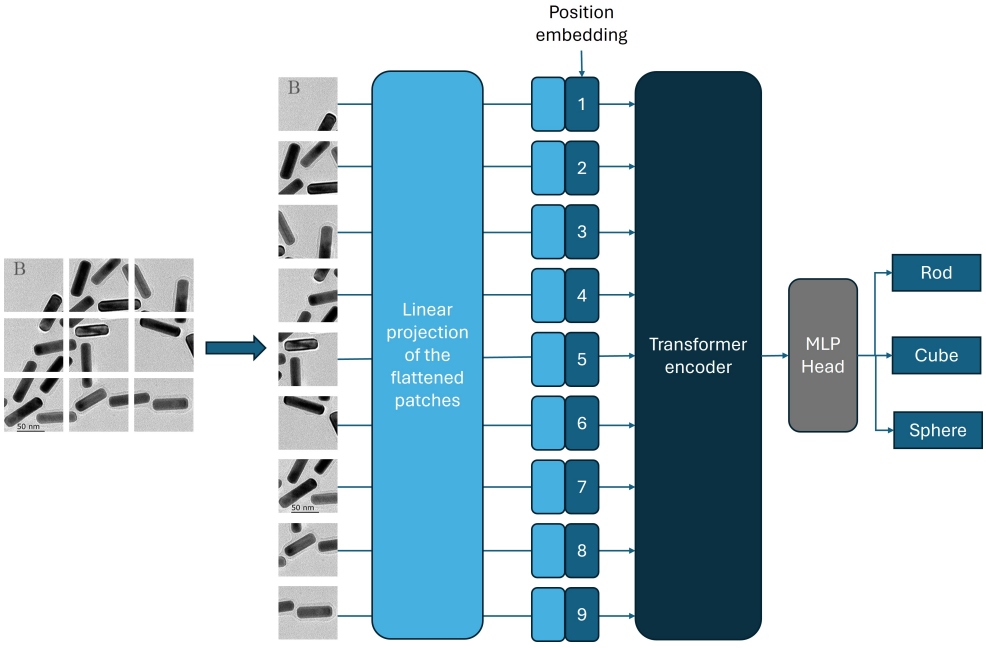

Figure 2.20: Schematic of the block structure of a visual transformer architecture.

To understand the basic principle of this technology, Figure 2.20 shows a schematic pipeline of how ViT works. First, the core idea of ViTs is to transform images into a sequence of fixed-size patches. The raw patches are flattened into 1D vectors, concentrating the raw pixel information into a suitable format for passing into the model.[113]

The flattened patches are passed through a trained fully connected layer to embed the encoded information of the pixels in a high-dimensional space suitable for analysis by the model with a formula similar to equation 2.1,

$$z = W \cdot x + b, \tag{2.18}$$

where z is the patch embedding, W is a learnable weight matrix, x is the flattened input vector, and b is the bias vector.

In patch embedding, the ViTs, unlike CNNs, do not preserve crucial spatial information of the patches that the model can work with to understand the spatial relationships within the images. Therefore, an extra step is required to add a vector to the path embedding to describe the spatial position of the patch in the image. The positional embedding is added to each patch embedding individually, and the exact mechanism for encoding it is also a trained feature learned from the data.

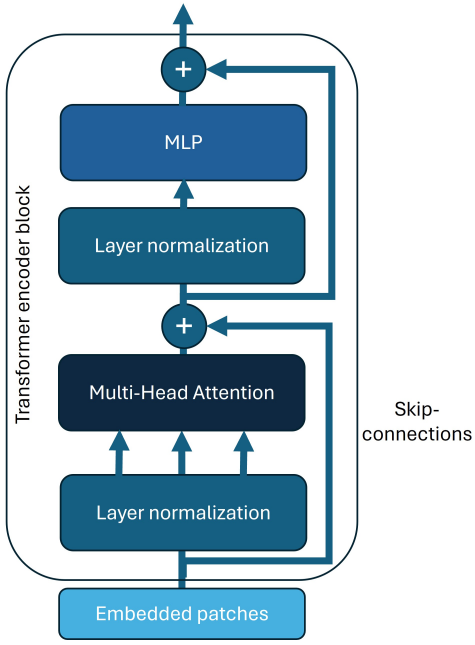

Figure 2.21: Schematic description of the working principle of a transformer block.

Once the patches are embedded along with the positional information, the data is ready to be fed into a series of concatenated transformer units that do the real work of learning and extracting the representations from the data. The innovative technique behind the transformer block is the so-called self-attention mechanism. This mechanism allows each patch to attend to every patch in the image, thereby capturing the global context. The work involves three key vectors for each patch:

- A query vector (Q) represents what the patch looks for in other patches.
- A key vector (K) represents what the patch offers to other patches.
- A value vector (V) contains the information that must be aggregated.

As the patches pass through the model, an attention score is calculated for each patch in combination with all other patches. This computes how much attention a patch should pay to the other patches, and is done by computing the dot product of the query vector with all the key vectors of the other patches. These results are then normalized to produce a set of attention weights. In the end, the output of each patch is a weighted sum of the value vectors of all the patches, allowing the model to dynamically focus on relevant parts of the image, regardless of their locations.[113]

Rather than using a single attention mechanism, Transformers run multiple attention "heads" in parallel. This well-known multi-head attention technique allows the model to simultaneously focus on different parts and aspects of the image. For example, one head can focus on various colors of nanoparticles, while another focuses on finding round shapes. The heads' outputs are then concatenated and combined through another linear layer, creating a rich, aggregated image representation. Following the attention layers, as shown in Figure 2.21, another feed-forward network is used. It consists of two linear layers connected by a nonlinear activation to add nonlinearity and complexity to the extracted image representation.[113]

In addition, two additional tricks are used: residual connections and layer normalization, which improve learning stability on deeper networks. First, residual connections are applied after a multi-head attention block to preserve information and ensure that gradients can flow backward through the model during training. This is done by adding the layer's input to the output, which improves the training of deeper networks and avoids problems such as the vanishing gradient problem. After applying residual connections, transformers apply layer normalization to stabilize and speed training. This technique normalizes the inputs over the features for each sample, rather than in batches as in CNNs, which is particularly useful in the case of transformers, where the input can be of varying size.[113]

This was the explanation for the transformer block. In typical ViTs, however, the encoder consists of several stacked layers, each with its multi-head attention block followed by a feed-forward block. Analogous to CNNs, the stacked layers allow the model to capture increasingly complex and abstract features from the images. The first layer typically focuses on low-level features such as edges and colors, while deeper layers capture higher-level features such as whole particles or particle systems.

2.2.5 Adding an extra dimension: 3D networks

In this work, 3D models are often used to handle a third dimension, such as videos with an extra temporal dimension or tomographies with an extra spatial dimension. This extra dimension in the input data can improve the performance of the models, as it provides additional context for the model to make predictions in the task of interest. Using the nanoparticle example, if nanoparticle tomography is analyzed and the model can consider the spatial context of the layers before and after, it is easier to be sure where a particle is and drastically reduces the misdetection of artifacts or noise.

Surprisingly, the principles of 2D CNNs and ViTs do not need to be modified much to apply the networks to 3D data, since they are already adapted to handle multidimensional data *per se*. For the CNNs, the 2D convolution step uses a filter of size F^2 over an image of height H and width W with the shape $H \cdot W$. Modifying the network to account for an extra depth dimension D in the image $H \cdot W \cdot D$ is as simple as adding an extra dimension to the filter, creating a filter F^3. Then, the 3D convolution slides are in depth and the typical height and width.

For the ViTs case, the modifications are similar. In 2D cases, an image of shape $H \cdot W$ is divided into several patches of size P^2. The patches are then flattened into a vector of length P^2 and then projected into a high-dimensional embedding space, which is the input to the transformer block. For 3D ViTs, the difference is the dimension of the patches, which become P^3, also called cuboids. As a result, the flattened vector becomes larger than in the 2D case, but the projection into higher-dimensional space remains similar. However, the transformers must consider another modification: position embedding, which becomes more complex in the 3D case, including the new dimension encoded in each cuboid.

Integrating 3D architectures may be easier from a theoretical point of view, but it has serious consequences for learning DL models. A three-dimensional filter or cuboid directly impacts the model's complexity, drastically increasing the number of parameters. A highly complex model means a better ability to capture rich interdimensional relationships, but also increases the risk of overfitting the model when the datasets are limited. In addition, more parameters directly impact the computational cost of training and running the 3D models. This results in higher memory requirements, longer computation times, and the need to reduce batch sizes and their contextual consequences.

2.2.6 U-Net and derivatives: modern deep learning architectures for image analysis

The previous chapter introduced the basic concepts to understand how CNN and ViT models work in different dimensions. However, there is another gap between the building blocks of a model and basic architectures and the actual state-of-the-art architectures used in this work to address complex and multidimensional cases in materials science.

This section explains the working principle of the most widely used architecture for image segmentation: the U-Net model.[116] In this context, an architecture is defined as a specific structure of convolutional or transformer blocks, connected in a specific order and often with the introduction of additional blocks or functions that help the model to improve data analysis. The most basic example of an architecture for segmentation has been presented in Figure 2.19, where the blocks are linearly connected, first decreasing the image's resolution until the image information is encoded into a 1D vector. Then, using upsampling and transposed convolutional blocks, the original size of the image is restored as a segmentation map with the segmentation of interest.

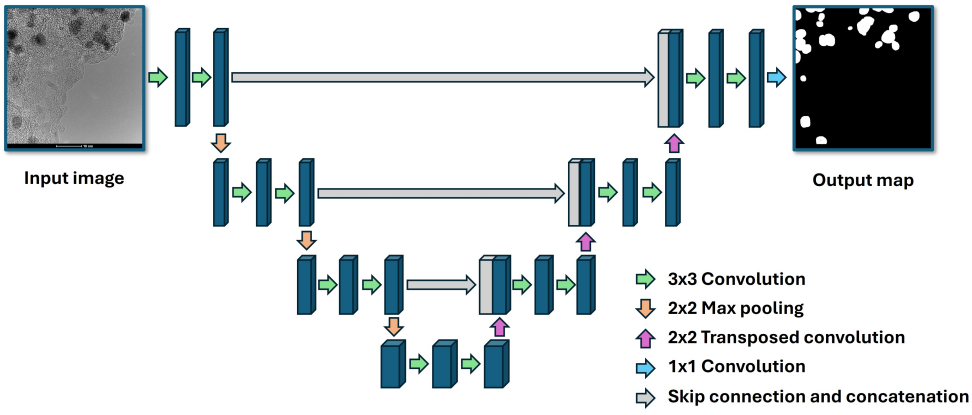

Figure 2.22: Schematic description of the working principle of the U-Net model with the corresponding connections.

Google proposed this first version as AlexNet in 2012.[32] From then on, several advances in architectural design were made to improve the results obtained by the models continually. But it was not until 2015 that Ronneberger *et al.* revolutionised architectural design with the U-Net model. This architecture is shown in Figure 2.22. While the basic structure of the U-Net model is similar to the encoder-decoder structure of the AlexNet presented above, the U-shape resulting from the additional skip connections and concatenations makes it unique and gives the architecture its name. The advantage of the U-Net model is that, in deeper layers, especially after the feature vector has been built and in the

decoder part of the architecture, some of the spatial information is lost due to the decrease in resolution, making it difficult for the model to locate the different objects accurately. This problem is addressed by concatenating the encoder phase output maps, which are still rich in information, with the corresponding decoder output maps. This concatenation has shown excellent results in many applications in all scientific fields.[58,59]

Starting from the basic U-Net architecture, researchers have tried several modifications to improve the learning paradigm of the network and achieve better results in terms of accuracy, speed, complexity, and even in the type of segmentation. A notable example is the StarDist model, which uses a U-Net backbone for image analysis, but with a modification in the loss function that allows the model to predict in an instance segmentation fashion and distinguish overlapping convex polygon-shaped entities.[58] Another possible modification is the dimensionality of the network, which can be transformed from a 2D to a 3D architecture as described by Ronneberger *et al.*.[117] Furthermore, it is also possible to partially or entirely replace the convolutional blocks with transformer blocks, obtaining the so-called Swin U-net.[71] In addition, several groups have aimed to increase the depth of the U-net model and the complexity of the model without compromising the computational cost. To do this, residual architectures, such as ResNets and ResNexts, were used as the backbone for the U-Net model, maintaining the same skip connection and concatenation characteristics, but allowing the increase of trainable parameters in the model.[74]

This work used the above-mentioned architectures, encompassing the classic U-net or derivatives that implement additional functions or blocks to increase the model's performance. The specific details of each architecture are explained in more detail in the corresponding chapter where it is employed.

2.2.7 Evaluation metrics

As mentioned in Section 2.2.1, the last piece needed to close the training loop is the assessment of whether what the model is learning is correct. The labels are used for this, since the performance evaluation is calculated by comparing the correct labels with the model's output. The quantification of the comparison is called the evaluation metric. Depending on the task, several methods exist to quantify the model performance, but in this work, one specific task is focused on: segmentation. Semantic segmentation is the methodology to identify, pixel by pixel, a particular object of interest from an image and classify the pixels into general classes.[112]

For segmentation tasks, accurate methods exist to quantify the performance of a model on a given task. The trained model makes predictions on previously unseen data for eval-

uation, for which true labels have been annotated. To quantify the model's performance, it is possible to compare the true labels and the model's predictions. The overlap will contrast the correctly identified pixels as the class of interest and those that were not. This information allows quantification of each prediction using a method called the confusion matrix. This method counts each pixel classification and calculates four different values:

- Number of True Positives (TP) is the number of predictions correctly classified by the model into their respective classes.
- Number of False Positives (FP) is the number of pixels incorrectly classified into the wrong class.
- Number of True Negatives (TN) is the number of pixels correctly identified as background and do not belong to any class of interest.
- Number of False Negatives (FN) is the number of pixels that belong to a given class, but are incorrectly classified by the model as part of the background.

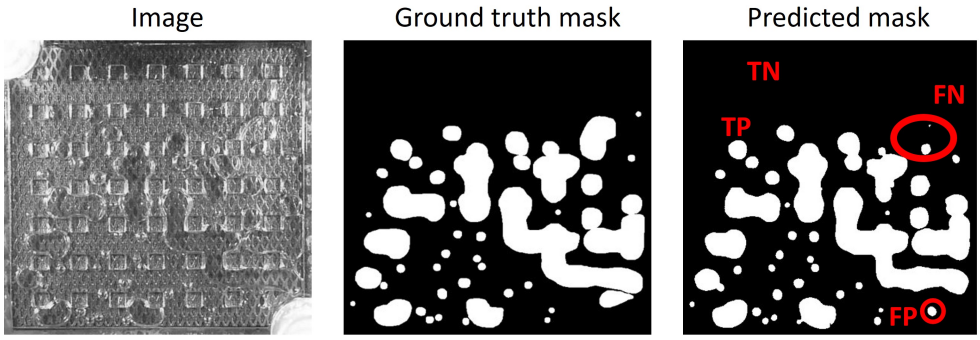

Figure 2.23: Demonstration of an optical image, ground truth, and predicted masks. Examples of true-positive (TP), true-negative (TN), false-positive (FP), and false-negative (FN) predictions are shown on the predicted mask.

Figure 2.23 illustrates these four classification results and provides a visual reference for understanding the distribution of TP, FP, TN, and FN within a segmentation task. More sophisticated metrics, such as precision, recall, and F1 score, can be derived from these basic metrics. Precision quantifies the accuracy of positive predictions, recall measures the ability of the model to identify all relevant instances, and the F1 score provides a harmonic mean of precision and recall, balancing the trade-off between these metrics for a comprehensive assessment of model performance. The three metrics are defined as

$$\text{Precision} = \frac{TP}{TP + FP}, \text{ Recall} = \frac{TP}{TP + FN}, \text{ and } F_1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}. \quad (2.19)$$

During and after training each model, the model's ability to effectively predict on unseen data is evaluated, quantifying the prediction performance with solid numbers. Furthermore, employing the same evaluation metrics in the following chapters enhances the ability to benchmark the performance of various models and even compare the results of different architectures across diverse use cases.

Chapter 3

UTILE-Oxy: Deep Learning to Automate Video Analysis of Bubble Dynamics in Proton Exchange Membrane Electrolyzers

Note: The following chapter is based on the publication "Colliard-Granero, A., Gompou, K. A., Rodenbücher, C., Malek, K., Eikerling, M. H., & Eslamibidgoli, M. J. (2024). Deep learning-enhanced characterization of bubble dynamics in proton exchange membrane water electrolyzers. Physical Chemistry Chemical Physics, 26(20), 14529–14537". The personal pronoun "we" is used throughout this chapter to refer to the group of researchers that were part of this specific study. In this chapter, the main contributions were developed by me with the assistance of Keusra Armel Gompou. Christian Rodenbücher collected the data, and the other co-authors took a supervisory role through this work.

3.1 Introduction

As mentioned in Chapter 1.1.2, polymer electrolyte membrane water electrolyzers (PEMWE) stand out for their high current densities, reduced propensity for gas crossover phenomena, and more compact cell design. However, challenges persist, including the high cost of fabrication materials, reduced membrane resistance at elevated pressures, and susceptibility to cross-osmosis phenomena.[15,118]

In particular, the generation of oxygen gas bubbles at the anode of the catalytic layer is a significant factor inhibiting water access to the reactive sites, particularly under condi-

tions of high current densities.[118,119] Consequently, several studies focus on optimizing the two-phase flow behavior to increase PEMWE performance. For instance, several works have been conducted to explore the effects of flow field channel geometry on cell efficiency, comparing parallel and serpentine flow fields. Majasan and coworkers found that under a range of relevant operating conditions, including variations in water inflow rate, fluid dynamics, and temperature, a parallel channel configuration is more efficient than a serpentine one.[120]

Another critical study area is the relationship between flow regimes and cell current density. Dedigama *et al.* provided evidence that the presence of larger bubbles can displace more water, facilitating the removal of smaller bubbles adhered to the electrode surface, and ultimately improving water transport.[121] Aubras *et al.* demonstrated that the transition from effervescent flow of smaller bubbles to slug flow dominated by larger bubbles enhances overall mass transport, diminishes ohmic resistance, and contributes to improved cell efficiency.[122] Generally, the dynamics of bubble coverage, along with the concurrent processes of bubble growth/detachment and electrochemical reactions, are crucial to cell performance. Su *et al.* proposed that a high bubble coverage can indicate a deficient water supply to the cell, leading to a higher voltage and decreased efficiency.[123] These findings collectively underscore the complex interplay between flow design, bubble dynamics, and cell efficiency.

Understanding the processes of bubble formation, growth, and detachment is pivotal in the context of PEMWE, where their interplay determines the two-phase flow dynamics.[122] A gap exists in characterization techniques, necessitating advanced methodologies to elucidate the complex behavior of bubbles in PEMWE. Conventionally, the evaluation of bubble dynamics has relied on analyzing extensive datasets acquired via indirect techniques such as acoustic emission (AE), electrochemical impedance spectroscopy (EIS), and pressure drop measurement. AE is a non-destructive, *operando* diagnostic tool utilizing a piezoelectric sensor to capture mechanical disturbances emanating from an object.[124] This methodology facilitates the determination of bubble size distribution and the elucidation of diverse gas-flow patterns through the analysis of acoustic waves generated by bubbles.[125,126]

However, AE has limited spatial resolution and susceptibility to background noise, which can impede bubbles' accurate localization and size determination. EIS offers insights into the individual contributions to an electrochemical system's total impedance, encompassing processes such as interfacial charge transfer and mass transport.[127] Nonetheless, interpreting impedance data, especially in systems like PEM electrolyzers with evolving bubbles, is challenging. Especially, distinguishing between the various contributing

factors in the impedance spectrum, e.g., charge transfer resistance, mass transport limitations, and bubble dynamics. Conversely, pressure drop measurement, utilizes pressure sensor probes to detect signals that reflect bubble behavior, allowing for the correlation of these signals to dynamic bubble phenomena, as discussed by Zhang *et al.*[128] Therefore, the integration of indirect techniques with direct observational methods can offer a more holistic understanding of bubble dynamics.

Optical photography facilitates the observation and quantification of bubble behavior.[129] While this technique is advantageous for transparent cells, offering unobstructed visualization, quantifying important parameters can still pose significant challenges. The manual processing of extensive video data sets is laborious and prone to subjective interpretations that can vary with the analyst's bias. The advent of sophisticated image analysis coupled with the evolution of artificial intelligence (AI) offers a compelling solution to these limitations. These technological advances have laid the groundwork for automated, objective data extraction from image sequences, effectively circumventing the issues of manual analysis.[59,130–132]

Recently, Sun *et al.* performed a study where a multi-task deep learning (DL) network was employed for instance segmentation to elucidate fission gas bubbles within nuclear fuel.[133] Anderson *et al.* utilized DL to identify helium bubbles in irradiated micrographs autonomously and to extract their radii and cumulative volumes. The model exhibited a high accuracy, achieving a 93 % success rate in detecting bubbles on high-magnification micrographs, and maintained robust performance in analyzing lower magnification samples.[134] Kim and Park introduced an instance segmentation model designed to autonomously discern and delineate the contours of bubbles across diverse flow conditions, attaining an average precision (AP50) of 98 %.[135] Nevertheless, DL for bubble analysis in PEMWE remains an under-explored field. The closest work in the literature employs a state-of-the-art object detection framework, YOLOv7, to identify anodic oxygen bubbles within a transparent PEMWE system.[118] While this method facilitates the extraction of bubble features, including the area, their count, and the extent of bubble coverage across the electrolyzer, it relies on bounding boxes to approximate bubble detection. This strategy is prone to inaccuracies, especially when confronted with bubbles that have coalesced or present irregular shapes. Hence, there is a widely recognized need for methodologies capable of delineating the actual contours of individual bubbles during their formation, growth, and release times, to enhance the precision of feature extraction and analysis.

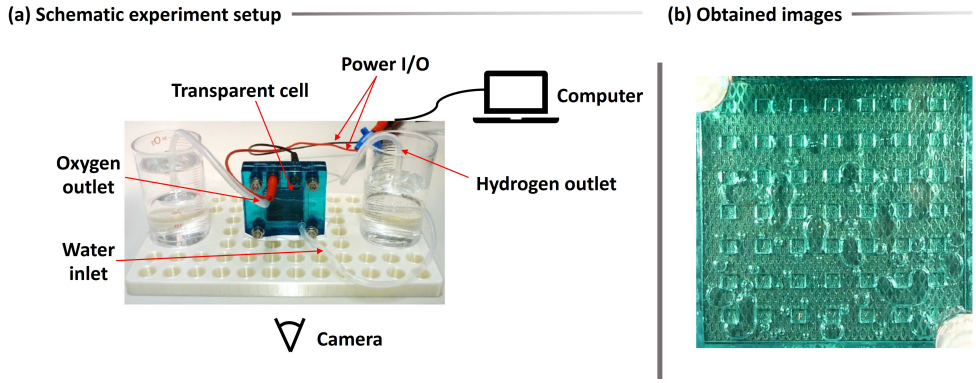

Figure 3.1: (a) Schematic experimental setup for recording the dataset. (b) Exemplification of the obtained data from the experiment from the PEMWE transparent cell and the oxygen bubbles.

3.2 Transparent PEMWE cell data acquisition and deep learning-based bubble analysis workflow

This chapter presents a detailed, automated deep learning-based workflow for analyzing video footage of oxygen bubbles in proton exchange membrane water electrolyzers (PEMWE). The proposed workflow is twofold: firstly, it involves the creation of a meticulously annotated dataset, which is then used to train models for the semantic segmentation of bubbles; secondly, it facilitates the automated extraction of bubble size and shape characteristics, which are crucial for the analysis of their distribution properties. The practicality of this approach is exemplified by successfully training models using a relatively small dataset of only 35 images, demonstrating that even limited data can yield insights for specific use cases. The dimensions of the dataset were determined based on the prior research, in which we empirically established the optimal quantity of images required to achieve saturation in the performance of the U-Net model. This investigation facilitated a balance between model efficacy and the time expended on annotations.[59] Furthermore, this study offers a flexible framework, allowing users to replace the provided model with a bespoke one tailored to their datasets, leveraging the robust computer-vision capabilities developed in this research for the extraction of time-resolved size and shape analysis of the regions of interest. This adaptability makes the workflow a powerful tool for researchers seeking to enhance their study of electrolyzer performance through advanced image processing techniques.

To demonstrate the application of deep learning for the localization of gas bubbles in the flow field of a PEM electrolyzer, we employed a commercial reversible fuel cell (FCSU-023,

Horizon Fuel Cell Europe, Czech Republic) operated in the electrolysis mode. The housing and the flow field of the cell consisted of green-blueish transparent plastic, allowing for *operando* optical inspection of bubble evolution (Figure 3.1 (a)). The cell was driven by a power supply in constant voltage mode (2651A, Keithley, USA), which also served to measure the current. Before starting each electrolysis run, the anode compartment of the cell was filled with deionized water, which was then consumed during the process as no water circulation was applied. Images of the anode, where the oxygen bubbles were released, were recorded by a digital camera (Lumix TZ 18, Panasonic, Japan) with a frame rate of 30 fps and a resolution of approximately 25 px/mm. The images were cropped to the active electrode area of approximately 25 × 25 mm^2 as shown in Figure 3.1 (b), with the water inlet visible in the bottom right and the gas outlet in the top left part. The cyclic voltammogram of the cell obtained at a speed of 10 mV/s is shown in Figure 3.2 (a). The onset of the electrolysis process can be seen as a steep current increase at 1.5 V. Hence, the images were recorded at four different voltages of 1.5, 1.6, 1.7, and 1.8 V while the voltage was applied for approximately 400 s each.

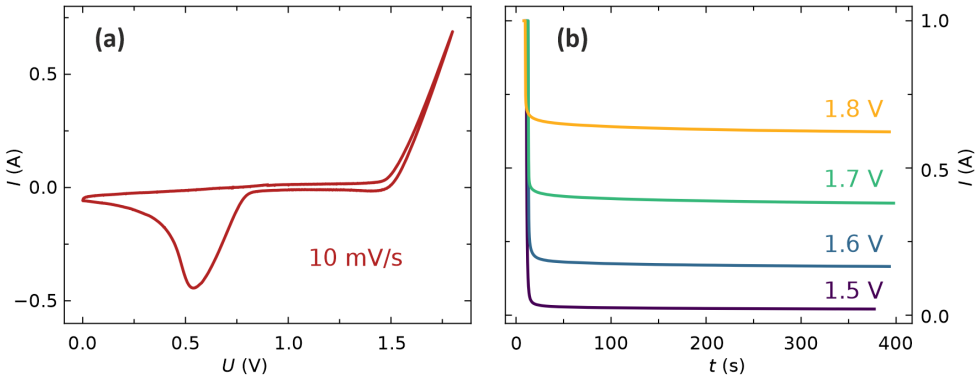

Figure 3.2: (a) Cyclic voltammogram of the electrolyzer. (b) Current as a function of time during the application of different constant voltages.

Our methodology for analyzing bubble dynamics in videos encompasses four principal steps, as illustrated in Figure 3.3: 1. manual annotation of oxygen bubbles to prepare the training dataset; 2. employing supervised learning for the semantic segmentation of bubbles; 3. extracting features automatically from the regions of interest (ROI); and 4. statistical analysis and visualization of the findings. Initially, we meticulously annotate the oxygen bubbles, as the precision of segmentation relies heavily on the quality of this labeled data. To facilitate this process, we used Label Studio,[136] a comprehensive and intuitive annotation tool. This software facilitates efficient annotation and the automated export of masks in the desired format.

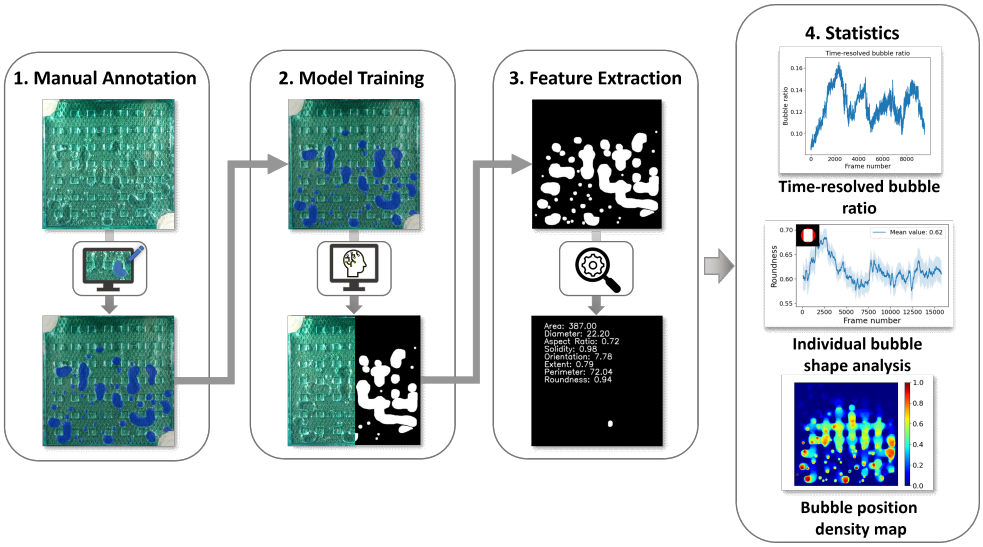

Figure 3.3: Methodological pipeline for the video analysis of oxygen bubbles in PEM electrolyzers.

For the training and validation of the model, we annotated thirty-five optical frames. Initially, each frame measured 640 x 630 pixels, but we resized them to 512 x 512 pixels for the training process. This resizing was crucial to optimizing the model's performance, as it specializes in analyzing squared cells. Furthermore, maintaining a minimum resolution of 512 x 512 pixels is essential to prevent any potential resolution-related issues. These frames were carefully chosen from various parts of the experiment to represent a broad range of bubble scenarios. This deliberate choice is aimed at improving the model's generalization abilities. Using visual cues, an expert performed the detailed annotation task, with a special focus on bubbles that were difficult to differentiate. To guarantee the reliability of the annotations, two additional reviewers examined them to achieve a unified consensus on the bubble selection.

In this study, utilizing the provided images and their associated annotated masks as ground truth, we focused on the pixel-wise identification of oxygen bubbles by implementing several U-Net-based architectures. These architectures are based on fully convolutional neural networks, which conduct a series of convolutions and down-sampling operations to identify the key patterns present within the images in a latent space. The encoding process progresses iteratively until the entirety of the information is consolidated into a singular vector. Subsequently, a reverse decoding process is initiated: up-sampling, coupled with transpose convolutions, restores the data to its original size. A salient feature of the U-Net architecture is concatenating the up-sampled decoder feature map with the

encoder feature maps of a matching resolution. This technique facilitates the discernment of object boundaries and edges, harnessing low-level and high-level features, culminating in a refined segmentation output.[116]

Within the scope of this research, we experimented with the canonical U-Net framework along with two advanced variations. The first is a U-Net model augmented with a ResNeXt101 network as its backbone, integrating residual connections. The second is the more recent Attention U-Net model, which capitalizes on attention mechanisms to enhance segmentation precision.

The ResNeXt101 U-Net architecture fuses the ResNeXt101 model's capabilities with the U-Net structure. ResNeXt101, a 101-layer variant of the ResNeXt series, utilizes the split-transform-merge strategy. Its unique feature is "cardinality" - parallel paths in a block. Instead of increasing depth or width, elevating cardinality enhances performance. By using ResNeXt101 as U-Net's backbone, the network captures intricate image features more effectively.[74]

Attention U-Net augments the traditional U-Net with an attention mechanism, enabling the model to concentrate on specific image regions. This is vital for detecting various object scales or minor yet crucial image sections. Attention gates, applied before each decoder concatenation step, weigh encoder features, intensifying key pixels, and suppressing irrelevant ones. This heightens segmentation precision, especially when target regions are surrounded by noise.[137]

In our study, we evaluated the performance of our trained model using standard evaluation metrics, notably the Intersection over Union (IoU) threshold. IoU is a commonly used metric in image segmentation, providing a quantitative assessment of the overlap between the model's predictions and the actual annotated masks. It calculates the proportion of overlapping pixels to the total number of pixels present in the target class's predicted and true masks.

We performed further statistical analysis using the segmentation maps generated from model predictions. We calculated the pixel count representing bubbles in each frame using elementary mathematical procedures, estimating frame-wise bubble coverage. This analysis facilitated the creation of plots illustrating time-resolved visual coverage and bubble area distributions.

Furthermore, by overlaying these masks, we derived density maps that provide insights into the spatial distribution of bubbles throughout the experiment. Focusing on the distinct contours of bubbles in the masks, we formulated computer vision algorithms using the OpenCV library.[138] These algorithms were adept at identifying individual bubbles and

extracting many shape characteristics, enhancing our understanding of bubble dynamics and morphology.

3.3 Benchmarking of cutting-edge DL architectures and visual results of the semantic segmented bubbles

In the study of PEMWE videos, the accurate segmentation of bubbles is of great importance for detailed analysis. Our methodology for this segmentation task begins by choosing an appropriate neural network architecture tailored for optimal feature extraction and segmentation capability. To enhance the model's performance, we initialize the model's weights using a pre-trained model, capitalizing on previously learned features. The learning rate, which plays a crucial role in the convergence of the training process, is then optimally selected to determine the magnitude of steps taken toward the minimum of the loss function during training. The primary objective during training is to minimize a loss function that quantifies the discrepancy between manually annotated masks and the model's predictions. A reduced loss indicates superior model performance. After training, we critically assess the model's effectiveness using a validation set to determine its robustness and generalization capability.

In our research, we undertook a benchmarking study, comparing three distinct variations of the U-Net model. Firstly, we evaluated the conventional U-Net 2D model, which was trained using the ZeroCostDL4Mic approach.[139] This was compared with a custom U-Net 2D model that integrated a pre-trained ResNeXt101 from the ImageNet dataset as its backbone. The third model was the more contemporary Attention U-Net. The models were trained on the same dataset consisting of 28 images for training and 7 for validation for 100 epochs with a batch size of 8. This training involved segmenting the images into patches and employing a dynamic learning rate, which was adjusted in real time based on model performance across epochs.

As illustrated in 3.1, we present a comprehensive array of metrics for these models. Regarding F1-scores, the models showcased similar performance levels; however, the U-Net model augmented with the ResNeXt101 backbone slightly outperformed the others. It was observed that the standard U-Net architecture achieved lower precision metrics but with a compensatory spike in recall compared to the other two architectures. In the context of our bubble segmentation task, the model's precision is especially relevant. A diminished precision hints at the introduction of non-existent bubble pixels, which reduce the interpretability and accuracy when analyzing features extracted from videos.

Table 3.1: Comparison of various U-Net models trained on an identical dataset, evaluated using the metrics: precision, recall, and F1-score.

Model	Precision [%]	Recall [%]	F1-score [%]
U-Net 2D	81	89	85
U-Net with ResNeXt101 backbone	95	78	86
Attention U-Net	95	75	84

In addition to the models' performances, the decisive step remains the visual evaluation of the segmentation outcomes. To this end, Figure 3.4 offers a comparative visualization with sub-figure (a) depicting a frame randomly sourced from the validation set, and sub-figure (b) showcasing the segmentation results as generated by the different models under study. A close inspection of these frames bolsters our conviction in the superior performance of the U-Net model incorporating the custom ResNeXt101 backbone. Notably, this model yields the most seamless contours around the segmented bubbles and achieves good results in separating bubbles near, a feat not as pronounced in the other models in our dataset.

We have rendered videos of the segmentation outputs to provide a more granular view of our results. These videos, available on our GitHub repository, allow for a detailed examination and facilitate a deeper appreciation of the differences in each model's performance.

3.4 Autonomous video analysis and computer vision-based time-resolved bubble quantification

Implementing automatic video segmentation paves the way for an enhanced software utility capable of extracting intrinsic data from the bubble dynamics unfolding within the cell. These parameters are instrumental for parametrization in simulation sciences and hold the potential to expedite investigations within experimental cohorts. Further, model predictions reveal multifaceted insights about the bubble dynamics of the utilized PEM electrolyzer.

3.4.1 Time-resolved bubble ratio evolution and total bubble area coverage

Figure 3.5 delineates the time-resolved bubble ratio and presents its corresponding histogram for four distinct voltages, retaining an identical experimental configuration. A comparative analysis reveals differences across the different experiments. As anticipated, a reduced voltage corresponds to decelerated gas generation, marked by merely four significant gas releases and a diminished mean bubble coverage. Conversely, an increased voltage

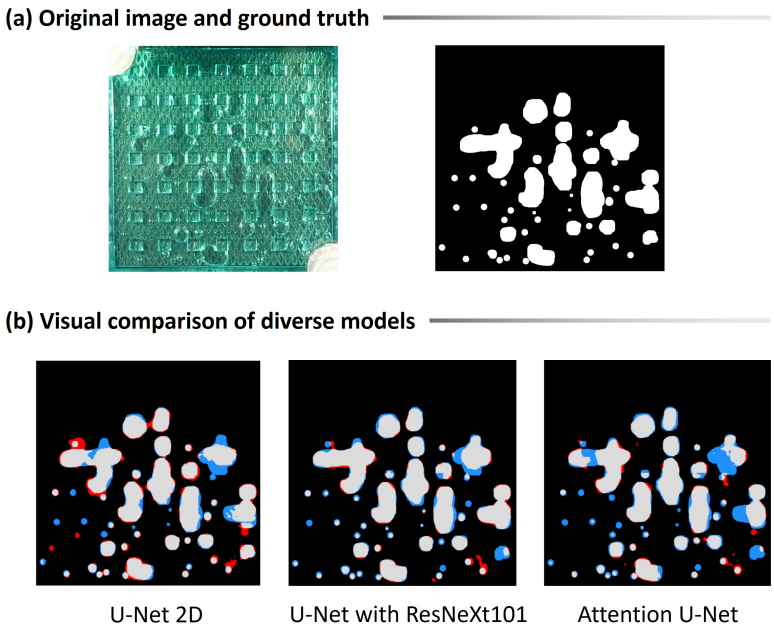

Figure 3.4: (a) Original image and the corresponding ground truth. (b) Prediction output for the three benchmarked models. Gray pixels indicate true positives for correctly detected bubbles, blue pixels represent false negatives where bubble pixels were missed, and red pixels signify false positives for erroneously detected non-bubble pixels. The U-Net model with the ResNeXt101 backbone achieves the smoothest predictions and the highest model performance compared to the other tested models. A clear video comparison is available under the project repository at https://github.com/andyco98/UTILE-Oxy.

shows an augmented visual bubble coverage of the electrode, and the time-sequence plots show a higher periodicity in gas releases.

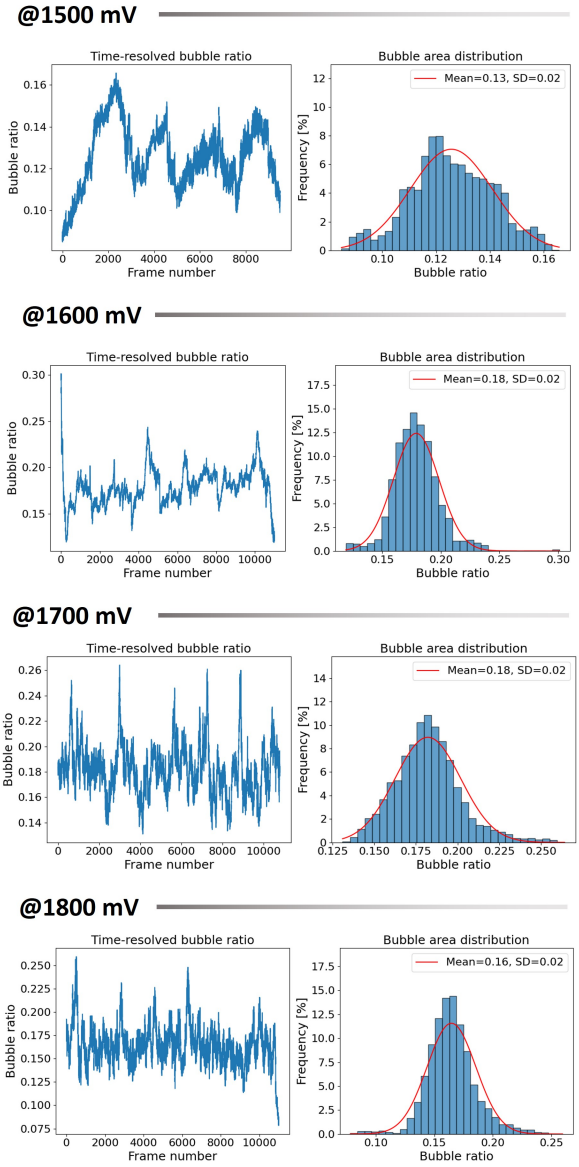

Figure 3.5: Time-resolved bubble ratio evolution alongside the total bubble area distribution for four distinct voltages, all conducted under the same experimental setup. The total bubble area distribution is further analyzed with a Gaussian fit, providing insights into the mean visual bubble coverage and its standard deviation.

3.4.2 Spatiotemporal bubble density heat maps

An interesting result obtained through our methodology is the spatiotemporal distribution of bubbles, visualized as a density map, showing bubble position probabilities over a time series. These maps are constructed using model predictions, where each frame's pixels are classified into two categories: '0' representing the background and '1' for the bubbles. Upon aggregating all these predictions, the regions most consistently occupied by bubbles across the time series attain a higher value, which is depicted in the map with warmer hues. Conversely, areas with infrequent bubble presence are represented with cooler colors. Intermediary zones, represented in green, elucidate regions within the cell where bubbles coalesce and accumulate before release. Such maps correlate with bubble nucleation points, shedding light on predominant bubble generation zones and their distribution across the cell.

Figure 3.6 exemplifies the utility of these density maps. Four distinct maps correspond to a single experiment across varied voltage levels. The map reveals fewer warm-colored regions at lower voltage settings, indicative of bubble nucleation, and a relatively random distribution pattern across the cell. As the voltage is incremented, there's a palpable increase in these nucleation regions, complemented by a more precise depiction of bubble accumulation zones — the latter exhibiting a subtle proclivity towards the cell's right side. Another salient observation gained from these heat maps is the remarkable consistency in nucleation points across experiments, especially at higher voltage levels.

3.4.3 Time-resolved morphological evolution of individual bubbles

Our methodology facilitates an individual analysis of bubbles, focusing on their unique morphological attributes. Analyzing the predicted masks, bubbles fully encapsulated by background pixels are classified as distinct entities. Employing specialized computer vision techniques, we derive attributes elucidating the structural properties of individual bubbles detected throughout the experimental phase. These commonly calculated properties in the realm of materials science include:[140–142]

- Area: Denoted by the pixel count of an individual bubble.
- Diameter: Inferred under the assumption that the bubble's area aligns with a perfect circle.
- Aspect Ratio: Derived by inscribing a bounding box around the bubble and calculating the ratio of its longest to shortest side.

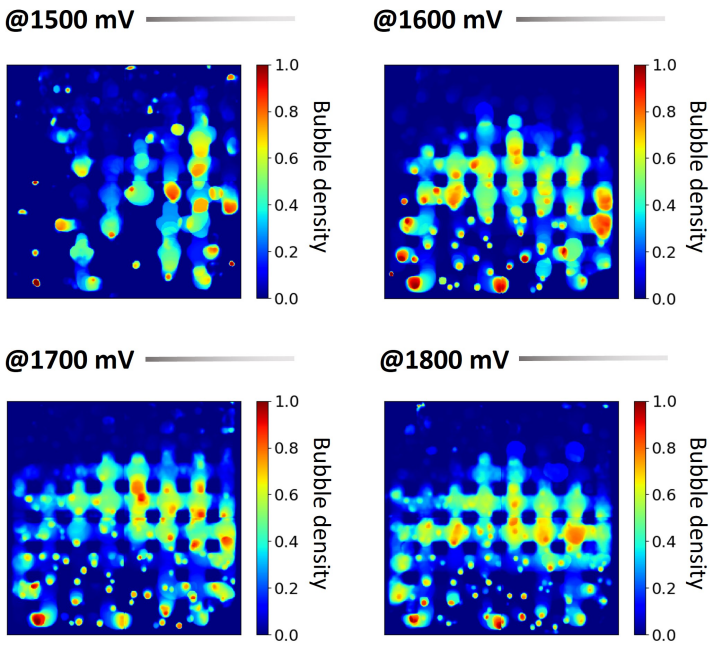

Figure 3.6: Density maps depicting the bubble spatial distribution probability throughout the experiment at four distinct voltages, all conducted under the same experimental setup. Areas with warmer colors indicate positions frequently occupied by bubbles, while colder colors suggest less frequented areas.

- Solidity: A metric indicative of the bubble's convexity. It is defined for a given shape as the ratio of the area of the shape to the area of its convex hull, represented as

$$\text{Solidity} = \frac{A_{\text{shape}}}{A_{\text{convex hull}}}. \tag{3.1}$$

  The convex hull of a shape is the smallest convex set that contains the shape.

- Orientation: Determined by superimposing an ellipse on the bubble and calculating the tilt angle of this fitted ellipse.
- Perimeter: Represents the contour length of a bubble.
- Extent: Defines the bubble's squareness.
- Roundness: Assesses the bubble's resemblance to an ideal circle.

When used together, this set of parameters provides a detailed geometric characterization of bubbles, aiding in a more thorough understanding of their behavior and interaction with surrounding environments.

Given that our data acquisition is frame-centric, it permits temporal analysis. Figure 3.7 illustrates time-series data for a specific experiment conducted at four distinct voltage levels. Initially, bubbles are individually analyzed and subsequently averaged on a per-frame basis. Given the volatility in individual data points, which can make the visualization of emergent patterns complex, we opted for data pre-treatment before visualization. We employed a moving average approach, wherein an average spanning 100 frames was taken, informing the central line of our plot. The accompanying standard deviation is depicted as a shaded region around this line.

This single bubble data provides a comprehensive perspective on the morphological trends of bubbles across temporal intervals. A particular revelation from this data is related to bubble uniformity. For example, smaller standard deviations in the area signify periods of systemic uniformity, indicating a preponderance of similar bubbles. Conversely, pronounced standard deviations, e.g., with respect to area, suggest a bifurcation between small and large bubbles, suggestive of system heterogeneity.

The results shown in this work establish a basis for a detailed, time-resolved characterization of bubble system dynamics within the cell, facilitated through automation.

Additionally, combining the characteristics mentioned earlier can result in a solid manner to characterize particular systems of interest under examination. From the time-series bubble coverage data, we can discern vital information about cell cycling, the periodicity

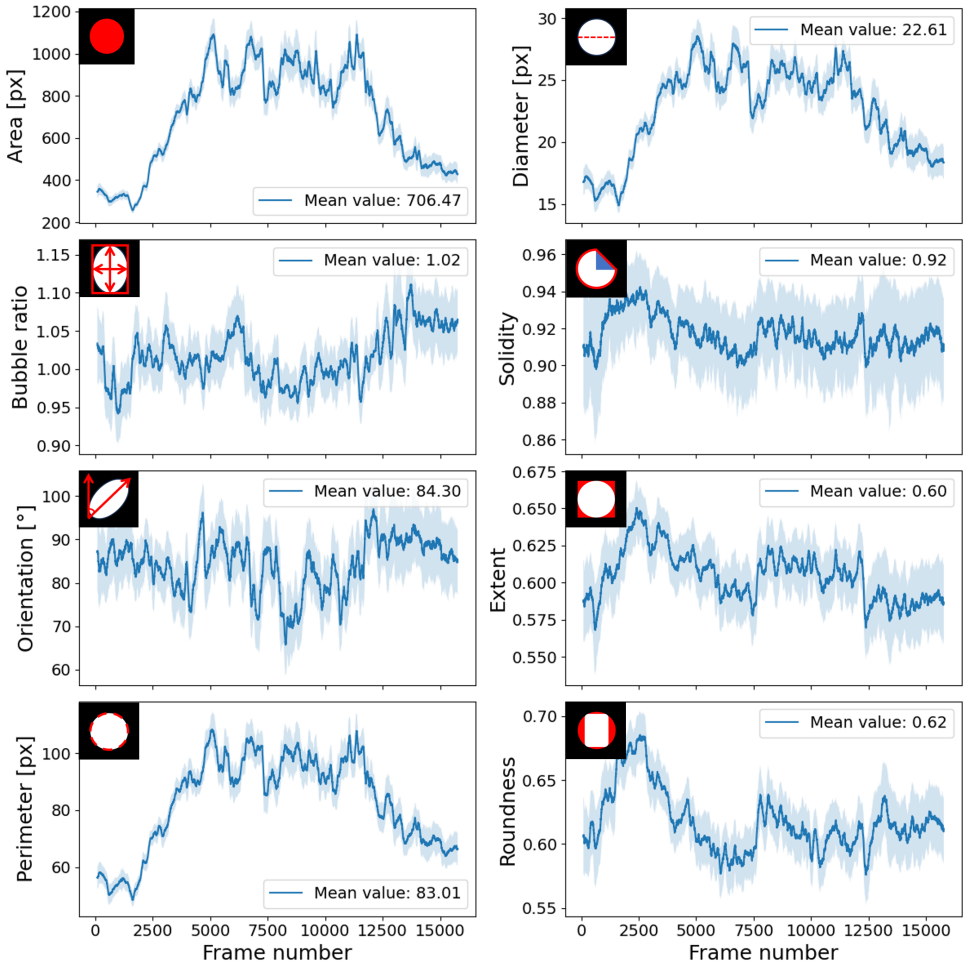

Figure 3.7: Time-resolved analysis of bubble property evolution for one experiment at 1500 mV. Results are visualized using a moving average calculated every hundred frames, with the standard deviation shown as a shaded region around the line.

of prominent gas releases, average bubble coverage, and its inherent volatility. The density maps are instrumental in visually pinpointing nucleation sites and frequent bubble aggregation zones, thereby shedding light on the distribution of active sites and the efficacy of the gas flow field. The contours and characteristics of individual bubbles further provide insights into the bubble formation dynamics and how the environment influences specific bubble shapes. For instance, comparing the average bubble area and its associated standard deviation with the average coverage metrics from the histograms and the active zones from the density maps enables elucidation of the underlying bubble dynamics.

3.5 Conclusions

This research marks a crucial advancement for the autonomous analysis of bubble dynamics in Proton Exchange Membrane water electrolyzers, utilizing optical videos as a rich data source. The autonomous software developed yields substantial results, including detailed visualizations of time-resolved bubble ratio evolution, bubble area distribution, bubble position probability density maps, and detailed analyses of individual bubble shapes. These results provide vital information on the cell's active area at different voltage levels, facilitating a deeper understanding of bubble dynamic developments across varied experimental settings and voltage ranges.

We envisage that our software will significantly enhance the depth and complexity of bubble dynamics analysis within a PEMWE cell. This is achieved by enabling the interpretation of data-rich yet efficiently conducted experimental recordings. Creating an accessible path to obtain such detailed data is essential for augmenting the sophistication of simulation models aimed at understanding these phenomena and resulting in insightful diagnostic techniques. This tool offers experimentalists the dual advantage of rapid yet robust characterization, facilitating iterative testing of diverse cell geometries and setups. Furthermore, the workflow can analyze several thousand frames in minutes, enabling an in-depth analysis that would be unfeasible to obtain manually in a reasonable time or cost.

Looking forward, this work will facilitate further research aimed at deciphering the complexities of bubble dynamics in relation to varying voltages and gas flow fields. Combining this approach with physical modelling makes it possible to unfold the full potential in terms of mechanistic interpretation and prediction of bubble dynamics.

Chapter 4

UTILE-Redox: Deep Learning-based Tool for Autonomous 3D Bubble Analysis of Vanadium Flow Batteries from Synchrotron X-ray Imaging

Note: This chapter presents results of a collaboration with the group of Roswitha Zeis on the accepted manuscript: André Colliard-Granero, Kangjun Duan, Roswitha Zeis, Kourosh Malek, Michael Eikerling, and Mohammad J Eslamibidgoli. Advancing vanadium redox flow battery analysis: A deep learning approach for high-throughput 3d visualization and bubble quantification. *Digital Discoveries*, accepted, 2025. The personal pronoun "we" is used throughout this chapter to refer to the group of researchers who have been part of this specific study. In this study, Kangjun Duan and Kerstin Köble collected and partially annotated the datasets. The main contributions were performed by myself, including further data annotation, model training, and software development. Roswitha Zeis, Michael Eikerling, Kourosh Malek, and Mohammad Eslamibidgoli had a supervisory role during the research.

4.1 Introduction

As stated in Chapter 1.1.3, the transition to renewable energy sources is imperative to combat global environmental and climate issues, necessitating a new power grid system reliant on renewable energy.[144–146] Energy storage is critical in regulating peak demand and dynamics across the grid.[147] However, the intermittency of renewable energy sources poses a challenge when they dominate power generation.[148–151] Redox flow batteries (RFBs),

notably vanadium redox flow batteries (VRFBs),[152,153] offer a promising solution by decoupling energy capacity and power generation, enabling long-duration energy storage and seasonal energy shifting.[144,147,154] Despite its potential, VRFB technology faces hurdles to commercialization, including high cost, maintenance, and side reactions during overcharging.[153,155]

Of particular concern is the hydrogen evolution side reaction (HER) at the negative half-cell, which can degrade electrode performance by consuming charge, reducing energy efficiency, causing electrolyte imbalances, and hindering desired redox reactions.[145,153,156,157] Additionally, this side reaction leads to the formation of bubbles, especially in the negative half-cell,[29,145,153,158,159] which further compromises battery efficiency by disrupting active sites, reducing surface area for redox reactions, and impeding reactant transport.[160,161] Addressing these challenges demands focus on identifying, quantifying, and systematically preventing HER and bubble formation.[162]

Over the last decade, a focused but limited number of studies have tackled HER challenges in VRFBs.[29,158,159,161–167] Innovative diagnostic approaches, including differential electrochemical mass spectrometry, various electrochemical measurements, specialized hydrogen sensors, and buoyancy-based collection devices, have been utilized to monitor and assess the implications of HER. These methods primarily aim to quantify the hydrogen gas produced via HER by employing indirect measurement techniques. Investigations have revealed that the HER activity is influenced not just by the operational conditions,[165,167] but also by inherent properties of battery components themselves.[159,161,162] Despite these advances, a comprehensive understanding of how hydrogen gas is distributed within the electrode during operation remains elusive. This aspect is crucial, as it could shed light on potential adverse effects on electrode integrity and ion transport, thereby guiding efforts to mitigate HER's impact on VRFB efficiency and durability.

Various methods have been employed to study the behavior of bubble evolution. These include optical imaging, such as high-speed camera technology, which captures the dynamics of bubbles.[61,168] Indirect characterization methods, including acoustic measurements and pressure sensors, provide insights into bubble size distribution and gas release patterns.[169–171] Additionally, radioactive imaging techniques, like X-ray imaging and neutron imaging, are used to investigate the internal structure of batteries and bubbles.[163,172] Although recent advancements in neutron radiography and transparent VRFB imaging have enabled investigators to succeed in visualizing hydrogen evolution, they primarily offer a two-dimensional view. This limits their ability to keep the exact distribution of bubbles within the electrode itself.[163,167]

Researchers have applied X-ray imaging for two primary purposes within VRFBs. First, they examined the process of electrolyte injection and its distribution across VRFB electrodes. Additionally, they investigated the presence of gas residues, various electrolyte properties, and the composition of electrode components. This approach has enabled them to uncover detailed information about the internal structures and defects of VRFBs, which is crucial for improving diagnosis methods, advancing research, and enhancing quality control measures.[164,172–176] Bevilacqua *et al.*[173] and Köble *et al.*[172] found it possible to resolve and image the hydrogen bubbles from the HER by synchrotron X-ray tomography. They concluded that hydrogen bubbles often accumulate in existing air bubbles. However, although three-dimensional images can provide more comprehensive information, they face difficulties in image annotation and reconstruction. This hinders the full utilization of experimental data and the in-depth exploration of bubble dynamics.

The process of analyzing 3D volumes from VRFBs is currently hindered by the' slow, inefficient, and often subjective nature of traditional image processing methods. In contrast to manual analysis, which can take weeks to annotate a single tomograph fully, simpler segmentation methods fall into two categories: algorithmic approaches (e.g., Otsu thresholding) and basic machine learning techniques (e.g., random forest pixel classifiers). Thresholding is limited to binary separations of the grayscale histogram and fails to support the multi-class differentiation needed in our analysis. While pixel classifiers can distinguish multiple classes, they rely only on intensity and ignore spatial and shape features, which is problematic given the intensity similarities among bubbles, membranes, and gaskets.

In contrast, deep learning integrates intensity, spatial, and shape information, enabling high-quality segmentation of an entire tomograph within minutes with minimal computational requirements. Therefore, deep learning (DL) and computer vision (CV) technologies emerge as powerful alternatives for automating tasks such as annotation, segmentation, and analysis of imaging data, particularly in energy materials. These techniques have received growing attention in recent years, enabling a wide array of applications, including automated analysis of particle morphology in electron microscopy micrographs[59,131] to elucidate time-resolved bubble dynamics through optical videos of transparent polymer electrolyte water electrolysers.[61] Tang *et al.*[177] developed a 3D deep learning algorithm capable of multi-class segmenting micro-CT tomographies of membrane electrode assemblies in polymer electrolyte membrane fuel cells. This algorithm facilitated the precise and swift segmentation and differentiation of components within the catalyst, membrane, and gas diffusion layers. Subsequently, it enabled the accurate extraction of the connectivity, porosity, and permeability descriptors of the gas diffusion layer and surface roughness between the catalyst layer and membrane. Li *et al.*[178] utilized synchrotron X-ray to-

mography to investigate intermetallic phases in recycled aluminum alloys, employing DL for automated phase segmentation and subsequent morphological and structural analysis such as phase diameter distribution and spatial Fe- and Al-phases distribution. In another study, Kopp *et al.*[179] demonstrated the application of DL for segmenting quantitative multi-class microdamage in heterogeneous materials using synchrotron radiation. These examples underscore the specific application of such models, which are limited by the availability and quality of training data.

The situation reinforces the necessity for further adapting DL algorithms in specialized areas such as 3D bubble analysis in VRFBs. The goal is to augment the accuracy and speed of segmentation, feature extraction, and volume analysis, thereby advancing our understanding and capabilities within these complex scientific domains. This chapter presents a DL-based workflow to analyze synchrotron X-ray tomography images of hydrogen bubbles and the internal structures of vanadium redox flow batteries (VRFB). The workflow comprises four phases: first, it starts with generating meticulously annotated datasets encompassing various examples. The segmented datasets are utilized to train deep learning models, facilitating the automation of semantic segmentation for various regions of interest (ROIs) in images. Upon successful training and demonstration of accurate predictions on unseen imaging data, we establish a framework for the automated analysis of the segmented bubbles' size, shape, and distribution. Lastly, the outcomes of this analysis are represented through graphs and 3D visualizations of the analyzed volumes and other characteristics.

The availability of this methodical pipeline allows researchers to streamline the time-consuming process of analyzing complex 3D volumes in VRFBs, cutting it to just a few seconds. Additionally, this framework will enable users to swap out the provided models with custom versions specifically designed for their unique datasets, thereby broadening its utility. This adaptability, paired with the sophisticated computer vision techniques introduced here for 3D tomography analysis and morphological and spatial analysis of bubbles, highlights the substantial impact of this research on the field.

4.2 Experimental synchrotron X-ray tomography data acquisition of VRFBs and deep learning-based tomography analysis workflow

The experiments were conducted by Roswitha Zeis's group,[164] where the design of the cell up came from previous work as shown in Figure 4.1 (a).[180,181] The carbon felt materials utilized in this study were one commercial material, SIGRACELL® GFA 6 EA. It is a

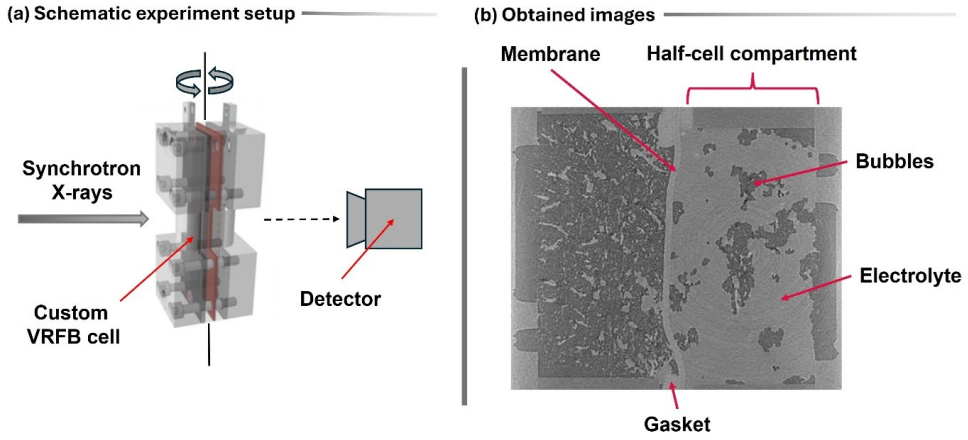

Figure 4.1: (a) Schematic experimental setup for measuring the synchrotron X-ray tomographies. (b) Exemplification of the obtained images from the VRFB cell and the regions of interest.

Rayon-based, graphitized carbon felt. All experiments used a 25% compression ratio to compress the nominal thickness from 6 mm to 4.5 mm. Before experimentation, the carbon felts underwent pretreatment in a muffle oven for 25 hours at 400°C in an air atmosphere to introduce active sites on the carbon surface.[29] The vanadium(IV) electrolyte was prepared by dissolving 0.1 molar $VOSO_4$ ($VOSO_4 \cdot 5H_2O$, chemically pure, GfE) in 2 molar H_2SO_4 (Suprapur®, Merck, diluted with purified Milli-Q water). This electrolyte was degassed with nitrogen to remove atmospheric air and then transferred to a syringe before each experiment. A standard procedure was applied to all examples for the HER experiments performed in this work. A potential of –0.5 V vs. Ag/AgCl (about –0.3 V vs. RHE) was applied for 5 minutes to induce the HER following the procedure indicated by Köble *et al.*[182]

The X-ray tomography scans were performed at the Biomedical Imaging and Therapy Bending Magnet (BMIT-ID) 05ID-2 beamline at the Canadian Light Source Inc. (Saskatoon, Saskatchewan, Canada).[183] A Budker superconducting wiggler generated high-intensity monochromatic synchrotron radiation with an energy of 30 keV in the storage ring.

The sample was positioned 58 m from the light source, with a 40 cm separation between the sample and the detector. An indirect detection scheme was employed, wherein the luminescent image produced by the X-rays was captured in a 500 mm thick YAG scintillator coupled with a CMOS camera (Orca Flash V2 sCMOS, Hamamatsu Photonics, Shizuoka, Japan), chosen for its fast decay time and low afterglow. The resulting horizontal field of

view was 26.6 mm, while the vertical field was limited to 8 mm due to the synchrotron beam's divergence. The estimated spatial resolution of the synchrotron beam was 40 mm.

Before each tomographic set, ten dark-field images (camera background, no beam) and ten flat-field images (background with beam, no sample) were recorded. During the tomographic process, the sample was rotated 200° around its vertical axis, with 2000 images for the three-dimensional tomographies captured within a 180° range to mitigate the effects of acceleration and deceleration during rotation. One image was taken every 0.09°. The images were reconstructed using the ufo-kit script developed at the Karlsruhe Institute of Technology.[184] To address ring artifacts in the raw images, low-pass filtering of sinograms in the frequency domain was applied before further reconstruction steps.

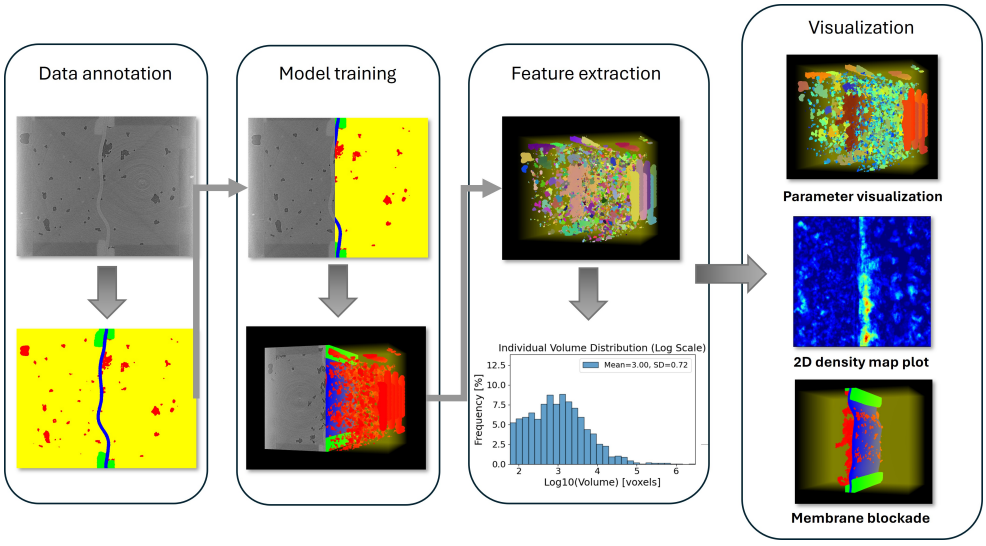

Figure 4.2: Methodological pipeline for the bubble analysis in VRFB tomographies.

Our methodology for autonomous analysis of bubbles within VRFB utilizing synchrotron X-ray tomography is shown in Figure 4.2. It manually annotates a comprehensive dataset comprising 2294 image/mask pairs. These pairs represent three distinct tomographic volumes with varying configurations: one half dry, half wet, and two thoroughly wet cells. This diversity in data ensures the model's adeptness at accurately analyzing the different potential conditions present within the VRFBs. Annotations were meticulously conducted on all three volumes using Dragonfly, a commercially available software that streamlines annotation, allowing thresholds to expedite data labeling.[185] Each slice was manually classified into four categories for detection: electrolyte, bubble, membrane, and gasket, as depicted in Figure 4.1 (b).

The dataset was divided into subsets for model training and validation. To optimize the training process, images were resized to 512 x 512 pixels, balancing resolution and memory requirements while preserving the images' contextual integrity. Post-prediction segmentation masks were resized to their original dimensions to maintain the correct aspect ratio. This research focuses on multi-class segmentation using variations of the U-Net architecture, employing the manually annotated masks as ground truth. The standard U-Net architecture, a fully convolutional neural network, executes convolutions and down-sampling to capture essential image patterns within a latent space during the encoding phase. This process condenses the image information into a singular vector, followed by a decoding phase that inversely reconstructs the image resolution through up-sampling and transposed convolutions. U-Net's distinctive feature is the concatenation of encoder feature maps with corresponding decoder up-sampled feature maps, enhancing the delineation of object boundaries by integrating both low- and high-level features. This approach yields a refined segmentation output, demonstrating improved object edges and contours recognition.[186]

Our study explored and evaluated four adaptations of the traditional U-Net architecture for image segmentation: U-Net with a ResNeXt101 backbone, Attention U-Net, U-Net 3+, and Swin U-Net.

- The ResNeXt101 U-Net variant integrates the ResNeXt101 model with the U-Net framework, utilizing the split-transform-merge strategy characteristic of ResNeXt101. This model is particularly noted for its "cardinality" feature, which refers to using parallel paths within a block to enhance performance without significantly increasing the network's depth or width. Integrating ResNeXt101 into U-Net enables more effective capture of complex image features, leveraging the strength of cardinality.[74]

- Attention U-Net enhances the standard U-Net architecture by incorporating an attention mechanism, focusing the model's processing on specific image regions. This is crucial for identifying objects at varying scales and minor yet significant image sections. By using attention gates before each decoder concatenation, the model emphasizes important pixels while diminishing the impact of irrelevant ones, thus achieving higher segmentation accuracy, particularly in noisy environments.[187]

- U-Net 3+ combines the original U-Net architecture with deep supervision methods, facilitating direct feature transfer between encoder and decoder at various levels. This method effectively bridges semantic gaps and maximizes the utility of hierarchical features, thereby enabling the detection of a wider range of image details from broad to fine.[188]

- Swin U-Net merges the Swin Transformer with the U-Net design, enhancing image segmentation by modeling long-range dependencies and capturing fine details across different scales. Incorporating the Swin Transformer's efficient self-attention mechanism with U-Net's established encoder-decoder structure results in superior segmentation precision. Swin U-Net's advantage lies in its ability to utilize global and local context information, offering a versatile and accurate approach to addressing segmentation tasks.[189]

Upon completion of the training phase, each model's performance was systematically evaluated and compared using established evaluation metrics.

After generating the segmentation maps for a given volume, we constructed a three-dimensional stack of these volumes to facilitate further statistical analyses. Utilizing established mathematical formulations, we extracted essential metrics from the battery cell, including total bubble volume, the bubble volume fraction R_{bv} defined as the ratio of bubble volume over the total negative electrode volume, and bubble density distribution across different planes. Additionally, we isolated individual bubbles and implemented advanced computer vision algorithms using the OpenCV library[190] to analyze the bubbles' shapes and spatial characteristics. These algorithms enhanced our comprehension of the bubbles' morphological features and their spatial arrangement within the cell by extracting the following bubble individual features:

- Volume: Calculated by the count of voxels within an individual bubble.
- Orientation: This is determined by computing the vector along the bubble's most significant axis, with polar coordinates derived from this vector and represented in a polar coordinate system.
- Flatness: This metric quantifies the degree to which a bubble's shape is flattened, defined as the ratio of the bubble's intermediate axis length to its shortest axis length,

$$\text{Flatness} = \frac{\lambda_2}{\lambda_3}, \tag{4.1}$$

  where the axes lengths are ordered in descending magnitude $\lambda_1 \geq \lambda_2 \geq \lambda_3$.
- Elongation: Measures how stretched a bubble is, expressed as the ratio of the bubble's longest axis length to its intermediate axis length,

$$\text{Elongation} = \frac{\lambda_1}{\lambda_2}. \tag{4.2}$$

- Sphericity: Assesses how closely a bubble's shape approximates that of a perfect sphere, calculated using the formula

$$\text{Sphericity} = \frac{\pi^{\frac{1}{3}} (6V)^{\frac{2}{3}}}{A}. \tag{4.3}$$

  V represents the bubble's total volume, and A represents its surface area.

- Wall Proximity: Indicates the distance in voxels from the bubble to the nearest cell boundary or membrane, facilitating the analysis of bubbles' spatial distribution within the cell.

In challenging examples, the predictive model occasionally introduced artifacts or noise into the segmentation maps. We implemented two corrective functions to mitigate these inaccuracies and preserve the integrity of our statistical analyses. Initially, a filter was applied across the entire 3D volume stack to eliminate detections of bubbles that appeared one-dimensional (1D), based on the fact that real bubbles should span at least two voxels in depth, with 1D detections likely arising from model-induced noise. Subsequently, a second function assessed the morphological characteristics of the bubbles, excluding artifacts characterized by significant non-convexity and irregular shapes, which provoked errors in the determination of bubble metrics.

The derived measurements were systematically visualized to elucidate the system's characteristics. Advanced 3D visualization techniques were developed using the Visualization Toolkit (VTK) library,[191] enabling interactive volume exploration with customizable color mappings based on the measured parameters. Furthermore, the software allows users to create and export short GIFs depicting the rotating volume, thus enhancing the presentation and visualization of the results.

4.3 Automatic DL-based 3D tomography segmentation and computer vision-aided bubble analysis

In the following sections, we evaluate and compare several deep learning architectures to achieve robust multi-class segmentation of VRFB tomograms. Based on the best model, we implement region-specific analysis and 2D bubble density mapping to quantify the spatial distribution across the positive and negative half-cells. We then extend our approach to detailed 3D morphological characterisation and interactive colour-coded visualisation of individual bubble properties. Finally, we demonstrate the utility of the entire pipeline by analysing how different electrode modifications affect hydrogen evolution and bubble formation in VRFB cells.

4.3.1 DL model benchmark and visual assessment of the multi-class semantic segmentation volumes

In analyzing tomographic images for vanadium redox flow batteries, achieving high segmentation accuracy is crucial for conducting a thorough examination. Our approach for segmenting the volumes into four distinct classes, i.e., bubbles, electrolyte, membrane, and gasket, starts with selecting an appropriate neural network architecture. To this end, we evaluated a series of models to identify the best performance. To further enhance the model's effectiveness, we have initialized each model with pre-trained weights derived from training on the ImageNet dataset, thus capitalizing on transferring previously acquired features. During the training phase, the primary goal is to minimize the loss function and quantify the discrepancy between the manually annotated ground truth and the model's predictions. The learning rate, critical for modulating the model's convergence speed, was optimally adjusted to dictate the step magnitude towards minimizing the loss function and dynamically reduced upon reaching loss reduction plateaus. A decrease in loss indicates improved performance of the task at hand. Post-training, the model's efficacy on unseen data has been evaluated using a validation set to ascertain the network's robustness and generalization capabilities.

We have explored four variants of the U-Net architecture: a U-Net model incorporating a pre-trained ResNeXt101 backbone, Attention U-Net, U-Net 3+, and Swin U-Net, based on visual transformers. These models have been trained using a dataset comprising 1894 training and 400 validation images across 200 epochs with a batch size of 8. The chosen optimizer was Adam,[192] and the selected loss function was categorical focal loss.[193] The learning rate has been dynamically adjusted in real-time, ranging from 10^{-4} down to 10^{-6}, depending on performance improvements across epochs.

Table 4.1 compares the performance metrics for each evaluated deep learning architecture. According to the F1-scores, all models show similar effectiveness on the validation dataset. However, the U-Net model with the ResNeXt101 backbone slightly surpasses others by achieving a higher recall. Maintaining high precision is crucial in our multi-class segmentation task, as misclassifying non-existent bubbles as present can introduce errors that potentially compromise subsequent statistical analyses.

Beyond quantitative metrics, we have performed a visual evaluation of model efficacy utilizing datasets not encountered during the training phase, presenting a range of configurations to test the models thoroughly. These configurations can essentially be categorized into four distinct scenarios: completely gas-filled (dry), half-filled (half-soaked), fully electrolyte-filled (soaked), and post-cell cycling (after hydrogen evolution reaction). In Figure 4.3, the models exhibit acceptable performance in the dry, half-wet, and post-

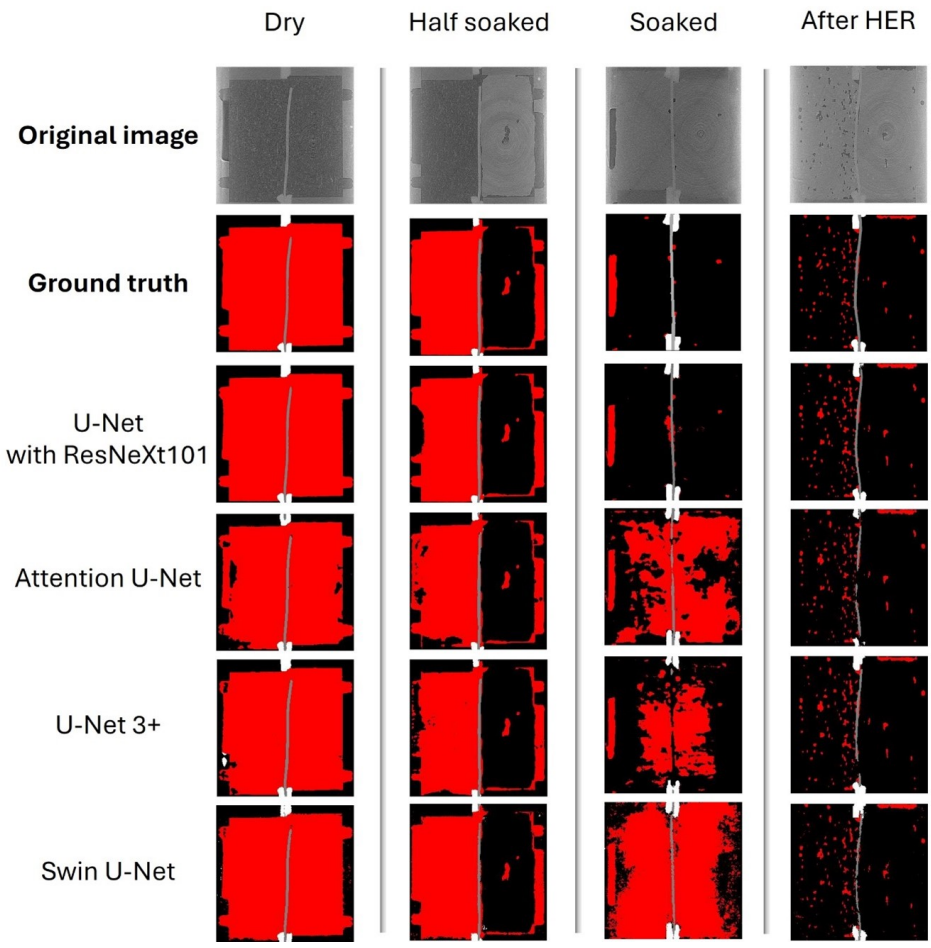

Figure 4.3: Comparative visualization of the original images, manual annotations, and benchmarked models across four distinct VRFB configurations with varying bubble/electrolyte ratios: completely gas-filled (dry), half-filled (half soaked), fully electrolyte-filled (soaked), and post-cell cycling (after hydrogen evolution reaction [HER]). These visualizations depict bubbles in red, membranes in grey, gaskets in white, and the electrolyte in black. This figure illustrates the models' segmentation accuracy in differentiating between these classes under various operational states of the VRFB.

HER scenarios, particularly noting that the U-Net model with a ResNeXt101 backbone achieved superior delineation of membranes and bubbles.

Furthermore, it should be noted that models encounter notable difficulties in configurations lacking or having sparse bubble representations. Under these conditions, models are prone to erroneously classifying empty electrode backgrounds as bubbles, resulting in a propensity for overprediction. This visual assessment further validates the exceptional

Table 4.1: Comparison of different U-Net models trained on a unique dataset. The performance of the models is assessed by employing precision, recall, and F1-score as metrics.

Model	Precision [%]	Recall [%]	F1-score [%]
U-Net with ResNeXt101 backbone	98	97	97
Attention U-Net	98	96	97
U-Net 3+	97	94	96
Swin U-Net	65	92	94

performance of the U-Net architecture merged with the ResNeXt101 backbone, which showcased a significant reduction in false positive detections of bubbles relative to other models. This model's adeptness at accurately distinguishing between true bubbles and background articulates its high level of discernment. It reinforces its applicability to complex segmentation tasks where precise detection and the reduction of false positives are paramount.

4.3.2 Positive and negative half-cell precise differentiation and multiperspective 2D bubble density maps

Automated tomography segmentation expedites the development of sophisticated software designed to extract advanced features from bubble dynamics within the cell. The parameters extracted play a critical role in deepening the system's understanding and facilitating experimentalists' research efforts. Moreover, these parameters are crucial for parametrization processes within simulation sciences, enabling more accurate and predictive modeling of cell dynamics.

A quantitative evaluation of the filtering process was conducted on an unseen dataset to assess its effect and influence on the predicted quantifications. Before filtering, the analysis detected 3255 bubbles, corresponding to a total bubble volume of 4.25%. After applying the filter, the bubble count was reduced to 488 while the overall bubble volume slightly decreased to 4.20%, a change of only 1.2%. These results indicate that the filtering process effectively removes noise and false positives without compromising the overall volumetric quantification, thereby ensuring reliable individual bubble analysis.

Furthermore, Figure 4.4 (a) demonstrates the software's ability to conduct region-specific analysis tailored to the user's requirements. Although analyses can encompass the entire cell, we have engineered specialized computer vision algorithms to discern between the positive and negative half-cells. To achieve this, we have delineated the membrane on both sides, enabling pixel-exact separation of the electrodes. This approach bene-

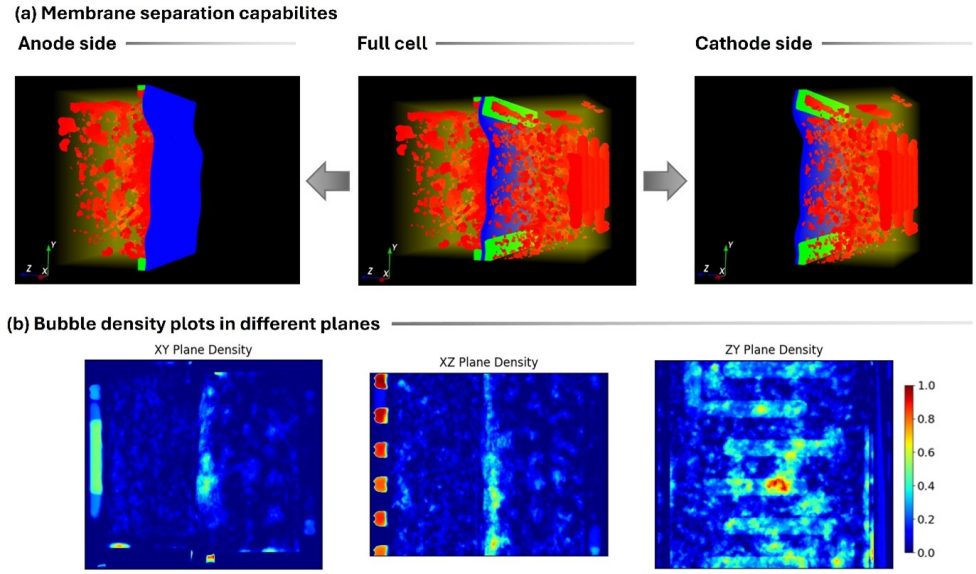

Figure 4.4: (a) Showcase of the precise electrode differentiation by the software. (b) Density maps depicting the bubble spatial distribution along the cell from three different planes, XY, XZ, and ZY. Areas with warmer hues represent high bubble density, while colder colors suggest less populated zones.

fits membranes with irregular shapes, which cannot be separated through conventional, arbitrary methods.

Identifying the region of interest can start the feature extraction process from the segmented volumes. Classifying voxels into four distinct categories makes it feasible to quantify the total gas volume within the cell and determine the bubble-to-electrolyte ratio to characterize the system.

An insightful outcome derived from the segmentation process is the generation of a 2D bubble density distribution plot across the cell, which is depicted as a density map from the orthogonal planes XY, XZ, and ZY, as illustrated in Figure 4.4 (b). We aggregate the predictions across each plane by filtering the predictions to retain only bubble voxels, set to '1', and designating the background as '0'. These density maps are then created by consolidating the segmented bubble data into a singular image from various perspectives. This aggregation results in regions densely populated with bubbles throughout the volume being assigned higher values, visually represented by warmer hues on the density map. Cooler colors indicate areas with fewer bubbles, while greener hues mark intermediate zones. The density maps from the XY and XZ planes in Figure 4.4 (b) qualitatively suggest that, in the given example, the bubbles are uniformly distributed across the cell,

with an increasing concentration near the central membrane. This visual representation aids in qualitatively assessing bubble distribution and accumulation within the cell from a 2D perspective, avoiding any 3D rendering approach.

4.3.3 3D morphological bubble analysis and 3D interactive color-mapped visualization of the properties of interest

Our methodology facilitates the rendering of volumes in three dimensions, enabling the subsequent extraction of specific features of interest. The process involves reconstructing a 3D volume, within which the segmented bubbles are isolated by a function that segregates each bubble surrounded by background voxels and assigns a unique identifier. This segregation permits the detailed analysis of each bubble, including extracting unique morphological attributes through computer vision techniques. These techniques provide insights into the structural properties of individual bubbles, elucidating characteristics commonly analyzed in material science.

These parameters collectively offer a comprehensive characterization of the bubbles' morphology, enhancing their understanding of their behaviour and interaction within the cell environment. As depicted in Figure 4.5, the derived measurements are subsequently visualized in corresponding histograms to provide an overview of the system. Furthermore, our methodology supports direct visualization of these measurements by rendering the 3D volume, where each bubble is color-mapped according to its measured values. This interactive 3D visualization capability significantly improves comprehension of property distribution throughout the cell and can be generated automatically. The software also features the ability to create 360° GIFs of the volume of interest, facilitating the later visualization or presentation of findings.

4.3.4 Voxel quantification of bubbly blockade at membrane level

Integrating automated volume segmentation with sophisticated computer vision algorithms has facilitated the creation of specialized functions tailored for specific use cases, enhancing our comprehension of the systems. By combining a function that delineates the detailed contour of the central membrane with precise, pixel-wise segmentation of bubbles, we have pioneered an approach for conducting membrane blockade analysis. As demonstrated in Figure 4.6, this methodology allows for quantifying and visualizing bubble voxels in contact with the membrane, laying the groundwork for subsequent correlation analyses. This function computes the extent of the membrane's surface area obstructed by bubbles and visualizes the implicated bubbles using an interactive 3D visualization tool, enabling detailed examination. The primary objective of this approach is to facili-

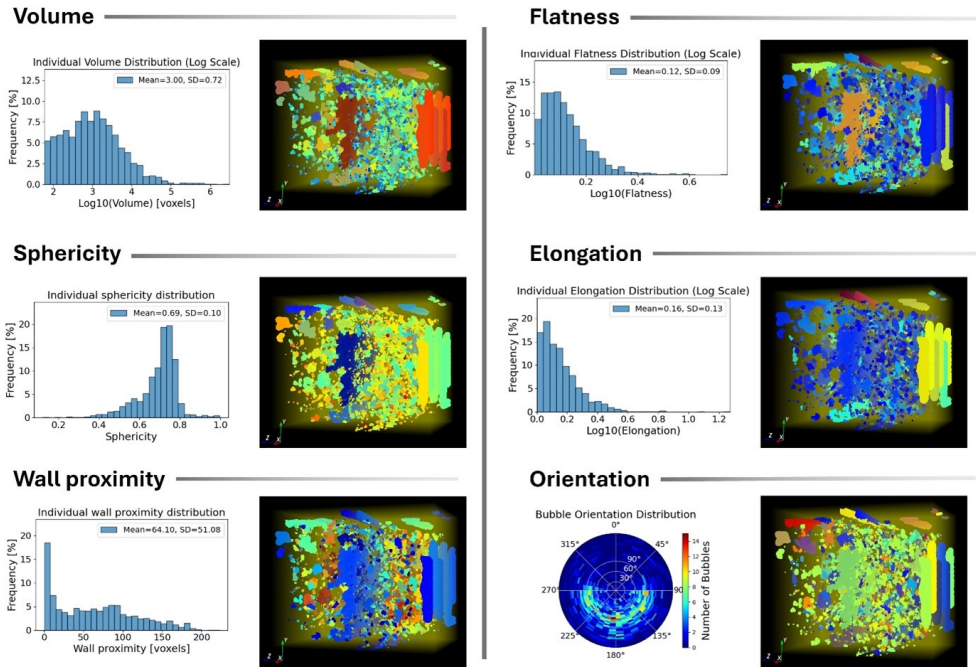

Figure 4.5: Analysis of individual bubble shape characteristics within segmented volumes. For each assessed morphological property, distribution histograms are plotted to detail the variance of properties across bubbles. Simultaneously, software-generated, color-mapped 3D visualizations represent these individual properties, where warmer colors indicate higher property values and cooler hues signify lower values.

tate future studies that will investigate the correlation between membrane blockage and potential performance degradation in the system. By quantifying the membrane area affected by bubble obstructions, we aim to identify and analyze factors that may contribute to performance losses, thereby offering insights into optimizing system efficiency and reliability.

4.3.5 3D DL-based tomography analysis of different electrode modifications and their influence on the HER reactions

We further investigated how different electrode modification methods affect the HER to illustrate our tool's capabilities. For this purpose, we employed a three-electrode setup featuring carbon felt electrodes modified with poly(o-toluidine) (POT), Fe–N–C, and Vulcan as working electrodes, a thermally activated carbon felt as the counter electrode, and an Ag/AgCl reference electrode (3M NaCl, E=0.209V vs. SHE).[162,182] The electrolyte was 0.1M $VOSO_4$ dissolved in 2M H_2SO_4. Before inducing hydrogen evolution, the elec-

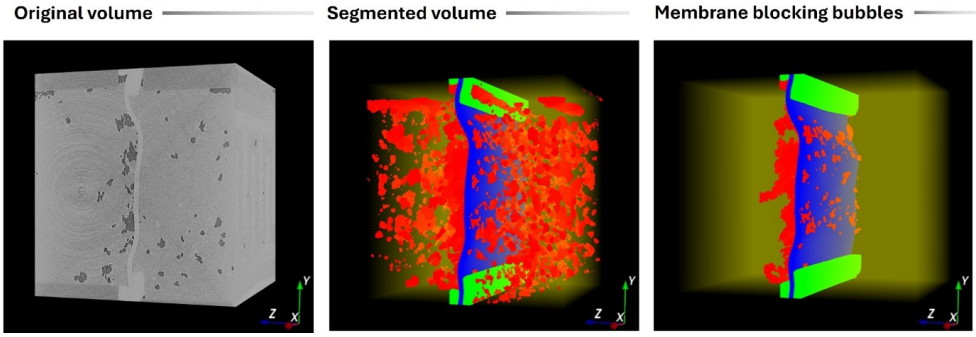

Figure 4.6: Visual comparison of the original volume alongside its segmentation and the representation of bubbles blocking the membrane.

trodes were thoroughly wetted and scanned once by X-ray tomography to identify residual bubbles in the electrode.

Figure 4.7 shows the resulting 3D reconstruction of bubble distribution and quantitative bubble analysis. The bubble volume fraction R_{bv} is defined as the ratio of bubble volume to the total negative electrode volume. In Figure 4.7(a), the R_{bv} in the POT modified electrode before HER was about 3.03%, indicating that a moderate amount of air was trapped even after complete wetting. Before HER, the R_{bv} in the Fe–N–C modified electrode was only 0.81% (as shown in Figure 4.7(b)), which means an excellent ability to remove gas phase and avoid large bubbles. In Figure 4.7(c), the Vulcan electrode had the highest residual bubble volume of 4.59% before HER, indicating that the additional carbon particles hindered the effective removal of gas during the permeation process. The bubble volume distribution histogram shows that the Fe–N–C modified electrode has mostly small and low-volume bubbles. In contrast, the Vulcan electrode tends to show a wider distribution. Moreover, the bubbles in the POT modified electrode and the Vulcan modified electrode show more obvious shape deformation (lower sphericity and higher elongation), suggesting that the local pore structure may limit the bubble release.

In addition, analysis of the bubble orientation reveals that in electrodes with a more anisotropic pore structure, bubbles tend to align along specific directions. This preferential orientation is especially evident in the POT and Vulcan modified electrodes, where the directional bias likely reflects the influence of the fibre alignment and pore connectivity on bubble nucleation and growth. These findings are consistent with the experimental data (dynamic vapor sorption and flow injection experiments) reported by Köble *et al.*,[182] indicating that the Fe–N–C modified electrode has good porosity and wetting behaviour under flow conditions.

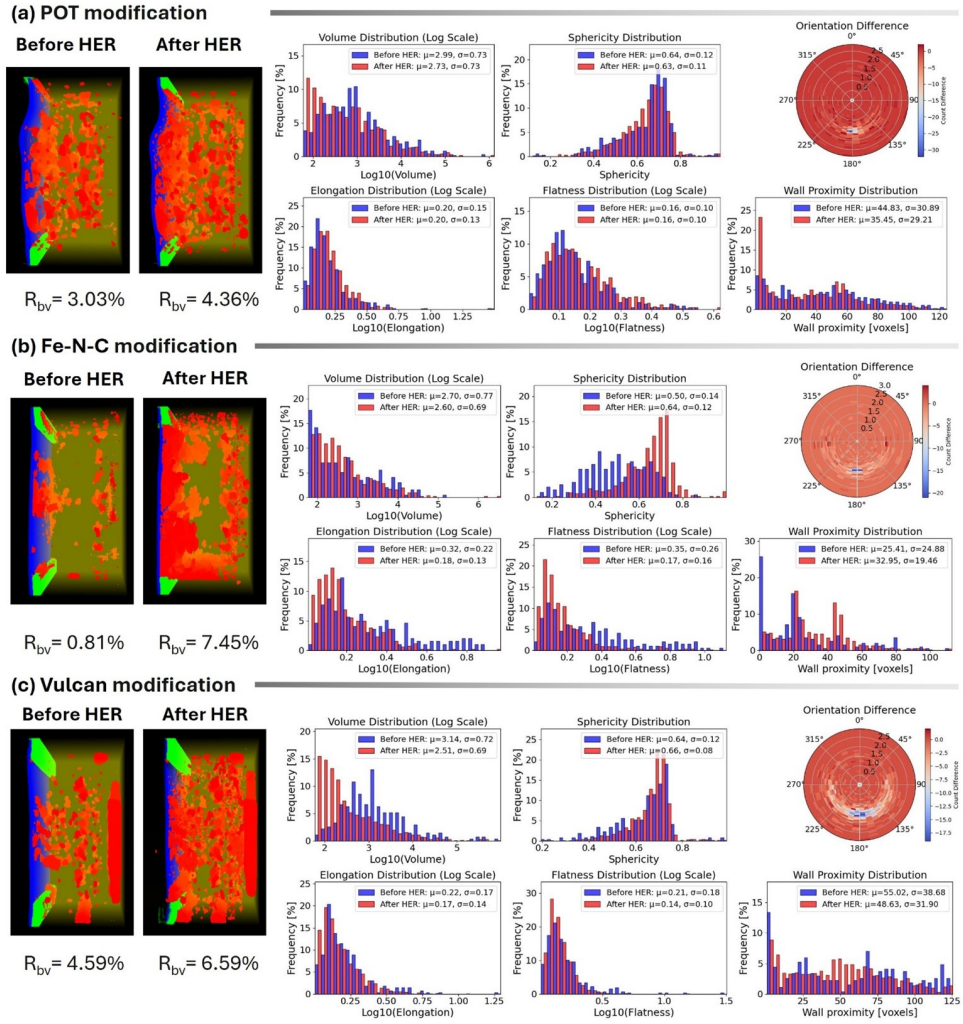

Figure 4.7: Bubble spatial distribution and bubble morphology characteristics of three modified electrodes: (a) POT modification, (b) Fe–N–C modification, and (c) Vulcan modification before and after the HER reaction induction. Each 3D rendering highlights gas bubbles in red, with the remaining electrode/electrolyte phase in contrasting colors. The bubble volume ratio (R_{vb}) is labelled for each sample. The histograms show the distributions of key bubble metrics, such as volume, sphericity, elongation, flatness, proximity to the electrode wall, and orientation, are superposed for comparison purposes in contrasting colors.

Subsequently, a potential of –0.5V vs. Ag/AgCl (about –0.3V vs. RHE) was applied for 5 minutes to induce the HER. A second tomography scan was then performed, and the results are compared against the preliminary tomographies in Figure 4.7. By comparing the bubble ratio before and after HER, we can quantify each modified electrode's catalytic or

inhibitory effect at negative potentials. For the POT modified electrode, the R_{bv} increased from 3.03% to 4.36%, a modest increase indicating that HER was partially suppressed, although the polymer coating caused some hindrance to mass transport. In contrast, the R_{bv} in the Fe–N–C modified electrode rises dramatically from 0.81% to 7.45%, highlighting the strong catalytic activity of the iron-based sites for hydrogen production, consistent with previous findings that Fe–N–C modified electrodes can significantly enhance side reactions at the anode of VRFBs.[182] The R_{bv} in Vulcan modified electrode increases from 4.59% to 6.59%, representing a modest increase of about 2%. Although not reaching the high increase of the Fe–N–C modified electrode, Vulcan still produces a large amount of additional gas, which can be attributed to carbon black agglomerates that promote bubble nucleation.

These tomographic results confirm the cyclic voltammetry and impedance analysis of Köble *et al.*,[182] showing that although the Fe–N–C modified electrode performs well in the V(IV)/V(V) reaction at the cathode, it significantly accelerates the HER at the anode. Meanwhile, the POT modified electrode seems to suppress the HER at the expense of partially blocking the transport pathway, while the Vulcan electrode is between these extremes. Overall, this study demonstrates how automated 3D segmentation and bubble morphology analysis can elucidate the initial gas removal efficiency and subsequent bubble growth triggered by hydrogen release, providing valuable insights for optimizing electrode design and mitigating parasitic reactions in VRFBs.

4.4 Conclusions

In conclusion, the work presented in this chapter establishes a generic framework for an expedited investigation into the effects of trapped bubbles in VRFB, utilizing 3D tomographic images as a rich data source. Through this study, we have successfully developed software designed for segmenting 3D synchrotron X-ray tomographic images into four pertinent categories: bubbles, electrolyte, membrane, and gasket. Moreover, our framework enables the autonomous differentiation between electrodes, quantifies bubble volumes, facilitates individual bubble shape analysis, generates 2D bubble density maps, and computes membrane blockage.

In our application example, we compared electrodes modified with POT, Fe–N–C, and Vulcan treatments. We demonstrated that our tool can efficiently distinguish differences in bubble dynamics before and after the HER. This use case reveals that, for example, the Fe–N–C modification exhibits a dramatic increase in bubble volume after HER, indicating strong catalytic activity for hydrogen production. In contrast, the POT modification shows a modest increase, likely due to its partial blockage of mass transport pathways.

These insights not only advance our understanding of bubble dynamics in VRFBs and inform electrode optimization to mitigate parasitic reactions, but they also underscore the broader potential of our approach.

Through the developed tool, experimentalists can now efficiently replace the challenging manual analysis of synchrotron tomographic images of VRFBs and expedite the quantification of observable bubbles to perform comprehensive analyses within a significantly reduced timeframe. Moreover, the software facilitates foundational research in physical modeling by automating the extraction of features such as individual bubble shape analyses and membrane blockade quantifications, tasks that would be impractical to perform manually.

Furthermore, the tool's advanced automation and visualization capabilities enable swift, visual comparisons across different volumes, setting the stage for comparative degradation analyses between various cells and potentially paving the way for 4D analyses of time-resolved tomographic images to monitor bubble dynamics over time. Indeed, this tool can be readily extended to other electrochemical systems, such as electrolyzers and fuel cells, where gas evolution is critical in performance and durability. We intend to make this tool publicly accessible, ensuring its design supports the seamless incorporation of additional feature extraction functions, new models, and various visualization techniques based on the needs of theoretical and experimental researchers.

Chapter 5

UTILE-Pore: Deep learning-enabled 3D analysis of porous materials in polymer electrolyte membrane technologies

Note: This work was performed during my research stay at the University of Toronto in the group of Aimy Bazylak, with further collaboration of Salvatore Ranieri, Jasna Jankovic and Tobias Morawietz. It culminated in the following publication: André Colliard-Granero, Salvatore Ranieri, Aimy Bazylak, Tobias Morawietz, K Andreas Friedrich, Jasna Jankovic, Michael H Eikerling, Kourosh Malek, and Mohammad J Eslamibidgoli. Utile-pore: Deep learning-enabled 3d analysis of porous materials in polymer electrolyte membrane-based energy devices. *Journal of The Electrochemical Society*, 172(7):074515, 2025. The personal pronoun "we" is used throughout this chapter to refer to the group of researchers who were part of this specific study. The employed GDL and MPL data were collected by me, while Jasna Jankovic and Tobias Morawietz provided the CL datasets. The research presented in this chapter was conducted by me with the assistance of Salvatore Ranieri, while the collaborators, along with Michael Eikerling, Kourosh Malek, and Mohammad Eslamibidgoli, played a supervisory role.

5.1 Introduction

As introduced in Chapter 1.1, electrochemical energy technologies that utilize polymer electrolyte membranes (PEMs), e.g., polymer electrolyte fuel cells (PEFCs) and water electrolyzers (PEWEs), offer high current densities, minimal reactant crossover, and vol-

umetrically compact designs, making them a promising option for hydrogen generation and conversion as compared to conventional alkaline systems, which typically operate at lower current densities and suffer from higher gas crossover rates.[195–197] Nonetheless, the high costs associated with precious metal catalysts and the need for enhanced mass transport remain significant barriers to broader commercialization. These issues are especially prevalent in both the catalyst and gas diffusion layers, routinely found within PEM electrodes. For this reason, understanding how their function and structures are related is critical for describing transport properties such as thermal and electrical conductivity and improving cell performance and degradation.[198–201]

The electrode layers in these cells are highly porous and exhibit a multi-scale characteristic length progression from a few nanometers within the CL to tens of microns within the GDL substrate, which supports a hierarchy of functions to minimize voltage losses of these cells. For instance, GDL porosity plays a critical role in governing reactant transport towards the CL while maintaining humidity, directly affecting the mass transport and proton conductivity, respectively.[202,203] On the other hand, within the CL, the porosity spans lengths from several to tens of nanometers to maximize the accessible electrochemical surface area, while maintaining high proton conductivity and reactant gas diffusivity.[204] On the other hand, within the CL, the porosity spans lengths from several to tens of nanometers to maximize the accessible electrochemical surface area, while maintaining high proton conductivity and reactant gas diffusivity.[205–207] Pore sizes in MPL span the range from hundreds of nanometers to a few micrometers, being an order of magnitude smaller than pores in the GDL. However, wide range of feature sizes represents a challenge when attempting to resolve the three-dimensional (3D) geometries with traditional imaging segmentation techniques.

Common techniques for resolving the 3D structure of these porous layers include micro-computed tomography (micro-CT), synchrotron tomography, and focused ion beam scanning electron microscopy (FIB-SEM).[208–210] However, translating volumetric data into quantitative properties requires accurately classifying each voxel as pore or material, which is an especially difficult task given the complex nature of pore networks.

Figure 5.1 displays three popular strategies for segmenting pore structures from tomographic data. Manual annotation is often used for simpler examples. In this case, 3D visual clues are employed to distinguish the different classes. AI can be employed to train a model based on manual annotations. The disadvantage of this method is the difficulty in asserting by eye a correct classification of the pores, resulting in a highly biased and time-consuming result.[211,212] On the other hand, Otsu thresholding is the most popular method for the segmentation of this type of tomographic volume. Often, a median fil-

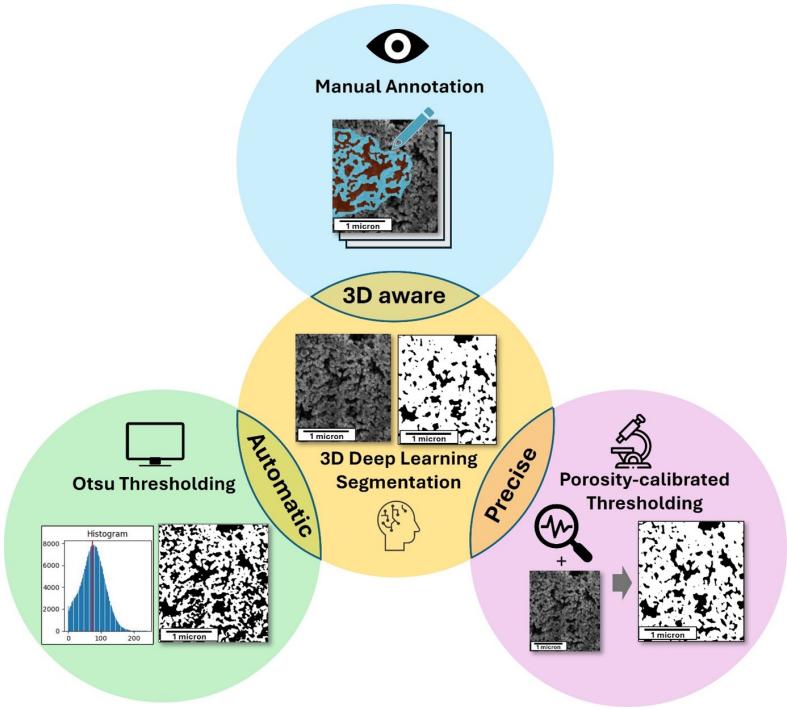

Figure 5.1: A schematic comparison of available methodologies for pore segmentation is presented on the example of catalyst layers.

ter is applied to the volume to smooth the images and reduce possible noise, and then the Otsu algorithm is applied to differentiate between the pores and the materials. The Otsu algorithm is a binarization algorithm that automatically finds the optimal threshold value by minimizing the intra-class variance between the black and white pixels. The disadvantage of this method is the lack of control over the output, which is solely based on the mathematical separation of the grayscale histogram, rather than corresponding to the real material structure.[213,214] Finally, porosity-calibrated thresholding is also widely used by experimentalists, and it introduces the employment of a physical characteristic of the sample of interest to tailor the thresholding toward a plausible output. For this calibration, the porosity of the sample must be measured externally, and then the threshold is adjusted to match the measured properties. The main issue of this technique is the need for an extra step with suitable equipment to measure the physical porosity, with an extra complexity layer of different methods yielding different results.[215,216] Furthermore, all three methods are limited to binary segmentation (i.e., solid vs. pore), preventing the distinction of additional layers such as the microporous layer (MPL) from the GDL.

Deep learning (DL)-based image analysis has gained increasing attention in research on PEM-based devices in recent years. Several studies demonstrated the utilization of DL to automate the characterization, ranging from the automation of particle size distribution analysis of catalyst nanoparticles from TEM images, to oxygen bubble dynamics elucidation from optical videos of flow fields, or direct permeability calculation from 3D tomographs of gas diffusion layers.[59,61,131,217]

Algorithms such as DL and computer vision (CV) also emerge as an effective alternative for automating annotation, segmentation, and analysis tasks in binary and multi-class settings for accelerating imaging analysis on energy materials.[218–220] As depicted in Figure 5.1, the employment of 3D DL models combines the best characteristics of each method; it enables the automatic segmentation of volumes, it is aware of the 3D nature of the structures, and it has the precision learned from porosity-calibrated datasets. DL is especially efficient in cases of large volumes of imaging data where the time-consuming nature of manual analyses represents significant bottlenecks.[221,222] Although DL and CV techniques can automate feature extraction across various domains, the diversity of research fields and the distinct characteristics of their imaging data such as differences in resolution, contrast, noise levels or employed equipment often necessitate the development of tailored AI models that incorporate specialized network architectures alongside domain-specific preprocessing and post-processing steps to effectively address each field's unique tasks.[32,59,61]

For the specific task of segmenting GDLs, CLs, and MPLs, multiple studies have explored conventional and machine learning-based approaches for the specific task of segmenting GDLs, CLs, and MPLs. Pfrang *et al.* approached the task conventionally by applying diffusion filtering and grayscale thresholding to separate GDL and MPL and analyze MPL thickness at the microscale.[223] Nevertheless, conventional image processing methods and grayscale-based algorithms struggle when pixel intensities lack a clear bimodal distribution or exhibit gradual intensity gradients. By employing DL, Mohidivara *et al.* compared two- and three-dimensional U-Net architectures against manual and semi-automatic annotation procedures for multi-class segmentation in wet GDL tomographies.[212]

Shum *et al.* investigated multi-class machine learning-based segmentation and basic image processing algorithms for water segmentation in GDLs. Their results highlighted that a limited dataset size can hinder the full potential of AI-based methods, emphasizing the need for large, well-annotated training sets.[224] However, while supervised machine learning approaches are advantageous in integrating pixel intensity with structural features, they often encounter the challenge of annotating complex porous structures, which can be both time-consuming and prone to user subjectivity. Moreover, synthetic datasets

have been employed to train deep learning models for separating GDL fibers from binder, streamlining data acquisition, but potentially compromising realism by introducing uncertainty when the models are applied to real images.[225]

Building on these advances in learned, multi-scale pore segmentation, we now introduce UTILE-Pore. This 3D deep learning-based framework integrates porosity-calibrated masks in model training to replicate calibrated 3D patterns for test data, eliminating the need for additional measurements. The framework also supports multi-class segmentation to differentiate between the MPL and the GDL substrate. The resulting pipeline reduces the time required for analyzing 3D volumes of porous media and performing porosity measurements to mere seconds, enabling additional class segmentation. Moreover, users can easily replace the pre-trained models with their custom models, broadening the framework's applicability.

5.2 3D tomographic data acquisition and porosity-calibrated DL-based segmentation workflow

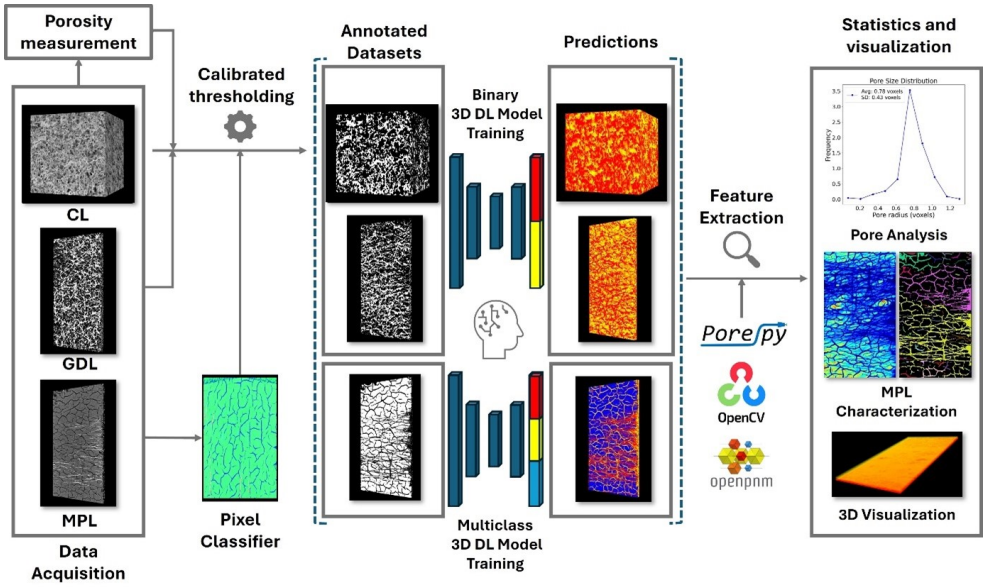

Figure 5.2: Methodological pipeline for the porous media analysis of CLs, GDLs, and MPLs from 3D micro-CT and electron microscopy-based tomographies.

Figure 5.2 illustrates the workflow developed in this work for autonomously analyzing porous structures in CLs and GDLs with and without MPL. The process begins with

acquiring a robust dataset for training deep learning models. In total, 33 samples comprising various CLs and GDLs samples from commercially available PEM components were collected. The dataset for model training was compiled by acquiring volumetric data from FIB-SEM of CLs and micro-CT measurements of GDLs with and without an MPL. Our primary goal is to characterize the porous architecture of CL and GDL by distinguishing between pore and material phases. For the GDL, a further distinction between fibers and PTFE binder from the collected datasets is not feasible due to nearly identical X-ray attenuation, and sub-micron PTFE coatings fall below our voxel size and cannot be reliably resolved. Furthermore, although the MPL is itself a porous medium, its pore sizes lie below the 3 µm voxel resolution limit of micro-CT scans. Consequently, resolving both GDL and MPL pores simultaneously with micro-CT is not feasible. To maintain accuracy, we therefore treat the MPL as a homogeneous solid and instead focus on macroscale MPL descriptors: crack density, layer thickness, and intrusion depth into the GDL, using the same imaging data. Due to the similarities between MPL and CL the same algorithms could be used on FIB-SEM data for the MPL in future work.

FIB-SEM measurements on catalyst layers were recorded following a consistent FIB-SEM protocol. A Zeiss Crossbeam 350 with Atlas Engine for 3D Tomography was used for the FIB-SEM procedure. The single-sided catalyst-coated membrane was glued with the cathode facing up using conductive carbon pads, while the catalyst layer was connected with silver glue to the SEM stub. The FIB tomography was carried out at 54° sample tilt (-36° image correction applied) and at a working distance of 5.1 mm (coincidence point of electron and gallium beam). All FIB operations were performed at a voltage of 30 kV. Before starting the tomography, the gas injection system (Platinum and Carbon) was used to deposit a platinum protective layer with the FIB beam at 700 pA, followed by milling of tracking marks at 100 pA with the FIB beam. Finally, these tracking marks were filled with carbon, and an additional carbon layer was deposited on top to protect these structures. These preparation steps were followed by milling the coarse trench with 7-nA FIB beam and milling the fine trench with a 1.5-nA FIB beam. Secondary electron (E-T) and material contrast (ESB) detectors were used for grabbing the slice images with a thickness of 10 nm per slice and 10 nm pixel size, resulting in 10 nm^3 voxels. The milling beam was chosen to be 700 pA at 30 kV. Drift compensation, auto contrast, and autostig were activated during the run. The slices were taken and interpolated to exactly 10 nm per slice with Zeiss Atlas software.

Meanwhile, the conducted micro-CT measurements for both plain and microporous-layer-containing gas diffusion layers were acquired employing a ProCon X-Ray CT-MINI (ProCon X-Ray GmbH, Sarstedt, Germany) using an X-ray source voltage and current of 30 keV and 200 µA, respectively, at a working distance of 38.11 mm, resulting in a voxel

resolution of 5 µm. A total of 2000 radiographic projections were acquired over a 360° stage rotation, and later reconstructed using VGSTUDIO MAX (Volume Graphics GmbH, Heidelberg, Germany). Each dataset was a reconstructed multipage TIFF file, only containing relevant slices with an average height and width of 1000 pixels. The depth of the datasets was highly dependent on the sample, which had an average of 40 to 60 slices for the GDL examples and an average of 700 slices for the CL examples.

First, the physical porosity was measured for each sample to generate the training masks for binary segmentation. Mercury injection porosimetry was employed for the porosity measurement of CL datasets, while a pycnometer was used for the GDL micro-CT datasets. For GDLs containing MPL, additional steps were required: First, a pixel classifier (Labkit-Tool in ImageJ)[226,227] semi-automatically segmented the MPL; second, the MPL voxels were isolated and removed from the volume, leaving only fibres and pores to be calibrated against the measured porosity. Finally, the removed MPL voxels were reintroduced into the volume, resulting in three distinct categories: MPL, fibres, and pores. The physical porosity used to calibrate each segmentation threshold has an absolute repeatability of ±2 percentage points for mercury intrusion porosimetry (CL samples) and ±0.5 percentage points for gas-pycnometry (GDL samples). Because the threshold shifts by only the instrument's error band, only voxels very close to the pore–solids interface can flip labels, a fraction too small ($< 2\%$) to affect overall training or performance ranking. Consequently, we train all models on the nominal masks without additional threshold jitter.

The next step is to preprocess the volumetric data to prepare it for the model training. Since the tomographic datasets varied in size and 3D models are computationally expensive, each volume was divided into 96×96×96 voxel cubes. Any cube at the periphery that could not fulfill the standard size was expanded via mirroring to fill the missing regions. 3666 cubes were generated for the binary case and 847 for the ternary case.

To handle the binary and multi-class tasks, we developed two distinct models: a binary model for segmenting GDLs and CLs and a multi-class model for distinguishing GDL fibres, pores, and MPL. For both cases, we benchmarked a selection of three well-established 3D architectures from the literature to identify the most effective one for each case: 3D U-Net with a ResNeXt101 backbone, 3D V-Net, and 3D Swin U-Net, with the calibrated masks serving as ground truth.

- 3D U-Net with a ResNeXt101 backbone combines the fully convolutional 3D U-Net encoder–decoder structure with the ResNeXt101 backbone for feature extraction. The standard 3D U-Net employs skip connections to merge low- and high-resolution features, improving boundary delineation. ResNeXt101 integrates a split-transform-

merge strategy called "cardinality," which provides multiple parallel paths within each block without increasing network depth or width. Incorporating ResNeXt101 into the 3D U-Net thus enables more efficient capture of complex spatial patterns.[74,117]

- 3D V-Net architecture retains the U-Net encoder–decoder framework with skip connections but replaces standard convolutional layers with residual blocks. These residual connections aim to expedite convergence and enhance segmentation accuracy, particularly for volumetric data where continuity along the depth dimension is crucial. In addition, this design may require fewer training samples than non-residual counterparts.[72]

- 3D Swin U-Net combines the U-Net encoder–decoder layout with Swin transformers, which effectively employ window-based self-attention to capture local and global dependencies. Unlike convolution-based networks, which rely on fixed receptive fields, Swin transformers dynamically adjust receptive fields based on dataset requirements. This flexibility enables more robust modeling of long-range spatial relationships in 3D volumes.[73]

During the training phase, the primary objective was minimizing the loss function, quantifying the discrepancy between the predictions and the porosity-calibrated ground truth. The 3666 binary and 874 ternary samples were split into an 80:20 ratio for training and validation. Each complete pass through the training dataset, called an epoch, refines the model's parameters based on the computed loss. Each model was trained over 100 epochs with a batch size of 4, using the Adam optimizer and binary or categorical cross-entropy as the loss function. Convergence occurs once the loss reaches a minimum, guided by an appropriately tuned learning rate. The learning rate was dynamically reduced from 10^{-4} to 10^{-6} once the loss reached a plateau, but constrained to avoid dropping to negligible values. After each epoch, the model's performance was assessed on the unseen validation set to evaluate its robustness and generalization ability. After the training phase, we systematically evaluated the performance of each model using standard metrics. To improve the positional robustness of the models, we augmented the employed data with on-the-fly random flips in the x, y, and z directions. Furthermore, the memory-demanding 3D model training with a batch size of 4 was parallelized on 4 Nvidia GPUs V100 with mixed precision protocols. Additional scripts and the employed datasets are available at the GitHub and Zenodo repositories cited in the data availability statement.

Once the model was trained, the input volumes also needed to be cropped into cubes of $96\times96\times96$ voxels to predict new examples. The cubes are then predicted by the model, and subsequently, they need to be repatched back to their original shape. To ensure a

smooth transition between the patches, the cubes are cropped with a certain overlap, and by repatching, we employ a Gaussian weight to blend the predictions.

After generating the segmentation volumes from the input data, we developed specialized computer vision and mathematical functions to perform additional analyses and extract quantitative information. These analyses are grouped into two categories: (1) functions for binary segmentations (i.e., GDL and CL), and (2) functions for multi-class segmentations (i.e., the interplay between MPL and GDL). For the binary segmentations, we introduced methods to measure porosity and surface roughness, and we integrated PoreSpy[228] for pore size distribution analysis and tortuosity simulation, which can be computationally intensive but benefits from GPU acceleration. Estimating permeability is more complex and typically requires manual parameter adjustments; thus, our tool employs the Kozeny–Carman equation as an approximation.[229] This equation describes the pressure drop of fluid flow through a packed bed of solids and is expressed as:

$$k = \frac{\varepsilon^3}{C\,\tau^2(1-\varepsilon)^2\,S_v^2}. \tag{5.1}$$

Where ε denotes the porosity, τ the tortuosity, S_v the specific surface area, and C the Kozeny constant. The Kozeny–Carman equation allows a permeability estimate by incorporating porosity and tortuosity values extracted from our predictive models. The Kozeny constant C, typically 4 to 5, encapsulates geometric factors, including pore packing and particle shape. The specific surface area is derived by converting the segmented structure into a mesh and measuring the polygonal surface areas.

Because the Kozeny–Carman equation provides only an approximation, and our goal is to support more advanced simulations, we included the functionality to export the segmented volumes directly into PoreSpy and OpenPNM[230] formats for further analyses. The surface roughness is calculated by cropping the upper layers of the binary volume, converting them into a mesh, and measuring depth variations at each vertex to quantify roughness across the entire structure.

For the MPL–GDL configurations, we created functions to analyze MPL cracks, map variations in MPL thickness, quantify the extent of MPL intrusion into the GDL, and calculate the surface areas of both the MPL and GDL. Specifically, for the MPL intrusion quantification, we employed the advantages of having a precise segmentation of the MPL and the power of CV to measure the thickness variations through the layer. For this reason, we first measured the total roughness of the MPL facing the GDL. Subsequently, we divided the MPL into smaller regions and calculated the local roughness of each patch. The standard deviation of the local roughness gave an insight into the depth variation of

the MPL through the layer. Additionally, we calculated the dimensionless coefficient of variation c_{R_a} for comparability purposes, given by,

$$c_{R_a} = \frac{\sigma_{R_a}}{R_a}. \tag{5.2}$$

With c_{R_a} representing the relative variability of the roughness, σ_{R_a} being the standard deviation of the surface roughness, and R_a the total roughness.

All derived measurements are automatically visualized in 3D using the Visualization Toolkit (VTK)[191] library, enabling interactive exploration of the predicted volumes. The software also supports generating short GIFs of the rotating 3D volumes, enhancing the ease of communication and presentation of the results.

5.3 Binary and multi-class autonomous semantic segmentation of CL, GDL, and MPL and computer vision-guided pore structure elucidation and quantification

In the following sections, we present a comprehensive application of our 3D segmentation pipeline: First, we benchmark several deep learning architectures on both binary and multi-class pore segmentation, evaluating F1-scores and porosity accuracy against physical measurements; next, we demonstrate binary feature extraction workflows to derive porosity profiles, pore size distributions, permeability, and tortuosity from segmented CL and GDL volumes. Finally, we leverage multi-class results to quantify MPL intrusion into GDL substrates and perform detailed crack analysis, demonstrating how these advanced metrics inform material design and performance optimization.

5.3.1 Three-dimensional DL architecture benchmark for binary and multi-class pore structure segmentation

Table 5.1 compares performance metrics for all binary models, with results indicating that the models performed similarly over the other architectures based on F1-scores. To further evaluate real-world applicability, an unseen GDL volume with a physically measured porosity of 73.7% was segmented by each model. For a better comparability of the models based on To evaluate the accuracy of the computed parameters, we added

columns to depict the porosity difference Δ_p, the tortuosity difference Δ_τ, and the specific surface area difference Δ_{SSA}, as defined in equations (5.3)–(5.5):

$$\Delta_p = |p_{\mathrm{real}} - p_{\mathrm{pred}}|\,, \tag{5.3}$$

$$\Delta_\tau = |\tau_{\mathrm{real}} - \tau_{\mathrm{pred}}|\,, \tag{5.4}$$

$$\Delta_{\mathrm{SSA}} = |\mathrm{SSA}_{\mathrm{real}} - \mathrm{SSA}_{\mathrm{pred}}|\,. \tag{5.5}$$

Here, Δ_x denotes the absolute deviation between the predicted parameter x_{pred} and the reference value x_{real}, which is derived from the physically calibrated volume.

Table 5.1 compares three 3D convolutional neural network (CNN) architectures on a binary pore-solid segmentation task. All models achieve similar F1-scores, but the 3D V-Net shows the closest match to physical porosity measurements. Although the 3D U-Net records the highest F1-score, the V-Net yields the smallest absolute porosity error ($\Delta_p = 1.7\%$), roughly half the U-Net's 3.9%. Here, Δ_p denotes the absolute difference between measured and predicted porosity, and values below the pycnometer's $\pm 0.5\%$ repeatability lie within calibration noise. The V-Net's superior performance stems from its volumetric residual blocks and symmetric encoder-decoder structure, which preserve fine boundary details and sharpen the fiber-void interface. Its larger effective receptive field and dense skip connections reduce over-segmentation of narrow throats without compromising tortuosity deviation (0.44%) or specific surface area deviation ($0.78 \times 10^{-2}\,\mu\mathrm{m}^{-1}$).

This improved porosity fidelity leads to visibly sharper segmentations in Figure 5.3. In the catalyst-layer (CL) scenario, all networks reconstruct the overall pore network. Still, the V-Net maintains the finest solid–void boundaries and preserves thin material regions near the edges. The U-Net delivers comparable bulk segmentation, whereas the Swin U-Net's windowed attention sometimes fractures the smallest CL pores. In the gas diffusion layer (GDL) case, each model captures the high-porosity texture so uniformly that visual assessment alone is insufficient; instead, quantitative metrics drive the evaluation. Overall, the 3D V-Net provides the best combination of segmentation accuracy and transport-property fidelity for binary pore–solid applications.

To benchmark the method against standard laboratory practices, we also compared predicted porosity to the physical measurement and the porosity derived from Otsu thresholding. The Otsu-based approach yielded a Δ_p of 5%, deviating the most from the physical value compared to the evaluated models, and a tortuosity deviation over twice that of the CNNs (Δ_τ= 1.14%), even though the SSA error may appear low. Thus, demonstrating

that the model-based prediction was more accurate in the delineation of the fibers in a 3D fashion.

Table 5.1: Comparison of different 3D deep learning architectures trained on a unique binary dataset of CLs and GDLs. The performance of the models is assessed by employing the standard metrics of precision, recall, and F1-score, and additionally by the porosity difference Δ_p, the tortuosity difference Δ_τ, and the specific surface area difference Δ_{SSA}.

Model	**Prec.** [%]	**Rec.** [%]	**F1** [%]	Δ_p [%]	Δ_τ	Δ_{SSA} [$\mu\mathrm{m}^{-1}\cdot10^{-2}$]
3D V-Net	91	89	90	1.7	0.44	0.78
3D U-Net	89	91	91	3.9	0.03	0.52
3D Swin U-Net	83	88	85	2.4	0.53	0.46
Otsu thresholding	–	–	–	5.0	1.14	0.44

Table 5.2 presents the multi-class segmentation results on GDLs containing MPL. The models were trained using 847 cubes of size 96×96×96 voxels extracted from 10 tomographs of commercially available GDLs with MPL, employing the same training parameters as before. The metrics indicate a high similarity in the performance of the diverse models. To validate the highest performing architecture, we compared the predicted porosity of the GDL (excluding the MPL) with the physically measured porosity for a sample without MPL. The 3D Swin U-Net achieves the smallest porosity error of Δ_p of 3.3%, despite having a lower overall F1. We hypothesize that its windowed self-attention mechanism excels at distinguishing the thin MPL cracks and the subtle grey-level differences at the MPL/GDL interface, features that can be easily lumped together as "solid" by pure convolutional models. By capturing long-range correlations across the volume, Swin U-Net more accurately segments these multi-class boundaries, giving it the edge in porosity fidelity for three-phase samples. This is confirmed in Figure 5.3, where Swin U-Net most faithfully segments the fine MPL cracks and cleanly separates the two layers; by contrast, V-Net tends to smooth narrow fissures at the interface, and U-Net mislabels small MPL regions as GDL material, slightly blurring the layer boundary.

Table 5.2: Comparison of different 3D deep learning architectures trained on a unique binary dataset of CLs and GDLs. The performance of the models is assessed by employing the standard metrics of precision, recall, and F1-score, and additionally by Δ_p, which is the positive difference of the predicted porosity from the physically measured porosity.

Model	**Precision** [%]	**Recall** [%]	**F1-score** [%]	Δ_p [%]
3D V-Net	83	92	88	9.8
3D U-Net	88	90	89	8.4
3D Swin U-Net	85	90	88	3.3

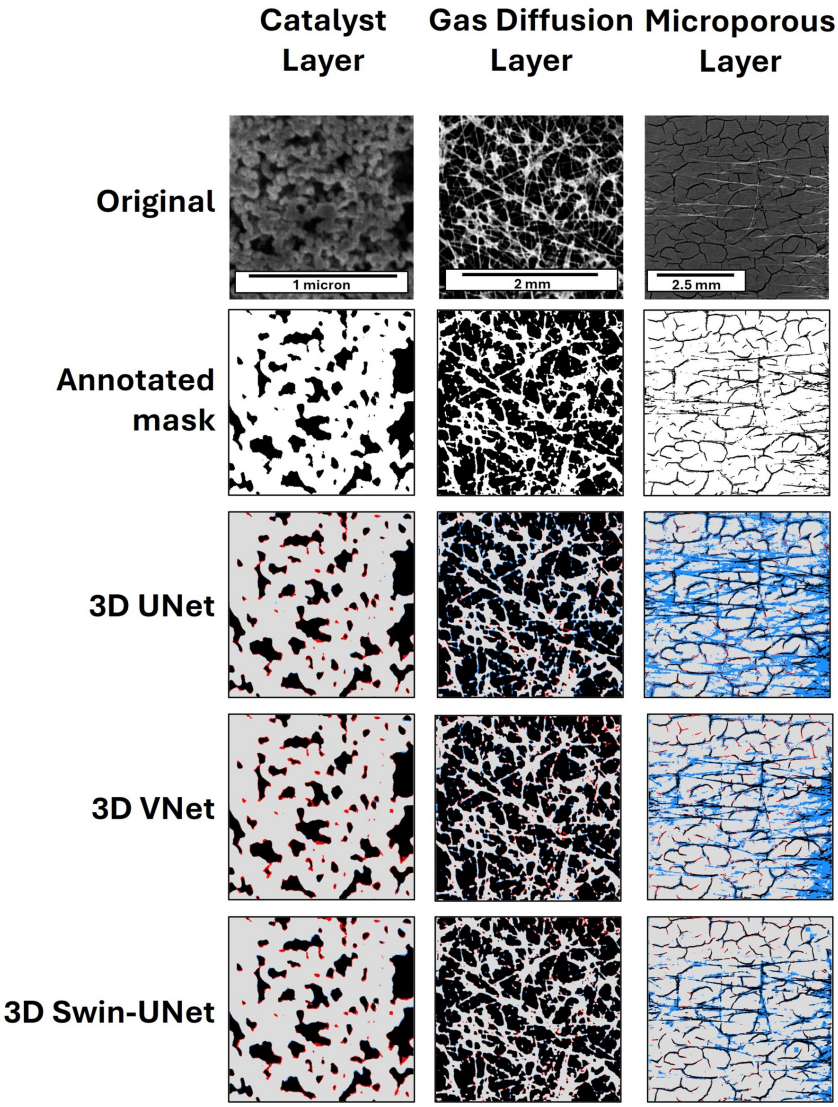

Figure 5.3: Comparative visualization of the original images with the corresponding annotated mask for the three types of datasets in consideration in this work: CLs, GDLs, and GDLs with MPLs. Gray pixels indicate true positives for correctly detected material, blue pixels represent false negatives where material pixels are missed, and red pixels signify false positives for erroneously detected non-material pixels.

The implementation of an automatic tomographic volume segmentation approach based on porosity-calibrated volumes combines the precision of the thresholding approach based on real physical information from the sample with the advanced 3D pattern recognition

capabilities of 3D deep learning architectures. The model is trained to analyze the three-dimensional context of the materials and precisely delineate the contours of the pores based on the physical properties of the sample. This approach not only removes the need to calculate the physical porosity of each analyzed volume but also enhances the segmentation on complex cases where thresholding may fail.

5.3.2 Feature extraction for pore-material structures in CL, GDL, and MPL

Furthermore, we developed a unified software that enables experimentalists to quantify structural properties in all scenarios using computer vision algorithms. In Figure 5.4, we showcase the implemented functions for binary segmentation, such as an algorithm to count pore and material voxels slice by slice and thereby estimate the sample's porosity from the model's predictions, enabling the porosity profile plots. We further integrated functions from the PoreSpy library to characterize pore sizes and generate pore size distribution plots, offering insights into the pore structure of each tomograph. These findings improve material-specific understanding and aim to help experimentalists correlate the structures with other present phenomena, such as wettability or water retention.

The CL presents a clear bias in the porosity profile towards the membrane with a more homogenous porosity towards the bulk of the structure, where we see a sharp increase in porosity from 50-65% within the first 100 slices of the CL, followed by a consistent 65% porosity thereafter. Again, this porosity inhomogeneity is confirmed by the pore size distribution plot, which shows one main peak, but over a wider range of voxel radii with a minor modal bias. The left-skewed bias indicates the presence of an appreciable number of smaller pores at 4.84 voxels with a standard deviation of 3.20 voxels. The main peak is broad, indicating a wide range of pore sizes with a significant number of smaller pores, particularly concentrated near the low-porosity membrane side. Similarly, our predicted porosity analyses show strong resemblance to CLs reported in the literature. Berejnov *et al.* leveraged Scanning Transmission X-ray Microscopy (STXM) on PEFC CLs and demonstrated strong porosity gradients between the membrane and the surface of the CL, revealing a low porosity at the membrane interface caused by an abundance of ionomer.[231] The bimodal pore size distributions further confirm the validity of our technique, and have been both experimentally observed and used for stochastic generation of CLs.[232,233]

In the case of GDLs, we observe in the porosity profile a certain homogeneity in the porosity through the slices, around 62-76%, while the pore size distribution analysis confirms the regular pore distribution by showcasing a predominantly single normal bell at pore sizes of 5-6 voxels with a standard distribution of 2.01 voxels. Our predicted porosity

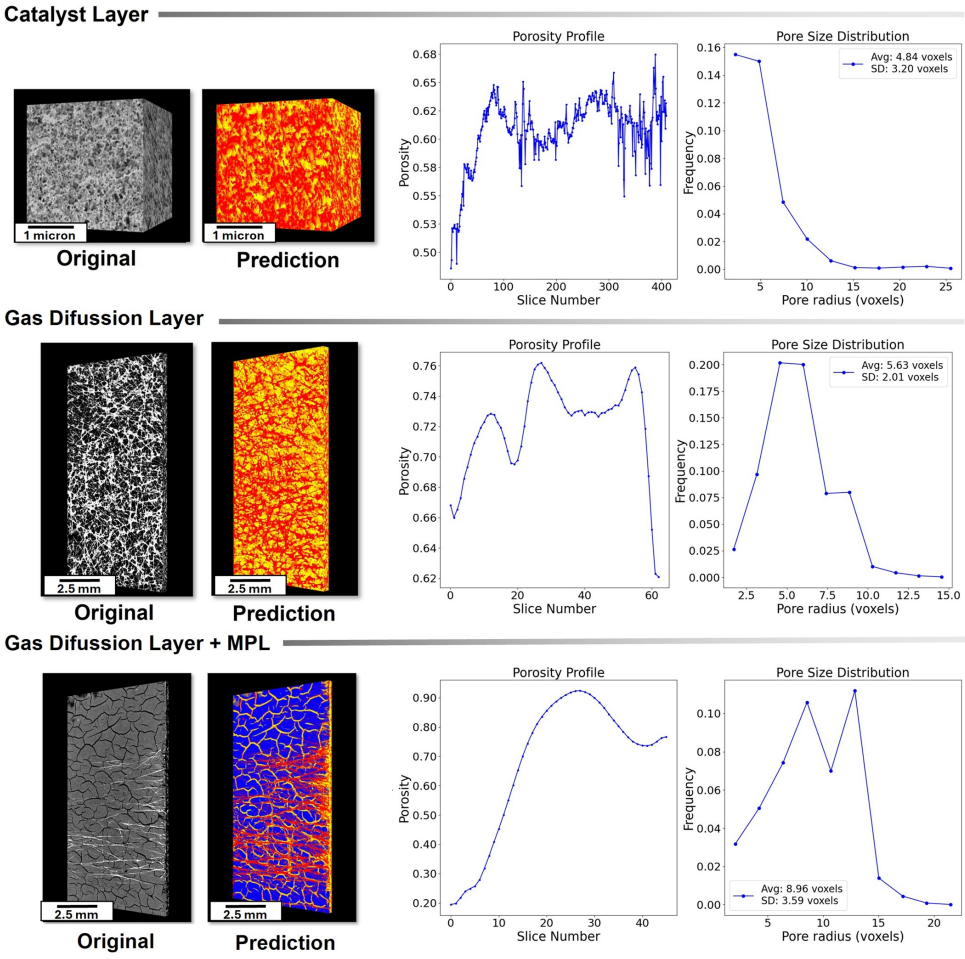

Figure 5.4: Showcase of the property extraction capabilities of the presented tool on CL, GDL, and GDL with MPL segmented volumes. For the segmented tomographies, the tool allows users to extract pore size distribution analysis, porosity, permeability, and tortuosity simulations.

profiles and pore size distributions align well with what has been previously reported in the literature. Ince *et al.* utilized X-ray tomography and revealed GDL porosity profiles above 70% and further revealed local porosity valleys, often towards the centre of the substrate, revealing binder and PTFE agglomerations.[234] Furthermore, our predicted pore scale distribution, mostly encompassing pore sizes between 25–30 µm, are typical of commercial carbon paper-based GDL.[235]

In the combined GDL/MPL volume, the slice-by-slice porosity profile reveals three distinct regimes. In the first 10 slices (pure MPL), porosity remains very low (20–25%), reflecting

the dense crack-limited architecture of the MPL. Between slices 10–20, the MPL/GDL interface porosity rises sharply from 25% up to 90% as MPL cracks penetrate into the substrate. Beyond slice 20 (bulk GDL) the porosity then settles into a plateau around 80–85%, matching the values seen in the GDL-only case. The pore size distribution reveals a broader, bimodal-like pattern with a mean radius of 8.96 voxels (SD: 3.59 voxels). The right-skewed distribution reflects the superposition of two characteristic length scales: larger pores associated with MPL cracks and finer channels from the underlying GDL fiber matrix. Together, these features reproduce the multi-scale pore network expected in an MPL-coated GDL, where fine cracks in the MPL facilitate through-plane transport and the underlying GDL provides the coarser diffusion pathways. Such a heterogeneous, bimodal pore architecture has been correlated with enhanced gas-diffusion performance in MPL-modified GDLs in prior experimental studies.[236,237]

5.3.3 Multi-class feature extraction for MPL/GDL intrusion quantification and crack analysis

For the multi-class segmentation model that includes the MPL, we concentrated on characterizing the added layer and its interdependency within the GDL substrate. Additionally, we isolated the MPL region from the segmented volume and performed a directed crack analysis of its surface. We began by extracting the first layer on the MPL side, then employed computer vision algorithms to identify and measure the isolated cracks. This process yielded the total crack area, the distribution of individual crack areas, and a visual representation of these cracks, as illustrated in Figure 5.5. In the presented example, the distribution is plotted on a logarithmic scale to showcase the crack size distribution visually more comprehensively. From the diagram, we can confirm the visual assessment of a small number of large, interconnected cracks and a higher number of smaller cracks through the MPL. The total crack count can be computed at 483, and the cracks represent 25% of the total surface of the MPL. Understanding such cracking behaviour can inform strategies for optimizing layer design and improving cell performance. For instance, MPLs with cracks have been shown to influence liquid water transport under high humidity conditions, helping to direct excessive water away from the CL.[238,239] However, despite the minor mass transport improvements, MPL cracks have more recently been linked with platinum migration and raise concerns about long-term durability.[240]

We also quantified the extent to which the MPL intrudes into the GDL substrate. First, we measured the average thickness of both layers to contextualize their dimensional scales. To visualize MPL thickness variation across the volume, we generated density maps of the MPL layer (Figure 5.4), along with a 3D rendering, revealing how the MPL distribution

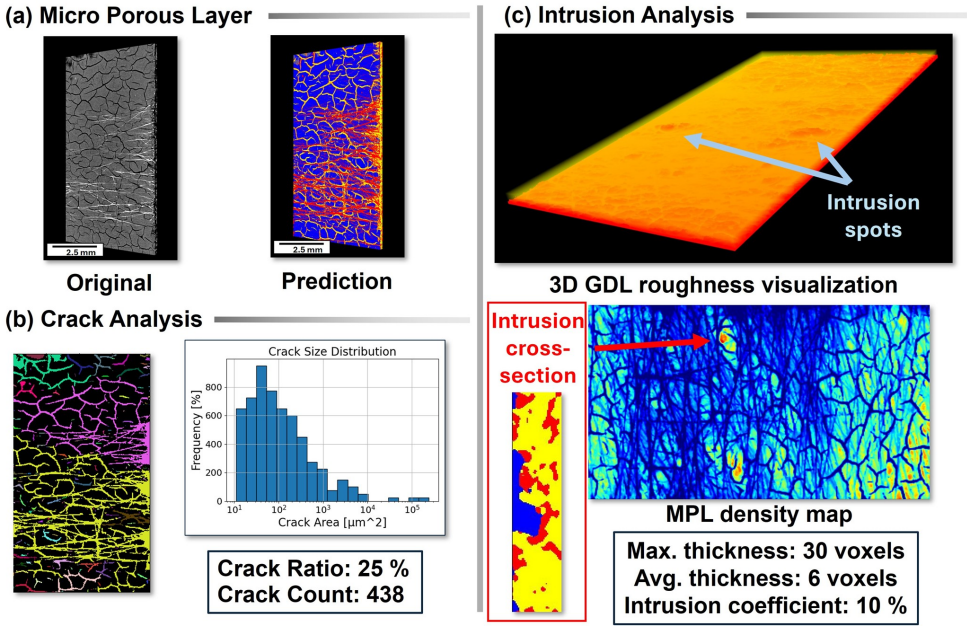

Figure 5.5: (a) Visual comparison of the original GDL with MPL volume and the obtained multi-class segmentation. (b) Example of crack analysis and visualization capabilities of the software, showcasing the segmented MPL. (c) Visualizations of the intrusion analysis of the MPL into the GDL substrate. To support the intrusion analysis, the software presents a 2-dimensional visualization of the intruded MPL as a density map and as a cross-sectional view. It computes a 3D interactive visualization for further inspection of the regions of interest. The intrusion is quantified based on local roughness analysis and its deviation from the normal distribution. Additionally, our software computes the maximum and average thickness of the MPL.

changes spatially. For further understanding of the local high-intrusion zones, a function to represent the cross-sectional view of the MPL/GDL interface was included. To go beyond visualization, we then developed a method to quantify the degree of MPL intrusion into the GDL by focusing on the surface roughness of the MPL side in contact with the GDL. Specifically, we calculated the global roughness of this interface, characterized by the standard deviation of height measurements. We expressed it as a coefficient of variation, defined as the ratio of the standard deviation to the mean roughness value. This quantitatively measures how deeply the MPL intrudes into the GDL.

In this example, we observe in the density map two highly intruded regions toward the middle of the MPL. In the 3D reconstruction, it is possible to observe both intrusions and a certain homogeneity in the rest of the MPL with slight increases in the intrusion towards the top and bottom parts. The maximum MPL thickness of 30 voxels and the average thickness of 6 voxels indicate that the intrusions achieved a high intrusion on the red-

hued areas. Since the areas are relatively small in comparison to the rest of the MPL, we observe a small intrusion coefficient of 10%. Accurate quantification of MPL intrusion is critical when optimizing for the next generation of GDL materials. Specifically, the MPL intrusion dictates the degree of diffusion scale overlap and can be tailored to result in a more continuous in-plane pore-scale distribution. These features are both key parameters for developing PEFC and PEWE sub-component models and achieving desirable in-plane porosity gradients, respectively.[241,242]

By applying deep learning to MPL-coated GDLs, we derive advanced metrics that were previously inaccessible. Beyond the functions described above, we further calculated the interface area between the MPL and the GDL. Specifically, we examined the neighbors of each MPL voxel and counted those that were adjacent to a GDL voxel. This surface area metric provides an additional quantitative measure of MPL intrusion into the GDL.

5.4 Conclusions

UTILE-Pore is an integrated 3D deep-learning pipeline that combines porosity-calibrated ground-truth masks with cutting-edge volumetric convolutional neural networks and transformer models to automate the segmentation of FIB-SEM and micro-CT tomograms. It focuses on CL, GDL, and MPL, replacing the time-consuming manual annotation and ad hoc thresholding by incorporating physical porosity measurements during model training. This approach enables both reliable binary (CL/GDL) and multi-class (MPL, GDL, and pores) segmentation within a single platform.

In seconds, UTILE-Pore generates binary and multi-class segmentations and immediately extracts vital structural and transportation properties, such as pore-size distributions, porosity, tortuosity, specific surface area, and permeability estimates via the Kozeny–Carman equation. For MPL-coated GDLs, it also assesses interlayer characteristics, such as crack density, crack size distributions, MPL thickness, and intrusion depth, using connected-component and roughness analyses. Interactive 3D visualizations and quantitative plots enable rapid correlation between microstructure and performance metrics, while automated export to PoreSpy/OpenPNM formats facilitates subsequent theoretical modeling.

To achieve a comprehensive multiscale characterization, we intend to integrate higher-resolution and phase-contrast imaging modalities (e.g., nano-CT, hierarchical CT, advanced FIB-SEM, and synchrotron tomography) in the future to allow for direct 3D segmentation of sub-micron MPL pores and PTFE binder distributions. We plan to release UTILE-Pore as a modular design publicly, encouraging the community to integrate

extra computer-vision tools, larger neural network architectures, and enhanced visualization capabilities, which will promote widespread usage and ongoing improvements in polymer-electrolyte-membrane devices.

Chapter 6

UTILE-Meta: FAIR Infrastructure for Imaging and Characterization Data Assets with Large Language Model-aided Ontology Standardization in Materials Science

Note: This chapter summarizes the work done as a collaboration between Max Dreger, a PhD student at IET-3 at Forschungszentrum Jülich, and me. Therefore, the personal pronoun "we" is used throughout this chapter to refer to the group of researchers that has been part of this specific study. The work focuses on developing metadata schemas for various characterization methods in materials science. These schemas were then integrated into a graph database using a graph data model and LLMs to allow advanced node organization and querying. The metadata schemas and the first graph database implementation were developed by me in collaboration with the experimental scientists Ingo Manke, Joachim Pasel, Tobias Morawietz, Roswitha Zeis, Andreas Friedrich, and Jasna Jankovic. Mr. Dreger's contribution was in the final integration of these metadata schemas into the graph database, the LLM-guided ontology alignment, and the creation of interfaces for data upload and search functionalities. This work is being submitted for the publication: André Colliard-Granero, Max Dreger, Ingo Manke, Roswitha Zeis, Joachim Pasel, Tobias Morawietz, Andreas Friederich, Jasna Jankovic, Kourosh Malek, Michael Eikerling, and Mohammad J Eslamibidgoli. Fair infrastructure for imaging and characterization data assets with large language model-aided ontology standardization in materials science. *Journal of Materials Informatics*, submitted, 2025. It was further

supported and supervised by Dr. Kourosh Malek, Dr. Mohammad J. Eslamibidgoli, and Prof. Michael Eikerling.

6.1 Introduction

As explained in Section 1.4, the standardized collection and management of data have become an essential pillar across all research fields. They are an urgent necessity that is not widely addressed in the scientific community.[244–246] In energy materials science, the vast amounts of data generated by advanced characterization techniques represent a largely untapped resource.[247–250] Without proper data storage practices and the adoption of data-driven models, the development of new technologies risks stagnation, as valuable insights from past experiments remain inaccessible.[251] A proven strategy to enhance data utility is adopting the FAIR principles (see Section 2.1.2), guidelines designed to maximize data value by focusing on metadata utilization and efficient database management.[77,252] The FAIR principles offer overarching objectives, but implementing them should be tailored to each specific research domain individually.

Imaging data, a core part in materials science, are particularly well-suited for FAIR principles due to their diversity across characterization methods and scales, from nanoscale to macroscale, as shown in Figure 6.1.[253] In materials science, imaging data often involves variances, where even minor differences in sample preparation or synthesis can produce significant changes in image characteristics.[254] Without a standardized protocol for storing such data, challenges arise in accurately identifying samples and retrieving critical synthesis and characterization details, which hinders reproducibility and complicates the management of large datasets.[77,255] Consequently, this lack of standardization leads to increased time, costs, and human resources, as experiments may need to be repeated unnecessarily.[256] Standardized metadata protocols also facilitate collaboration, enabling shared data structures and descriptors across institutions and companies.[165] Furthermore, a major advantage of standardized, machine-readable data storage is that it enables high-throughput, algorithmic access. This approach allows for the rapid retrieval of well-formatted data, suitable for training machine learning models to predict material properties, optimize fabrication workflows, and discover correlations between diverse data assets.[102,257]

The scientific community has shown increasing interest in this area, with various groups proposing potential solutions. For example, the Materials Data Facility is a database that provides user-friendly support for data publication and discovery, promoting open data sharing, self-service data publication, and encouraging data reuse through powerful data discovery tools.[258] Additionally, Romanos *et al.* made significant advancements in

metadata standardization in materials science with their CHADA metadata tables and associated workflow.[259] Using the EMMO ontology, they introduced a generalized protocol for experimental metadata collection, encompassing sample descriptions, method details, raw data, and post-processing steps.[100] Other initiatives, such as Nexus and NOMAD, also contribute to the ongoing effort toward metadata standardization.[147]

While these approaches align closely with FAIR principles, some limitations remain. For instance, CHADA metadata tables are not machine-readable and are stored as plain text, complicating data categorization and automated data extraction. To date, no fully integrated public database architecture employs its ontology and data structure.

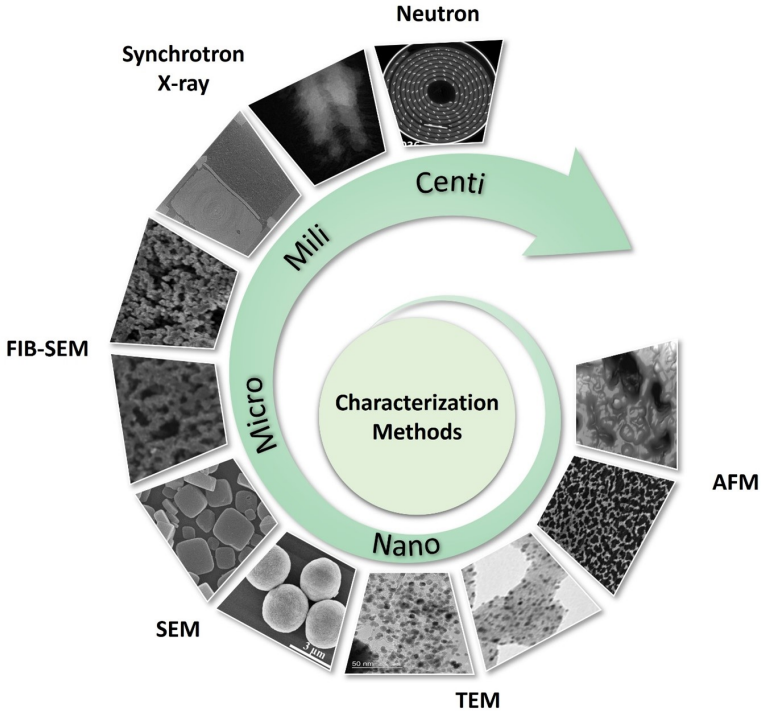

Figure 6.1: Depiction of different materials characterization methods scaling from nano to macro scale.[59,62,260,261]

This work aims to advance these initiatives by implementing a standardized, machine-readable metadata protocol with fixed entries but a dynamic format, accommodating various characterization techniques. Additionally, we employed a graph database implementation using Cypher[262] and Neo4J,[263] which automatically generates a graph-like data structure from the metadata schemas. To promote terminological standardization and interoperability across domains, we employ a large language model (LLM)-assisted

categorization of nodes, as demonstrated by Dreger *et al.*,[264] to align entries with the EMMO ontology. To illustrate the impact of this approach, we present a pilot app to ingest and query data into the database and illustrate its capabilities for a case study on nanoparticle analysis in transmission electron microscopy (TEM) images, highlighting the advantages of the proposed tools.

6.2 Metadata schema development for scientific imaging data management using graph databases

The proposed workflow for utilizing the standardized metadata protocols and subsequent integration into a graph database for data visualization, retrieval, and analysis encompasses several steps is depicted in Figure 6.2. The workflow is based on data collected from experimental partners. The developed metadata protocol is flexible enough to cover various characterization methods, from electron microscopy to synchrotron and neutron tomography. It offers storage of all data related to specific samples from fabrication to the analyzed characterization data, covering the whole sample life cycle.

Once the data have been collected, the metadata protocols and data are uploaded into the graph database. Inside the database, the collected metadata are transformed into a graph. The nodes are associated with ontology-based label nodes to preserve the standardization of the terminology, even if different words or characters are utilized during data collection. Finally, once populated, the graph database fulfills the FAIR principles, facilitating high granularity in data retrieval for AI model training, workflow optimization, and finding correlations inside the database.

For the creation of this workflow, we implemented an architecture that comprises the following parts:

- Metadata schemata, in which are agnostic characterization-techniques to capture experimental details and imaging data.
- A Neo4j graph database to store data and metadata, and enhance data query and retrieval capabilities.
- Semantic contextualization of data via LLM-guided ontology alignment to ensure the standardization of metadata regardless of the data inputs.

The following sections will explain each part of the pipeline in detail and illustrate possible applications with an exemplary use case.

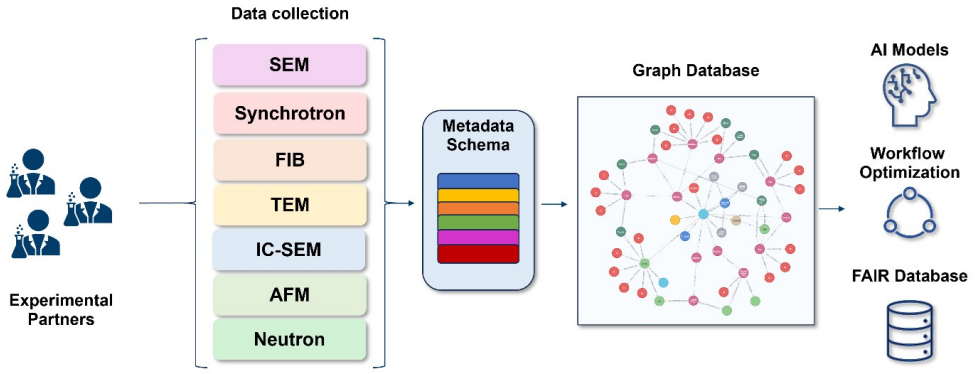

Figure 6.2: General workflow depicting the metadata schema creation, its implementation in a native graph database, and its subsequent employment for training AI models, experimental workflow optimization, and FAIR data storage.

6.2.1 Expert-guided metadata schema conceptualization

The main challenge in developing metadata schemata and standards is the lack of information and specific guidelines or tutorials on developing them in a high-quality manner and by other standards. Our investigation found previous works that set the ground for further development of FAIR metadata schemas, in accordance with AI-ready and machine-readable protocols. One of them was the CHADA tables, an innovative initiative introduced by Romanos *et al.*[259] to standardize the collection of metadata in materials characterization procedures. Their approach involves constructing a metadata framework capable of describing the process of any general characterization method. They proposed gathering comprehensive information on the use case, method, raw data, and post-processing steps to achieve this. This information enables the creation of flexible characterization workflows and allows for the storage of as much data as necessary. The primary method they suggest for collecting these data is by completing a structured Excel file, where the aspects of the experiments and processes are recorded in plain text.

While this approach presents a valuable concept, it currently lacks a practical solution for ensuring standardization and machine-readability of stored data. In our work, we aim to build upon the concept developed by this group and advance it toward a standardized, machine-readable, and FAIR implementation of a metadata protocol, with actual integration into databases for practical application of the proposed ideas.

Initially, our experimental partners were asked to complete the CHADA tables described by Romanos *et al.*[259] for a randomly selected experiment. We collected eight specially adapted CHADA tables, each about different characterization methods and areas within energy materials research. These tables included general textual descriptions of sample

characteristics, method-related parameters, summaries of the raw data obtained, and details of post-processing steps performed. We also added an extra section describing the fabrication/synthesis workflow, as we believe fabrication procedures are essential for understanding the data provenance and facilitating further analysis.

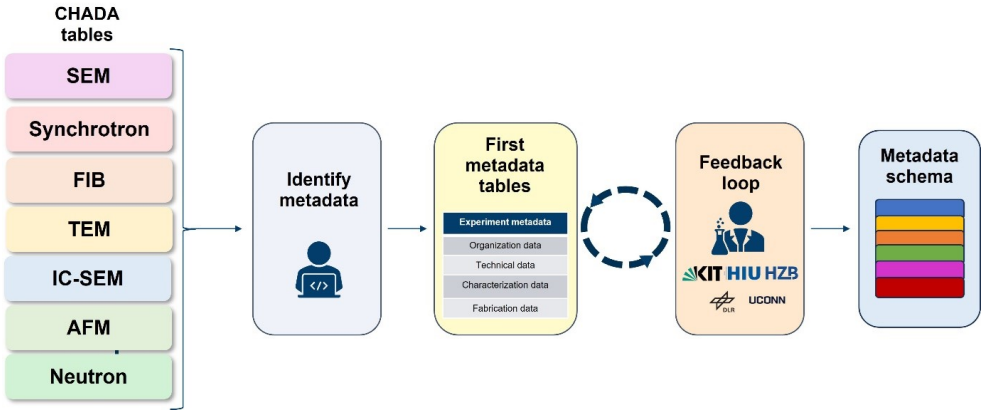

Figure 6.3: Workflow for the conceptualization of the universal metadata schema. Starting from the metadata collection with the CHADA tables filled by the experimentalists, the relevant metadata was extracted and refined into a final metadata schema in an interactive loop with the experts.

Next, as shown in Figure 6.3, we compared the information obtained from different methods to identify structural similarities and determine which entries were common or specific to each method. For each method, we created a preliminary, method-specific metadata table. This was achieved by converting the plain-text CHADA tables into a machine-readable, tabular format, with a complete description of the experiments. The initial metadata table versions were shared with our partners to establish a feedback loop, ensuring the metadata's accuracy and clarity for the experiment descriptions.

After refining the specific metadata tables based on feedback, we identified key parameters and structures that allowed us to simplify and generalize the metadata format. This led to developing a unified metadata protocol capable of flexibly describing each characterization method while adhering to the FAIR principles. The resulting expandable metadata protocol comprises six subgroups of information to describe each experiment fully are defined as:

- **Organizational Data** records who generated or owns each dataset and under what institutional or funding context (e.g., author names, affiliations, DOIs).

- **Synthesis/Fabrication Process** metadata detail the sequence of chemical or physical steps ranging from raw precursors to final material that produced the sample.
- **Sample Preparation Process** metadata describe the post-synthesis modifications (e.g., slicing, polishing, or staining) applied to ready the specimen for characterization.
- **Characterization Data** include all method-specific parameters (e.g., microscope model, accelerating voltage, or scan rate) and instrument settings critical to understanding how measurements were made.
- **Preprocessing Process** metadata document the transformations applied to raw measurement outputs (e.g., denoising filters, background subtraction, or normalization) before analysis.
- **Data Analysis Process** metadata capture the quantitative metrics and algorithms used to extract features (e.g., particle size distributions, phase volume fractions, count statistics), thus linking raw images all the way through to the scientific conclusions.

The different metadata subgroups can be of two types: key-value or process-based. Key-value data is straightforward and stored as static tables following the standard key-value format. We developed a standardized, machine-readable method to describe processes for the process-based subgroups. The diverse ways processes can occur necessitate a flexible yet standardized process description.

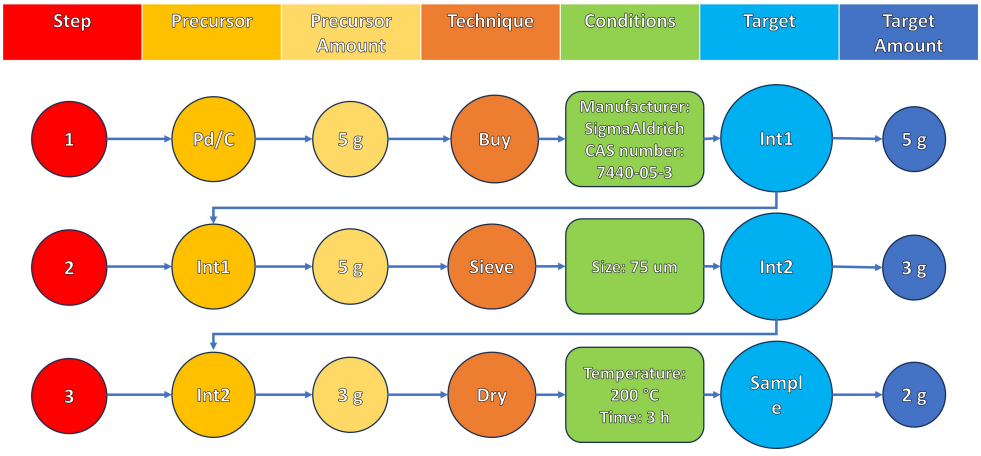

Figure 6.4: Detailed description of the process conceptualization schema created to describe complex and parallel experimental procedures.

As illustrated in Figure 6.4, processes are composed of steps performed chronologically. Each step involves one or more precursors with specified amounts. The precursors undergo modifications defined in the technique node, which may be associated with certain conditions or parameters, providing additional details about the technique. After applying the technique to the precursors under specific conditions, a target is produced in a defined amount. Since processes are often discontinuous and intermediary steps may lack specific names, we introduced numbered intermediates. Intermediates serve as precursors for subsequent steps without needing explicit naming, while preserving all prior step information in the sequence. Moreover, intermediates facilitate parallel steps that can converge in a future step. This approach accommodates processes that do not follow a linear, chronological structure, such as using "Intermediate1" as a precursor in a later step, e.g., Step 3.

Finally, the metadata tables are organized with fixed sections to ensure that necessary information is collected in a standardized manner while allowing the length of entries to be dynamically adjusted according to the complexity of each experiment. Experiments involving longer processes can seamlessly add steps, and tabular nodes can be expanded with additional parameters as required by the experimental requirements.

6.2.2 Autonomous metadata schema mapping into a standardized graph database

The fixed structures of the different metadata schemas allowed us to develop Python classes handling the ingestion of the metadata schemas into the graph database developed and published by Dreger *et al.*[264] These data handlers are the interface layer between the metadata schemas and the database. The scripts translate the completed CSV metadata tables into an appropriate graph database format, facilitating efficient data storage and querying.

The created graph from the tables depicted in Figure 6.5 follows a hierarchical structure, where the node with the organizational data stands in the middle. From this central node, the chronological steps from fabrication to sample preparation, measurement, data preprocessing, and data analysis are sequentially connected to form a circular experiment graph containing all the information and steps collected in the tables.

The representation of experiments as knowledge graphs enriches the data as they explicitly contain relationships that contextualize single data points within an experiment. This contextualization is entailed only implicitly within the tables. Converting the metadata into a graph database is a significant step toward achieving FAIR data practices. However, the content of the entries is not yet strictly limited. Imposing complex restrictions on

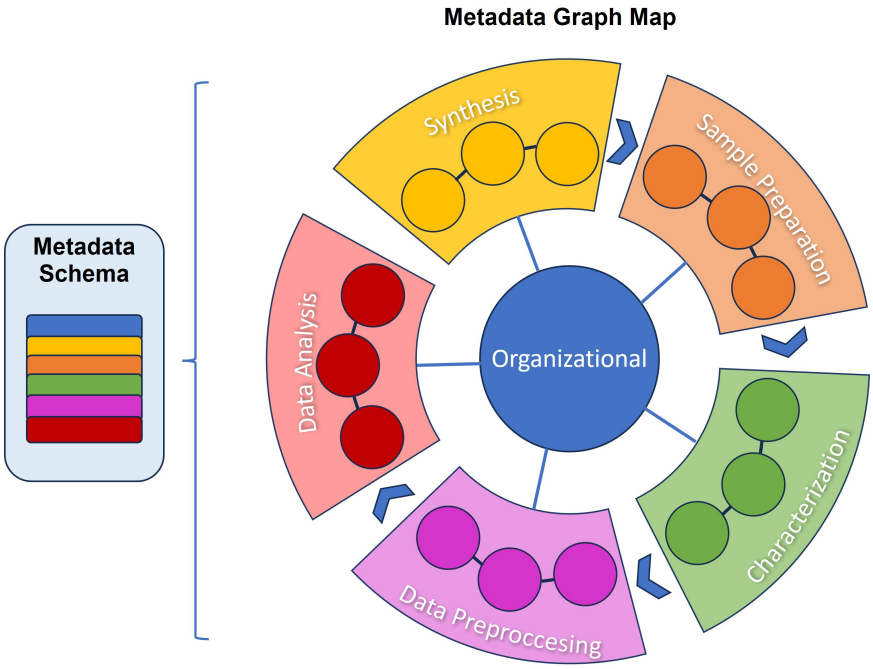

Figure 6.5: Overview of the conceptualized sections of the metadata schemas to cover all the experiment details.

user inputs could result in the approach failing to cover all possible cases or necessitate substantial human effort to update the terminology to encompass all potential entries. Therefore, it is crucial to categorize the data accurately and automatically group different parameters, techniques, and processes. This ensures a truly standardized terminology, regardless of the experimentalists' backgrounds or the use of synonyms.

Figure 6.6 shows the conversion from the metadata schema (left) to a graph representation of a single experiment (middle), and to the representation of multiple experiments (right). The graph containing several experiments shows that the single experiments can be interconnected as they might share the same researchers, conducting the experiments, and the same measurement instruments.

6.2.3 Metadata graph standardization via LLM-assisted ontology alignment

Even if the standardization of the metadata protocols and their subsequent transformation into an interconnected graph represents an essential advancement towards achieving an interoperable and standardized data repository, the individual graphs would remain

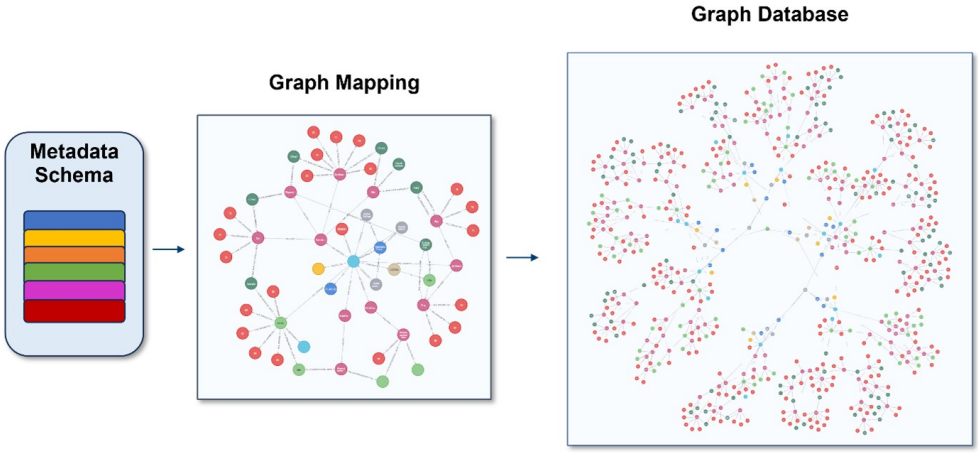

Figure 6.6: Schematic depiction of the conversion of metadata tables into a standardized graph, and its subsequent storage in the graph database along with other experiments.

isolated if no further step is taken. The storage of separated graphs would still lack the finding of similar experiments or related samples if the users stored different entries with synonyms or abbreviations. To tackle this issue, the graphs of the various experiments must be interconnected with the rest of the database by assigning each experimental node to a common standardized ontology structure. This is achieved by employing semantic contextualization of the nodes.

We first grounded our nodes in semantic context by importing the Matter, Quantity, and Process branches of the EMMO ontology (along with its BattInfo extension) into our graph database to achieve the semantic contextualization of the nodes. These ontology branches form the backbone of our label taxonomy, with each label node accompanied by a synonym and brief explanation. Because the combined branches include over a thousand labels, we automated the task of generating alternate names and descriptions using OpenAI's GPT-4-Turbo. We then computed embeddings for each label, based on its official name, synonyms, and description, to capture their whole meaning, and stored these vectors in the database linked to their respective label nodes.

Labeling an extracted node becomes a matter of matching its embedding against the stored label embeddings using cosine similarity. If the best match scores below 0.95, it suggests that no existing label sufficiently covers the node's semantics. In that case, an LLM-driven agent is invoked, equipped with a context-specific prompt detailing the node type, rules, and examples, and it either finds the most appropriate existing label or proposes extensions to the taxonomy. The agent can pursue one of three paths:

- No suitable existing label: It reviews all child labels under the relevant branch (e.g., "matter"), picks the closest parent, and suggests a new label—optionally with additional children—to enrich the taxonomy.
- Exact match found: It assigns the node to the identified label.
- Partial relationship detected: It examines parent or child candidates to determine the nearest semantic fit, then generates a new label under that branch.

Whenever new labels are created, they're integrated into the graph, linked to the original node, and flagged for human curation, ensuring that our ontology remains accurate, consistent, and searchable.

6.2.4 Hybrid lexical-semantic search

Data stored within a database needs to be findable and accessible, which requires a user interface that allows access to the data. This access needs to allow various ways of filtering the data, as simply dumping all the data would not be sufficient. Therefore, this work presents an innovative hybrid approach that combines lexical and semantic search capabilities.

Lexical search is usually employed to access data via identifiers, such as the names or ORCIDs of authors, the DOI of the data, or a publication about the data. Identifiers can also be assigned to materials or chemicals as CAS numbers and internal identifiers for experiments, sample preparations, synthesis, analysis, or data preprocessing. As identifiers are ideally unique, filtering the database with them can be done with a simple string comparison.

With semantic search techniques, we query the database via materials, workflows, parameters, or material properties. A user might want to get all TEM measurements done on Ir/C catalysts. The challenge is that the names of materials, measurements, properties, parameters, and manufacturing steps are not standardized. Ir/C could alternatively be named Iridium on Carbon or IrC. The nodes within the graph database are labelled, but the user does not know the labels; thus, a simple string comparison would not suffice for robust querying. The proposed pipeline in this work enables the user to define a workflow of interest to form a query graph that matches similar workflows within the database. For experiments collected with the proposed metadata schema, the pipeline is simplified since the schema already follows a standardized template. Then, semantic similarity is computed for the proposed query graph and the existing workflows in the database, and the experiments with relevant connections and matching processes are returned to the user.

Combining lexical and semantic search capabilities into a hybrid database access approach enhances data discovery. Hybrid search capabilities mean user query parts containing an identifier are processed with simple string comparisons. Everything that is not standardized, such as names of materials, processing steps, measurements, properties, and processing conditions, is initially processed with string comparisons, and if no match is found, similarity comparisons of embeddings. These similarity comparisons use the embeddings assigned to each ontology node. The filter argument is transformed into its vector representation and then compared to embeddings within the graph database. The most similar candidate is returned as a search result, if the similarity is below a predefined threshold (e.g., 0.95), no result is returned.

6.3 Pilot deployment of the database app and application use case

A pilot application for input and retrieval of experimental imaging workflows was developed to demonstrate the usefulness of the database and its implementation. The tool allows the user to introduce metadata protocols into the application with template storing capabilities to improve the user experience with faster data entry processes. Additionally, the presented application supports the entry and retrieval of metadata protocols in the two most common formats, CSV and JSON, to enhance the interoperability of the extracted metadata with automated console protocols and other external applications.

Furthermore, the user can select one of the two query functionalities explained in this work: lexical and semantic search, enabling the query of unique identifiers or whole experiments. A depiction of the app's user interface is showcased in Figure 6.7.

To further showcase the potential employment of the tool, we present an example describing a series of interactions between the user and the database following the successful data collection using the tool. We illustrate the case of a hypothetical group collecting TEM images of platinum nanoparticles on a carbon support to investigate their effects on polymer electrolyte membrane fuel cells. After each experiment, where synthesis parameters were varied, the images were recorded and stored in the graph database, accompanied by the detailed metadata protocol developed in this work, as shown in Figure 6.8.

Compiling a high-quality dataset with comprehensive metadata protocols enables rapid and straightforward retrieval of all related data. Accessing TEM images of platinum nanoparticles allows us to quickly assemble a large dataset of images across the database, which can then be used to train a DL model for individual particle segmentation. The enhanced data retrieval addresses a significant bottleneck in machine learning due to the

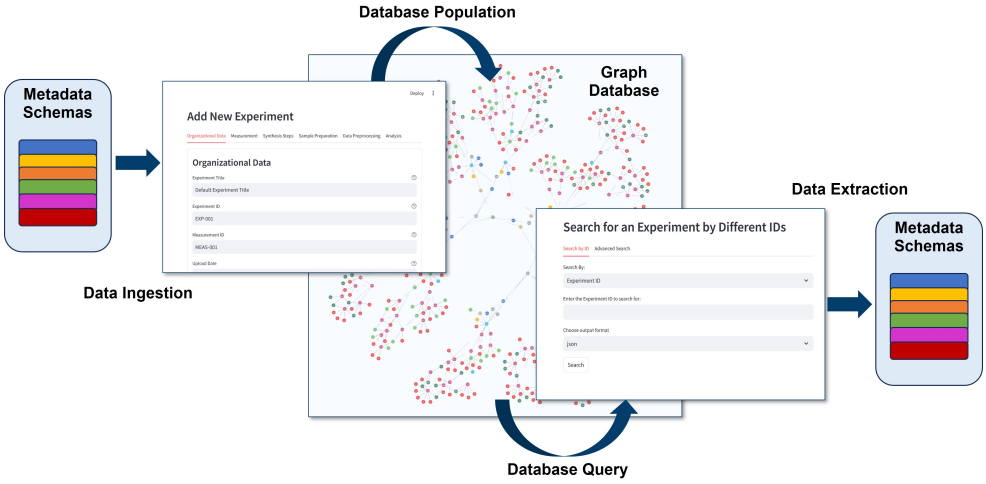

Figure 6.7: Showcase of the graphic user interface for the entry and retrieval of data from the application

lack of high-quality data. Then, the developed model can automatically perform particle size distribution analysis, providing quantitative measurements of particle sizes.

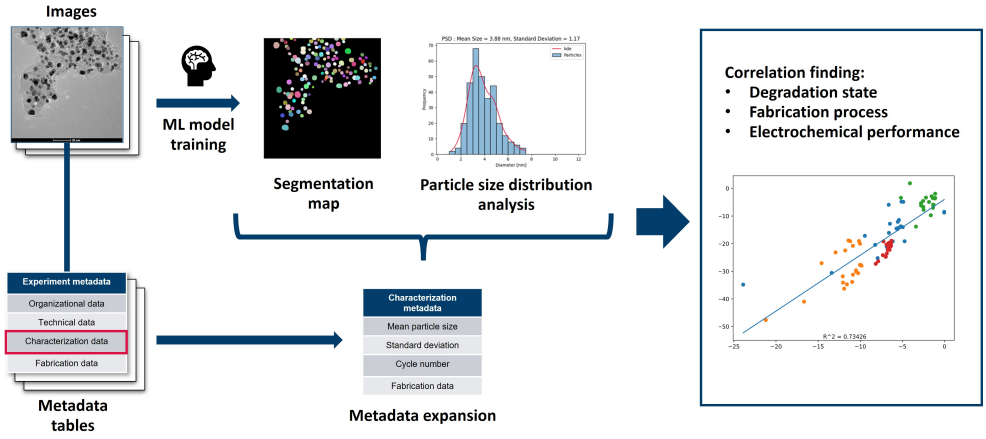

Figure 6.8: Schematic illustration of a potential use-case of the standardized metadata collection for accelerated AI-model training and correlation finding for the case of TEM images of Pt/C nanoparticles.

We proceed with the database approach once the model is successfully trained and deployed. Particle size distribution analyses can be automatically applied to each image of platinum nanoparticles in the database, and the resulting quantifications can be added to the image metadata by expanding the data analysis process. With standardized im-

age metadata, we can now access more insightful discoveries. For instance, if additional procedures and parameters, such as synthesis methods, temperatures, or different brands of ink powders, were recorded, they can be correlated to the particle size distributions. This would allow the user to understand how fabrication conditions affect particle size.

Another example is correlating the electrochemical performance of fuel cells fabricated with the characterized platinum nanoparticles by comparing performance data with the extracted size distributions. By utilizing time-resolved TEM images of catalyst layers, we can similarly perform degradation analysis based on size differences observed at different time points, facilitated by the tool. Moreover, outliers can be easily analysed by examining the detailed fabrication and measurement protocols to identify errors or explain anomalous behaviours.

6.4 Conclusions

This research represents a significant advancement in the standardization and digitalization of metadata into image databases for materials science. The proposed tool focuses on two main aspects: firstly, the creation of a standard metadata protocol for understanding data points; and secondly, the efficient, ontology-based storage of data using LLM-oriented ontology alignment and graph databases developed with the support of Max Dreger. This work bridges the gap between FAIR guidelines and their practical implementation, reducing protocol redundancy and facilitating scientific community adoption.

To achieve this goal, we developed a metadata schema by consolidating the expertise of eight different groups, each specializing in distinct characterization methods within the energy materials field. Together, we established a FAIR metadata collection protocol. The protocol encompasses all relevant aspects of the stored images, from organizational data that clarifies the topic and origin of the data to the data analysis processes the samples underwent to extract quantitative information from the imaging data.

Then, the developed metadata schemas were mapped into a graph database and aligned to a standard ontology based on EMMO and BattInfo via LLM-guided node categorization. This approach enabled the dynamic standardization of user inputs and the evolution of the employed taxonomy when users introduce newer concepts. This principle ensures the flexible automatic implementation of newer concepts in a standardized fashion.

Another significant concept in the development of databases is how easy it is to query and retrieve data. For this, we implemented in this work a hybrid approach combining string comparison capabilities and the advanced capability of querying whole workflows

for materials and experiments based on semantic search. This development is critical for finding and accessing complex datasets within the database.

For practical demonstrations, a pilot application with a graphical user interface was developed. Users can upload their experimental metadata protocols and query and search in the database with the hybrid search capabilities. The resulting data retrieval can be additionally downloaded in two of the most standardized metadata formats: CSV and JSON, ensuring seamless integration into existing workflows or external tools.

Furthermore, we demonstrate the tool's advantages through a hypothetical use-case scenario involving TEM images of platinum nanoparticles on a carbon support for PEMFCs. We demonstrate this by utilizing an advanced, AI-ready metadata schema, enabling models to be trained more rapidly due to faster dataset collection from high-throughput data retrieval capabilities. Additionally, data extracted from AI models or manual methods can be directly stored in the database, enabling rapid correlation and error detection between datasets.

One of the most significant bottlenecks in accurate metadata collection is the tedious manual data entry into correct formats and tables. Therefore, this work serves as a first step towards a user-friendly approach to bring metadata collection closer to experimentalists. Looking ahead, we aim to implement this schema into user-friendly tools that experimentalists can easily use. The substantial overlap between the data collected in the metadata tables and the entries of a standard laboratory journal is intentional, and the ability to digitize, standardize, and make this data easily accessible facilitates further data analysis. Template formats can be created to further reduce the effort required for formatting data into the database to avoid re-entering information for similar experiments. An innovative approach would be to use LLMs to automatically translate experiments documented in scientific publications, books, or handwritten laboratory journals into the metadata schema proposed in this work.

Chapter 7

Conclusions & Outlook

7.1 Conclusions

In conclusion, the research presented in this thesis represents a significant step forward in advancing automated image analysis in the field of energy materials. By providing high-performance deep learning (DL)-based image analysis tools to accelerate their characterization workflows, and supported by solid metrics to assess the quality of their analyses. This work has involved the development of three pieces of image analysis software to automate three different high-dimensional image analyses. Case scenarios involving autonomous video and 3D tomography analysis with 2D and 3D DL architectures were automated. At the end, together with experimental collaborators, a novel standardized metadata protocol was devised to store imaging data related to materials science according to FAIR principles.

In Chapter 3, an automated image analysis approach tailored to the study of bubble dynamics in PEMWEs was presented. For this purpose, an autonomous analysis software was developed to transform optical video recordings of transparent cells into comprehensive quantitative insights of the bubble system. The developed tool analyzes the 2D images via DL and performs a semantic segmentation of each bubble in the images. By stacking the results, a time-resolved analysis of the datasets can be extracted using computer vision algorithms. Detailed visualizations of bubble size distributions, bubble density maps, and morphological analysis of individual bubbles are of particular interest. These analyses provide valuable insights into how bubble formation and evolution vary across different operating voltages, thereby enhancing the understanding of the underlying physical phenomena affecting cell performance that cannot be achieved by standard or manual analysis due to the complexity and extensive amount of data per dataset.

Similarly, this work has established a generic framework for studying bubble dynamics in VRFBs, as described in Chapter 4. The software is capable of automatically analyzing 3D synchrotron X-ray tomographic images. The developed approach performs accurate DL-based multi-class semantic segmentation of complex datasets into bubble, electrolyte, membrane, and gasket classes. It further automates the analysis of bubble volume distributions, shapes, and spatial arrangements, enabling rapid identification of performance-related features related to membrane blockage and electrode efficiency. To highlight the utility of the developed tool, a comparative analysis of different electrodes modified with different catalysts demonstrated the ability of the tool to detect differences in bubble dynamics and reveal the catalytic activity of each catalyst in bubble formation following hydrogen evolution reactions. These findings directly inform electrode optimization strategies to mitigate these parasitic reactions, potentially paving the path for improving overall battery performance and extending cycle life.

Another key achievement in this thesis is developing an advanced deep learning framework tailored explicitly to analyzing porous materials relevant to PEM-based technologies, as described in Chapter 5. This work focused on training three-dimensional convolutional neural networks and visual transformers to accurately segment complex tomographic datasets obtained from micro-CT and FIB-SEM methods. The tool robustly segments key structural components of PEM technologies, such as CL, GDLs, and MPLs, in a binary and multi-class manner. In addition to the segmentation capabilities, the developed software implements state-of-the-art computer vision algorithms to quantify essential morphological parameters such as porosity, pore size distribution, tortuosity, crack analysis, and the interplay between GDL and MPL. The high performance of the approach was achieved by the special use of porosity-calibrated thresholding to train the models, a technique that aligns segmentation parameters with physical porosity measurements, significantly improving the reliability of annotated datasets used to train neural networks.

Furthermore, the AI tools developed in this work were equipped with sophisticated visualization and interactive 3D rendering for rapid visual analysis of structural data, allowing researchers to correlate morphological features with performance data quickly. In addition, automated extraction of key parameters supports theoretical modeling by providing accurate input data that was previously difficult or impossible to obtain manually. By publicly releasing this comprehensive toolset under an open-source framework, it is expected to stimulate broader scientific engagement and collaboration, inviting the integration of additional CV methods, advanced and further refined DL models, and tailored visualization tools to meet evolving research needs.

Chapter 6 demonstrated the urgent necessity to enhance data collection and storage in the energy materials field, aiming to accelerate image analysis via AI. Therefore, this work dealt with the standardization and digitization of image data storage in materials science, focusing on energy-related applications. In this part of the doctoral research, the creation of a standardized metadata protocol was being developed, alongside an efficient ontology-based storage system based on LLM-driven ontology alignment within a graph database structure. This dual approach effectively bridges the gap between the generic FAIR guidelines and a practical implementation for the scientific community. By significantly reducing redundancy in existing data protocols, the presented methodology lowers the barrier for broader adoption by researchers, as it improves the accessibility, understanding, and reusability of complex experimental imaging datasets.

The expertise of eight research groups, each specializing in different imaging techniques relevant to energy materials, was brought together to create a standardized metadata schema covering all relevant information from sample fabrication to data processing. This collaborative effort yielded a universal metadata collection protocol that is able to cover all essential aspects of imaging data. The range of metadata collected spans from organizational identifiers and provenance information to detailed descriptions of data acquisition, instrument parameterizations, and post-processing methods. The effectiveness of this metadata approach was demonstrated in several use case scenarios central to energy materials science. One prominent example was presented to better understand the potential for broader use of an organized FAIRified graph database. The use case scenario involved the TEM analysis of platinum nanoparticles on a carbon support, a common catalyst for PEMFC technologies. Here, the metadata protocol provided accelerated and standardized dataset curation, enabling faster training of AI models and subsequent data correlation and troubleshooting within the dataset. Such streamlined data handling and storage demonstrate the direct benefits of using this FAIR metadata management framework to accelerate scientific discovery and data-driven innovation.

In conclusion, this work represents a significant step forward in facilitating systematic, standardized, and automated analysis in materials science, particularly for energy technologies. In addition to streamlining metadata collection, standardizing data storage, and automating complex image analysis, robust tools have been provided to bridge the gap between experimental data acquisition and advanced computational analysis employing state-of-the-art deep learning-based image analysis capabilities. Therefore, this work accelerates experimental cycles, fosters collaboration, and improves reproducibility across the materials research community.

7.2 Outlook

Through this work, several milestones have been reached in training DL models and developing CV capabilities to accelerate the analysis of complex imaging data. Additionally, with the knowledge gained, it was possible to standardize and automate the storage of energy materials by bridging the gaps and developing a functional common language for image data. Nevertheless, these achievements have opened up new opportunities to advance further and accelerate the scientific workflows of researchers based on AI protocols that were not feasible before. This section analyzes and proposes different ways to continue the journey started in this work.

Several automated image analysis tools were developed in the context of the UTILE project. Still, a few other use cases related to the project collaborators were not addressed due to time constraints. An immediate follow-on path for the development of dedicated image analysis tools for energy materials imaging challenges is therefore possible. For example, the collaborator, Dr. Ingo Manke from Helmholtz Zentrum Berlin, proposed analyzing the degradation mechanisms of battery materials based on *operando* synchrotron X-ray and neutron tomography volumes. Automating the 4D analysis of the battery materials during battery usage and extracting quantities that correlate with the degradation state of the cell would provide valuable insights for future improvements into processes and parameters that influence the battery's lifetime. Another interesting use case scenario proposed by Dr. Mack from Forschungszentrum Jülich combines DL-based analysis of CL microstructure damage with critical mechanical properties to predict spallation life under cyclic thermo-mechanical loading. Automating the analysis via DL and the correlation finding between the mechanical properties could offer unprecedented insights into the real dependencies of the CL structure on the spallation life, not easily accessible with manual approaches.

Therefore, it is possible to continue solving individual use case scenarios that apply to a particular subset of researchers in the field, or it is possible to go deeper and understand the underlying problem and try to contribute from a fundamental perspective by developing a more general approach. For the latter purpose, two feasible ways exist to expand the collection of tools developed in this thesis further. The main bottleneck of data-driven software is in its very name, its data-driven nature. Deviations in the unseen dataset from the data used for training can lead to a reduction or even total failure of the model's predictive capabilities. Therefore, constant retraining and fine-tuning are required to adapt the tool to newer use cases.

Furthermore, retraining or fine-tuning models on new data requires computational resources and medium to high domain expertise to generate the newer datasets and train

the model. Therefore, a possible perspective is the development of a user-friendly layer for experimentalists to facilitate the improvement of specific models, supported with a user interface and cloud-provided resources for training. This would pave the way for experimentalists to further implement AI-driven tools in their daily routines with a reduced initial effort.

A new trend in AI is shifting the community's interest towards creating domain-specific foundation models. These large models are pretrained on large amounts of data and can be applied to various downstream tasks. Foundation models are the equivalent of the large vision models (LVM), but are used to solve field-specific functions that are not considered in general imaging of large datasets. Usually, LVMs are pretrained on a large number of images in a self-supervised fashion, which allows the model to independently learn to recognize features from pictures without the need for labelling data. After pretraining, these models can be further upgraded on a specific task by providing a labelled dataset later, tailoring the model segmentation capabilities toward the field of interest.

However, these sophisticated approaches face the problems of data scarcity in science and the added difficulty of abstract quantification from a highly heterogeneous niche of imaging data. Approaches such as the segment anything model (SAM) and the segment everything everywhere model (SEEM) are promising starting points that can reduce the overall effort required to collect training images. The development of functional foundational models would reduce the need to develop complex models for each specific use case scenario individually.[265,266] Therefore, an interesting next step based on the results of this work would be the development of a materials science-tailored foundational model. This can be achieved by collecting large amounts of scientific materials data from all possible sources and fine-tuning an available LVM, such as SAM, to increase its performance on the type of images of interest. Nevertheless, foundational models require a considerable collective effort to collect large annotated datasets in various use cases. The tools presented can help solve this issue. Implementing the presented standardized image data repository for AI-ready data collection and the developed case-specific data annotation models will help to rapidly create annotated datasets for training, accelerating the creation of a new generation of large image models for scientific applications.

A further extension of the work presented in this dissertation regarding metadata and standardization protocols is also feasible. For this, it is crucial to identify the weak points in the current approaches to image data storage and organization. One of the major bottlenecks in data management is the often manual effort required to enter the necessary information. In future versions of this database, incorporating additional user-friendly abstractions is envisioned to simplify and automate metadata collection processes, which

will be crucial for a broader adoption of the tool in the community. Therefore, facilitating the usage of the developed metadata schema in experimentalists' workflows is essential. This can be achieved by creating intuitive software interfaces that integrate seamlessly into laboratory routines in a user-friendly fashion. Given the significant overlap between the metadata schema and standard lab journals, this schema naturally complements existing lab management practices, allowing researchers to digitally standardize experimental efforts from the outset.

Another interesting approach would be incorporating template formats to streamline repetitive entries common to similar experimental setups, reducing manual effort. A longer-term possibility is using LLMs to automatically translate unstructured data into structured and standardized metadata tables, such as publications, lab notebooks, or handwritten entries. In addition, a new trend in LLMs called "agents" or "agentic AI" is emerging as a powerful assistant type of LLM system that can not only answer questions, but also plan the solution of complex tasks and execute steps to achieve a goal. Implementing such agents can highly automate data collection and entry into databases with high accuracy, drastically reducing the human effort required. Achieving this automation would significantly reduce the barriers to widespread adoption of standardized metadata protocols in the scientific community and increase reproducibility and transparency in scientific research.

The diversity of approaches developed during this dissertation for multiscale and multidimensional automated image analysis to accelerate the characterization and investigation of energy materials, combined with the standardized metadata protocols, ontology-guided and LLM-powered database for data collection, furnishes a powerful foundation for bridging the gap between science and AI. This work brings the scientific community one step closer to an actual data-driven investigation environment for the development of green energy technologies. By continuing to work towards the proposed next steps, the tools presented can be further developed to form the final UTILE goal of a user-friendly platform for data storage, model training, and data visualization in one place.

Bibliography

[1] Daniel J Soeder. Fossil fuels and climate change. *Fracking and the Environment: A scientific assessment of the environmental risks from hydraulic fracturing and fossil fuels*, pages 155–185, 2021.

[2] International Energy Agency. Global energy review 2025. Technical report, International Energy Agency, Paris, France, 2025.

[3] Phebe Asantewaa Owusu and Samuel Asumadu-Sarkodie. A review of renewable energy sources, sustainability issues and climate change mitigation. *Cogent Engineering*, 3(1):1167990, 2016.

[4] Keshani Attanayake, Isuru Wickramage, Udul Samarasinghe, Yasangi Ranmini, Sandali Ehalapitiya, Ruwan Jayathilaka, and Shanta Yapa. Renewable energy as a solution to climate change: Insights from a comprehensive study across nations. *Plos one*, 19(6):e0299807, 2024.

[5] Robert Gross, Philip Heptonstall, Matthew Leach, Dennis Anderson, Tim Green, and Jim Skea. Renewables and the grid: understanding intermittency. *Proceedings of the Institution of Civil Engineers-Energy*, 160(1):31–41, 2007.

[6] Zhenguo Yang, Jianlu Zhang, Michael CW Kintner-Meyer, Xiaochuan Lu, Daiwon Choi, John P Lemmon, and Jun Liu. Electrochemical energy storage for green grid. *Chemical reviews*, 111(5):3577–3613, 2011.

[7] Nicola Jones. The electric-car battery revolution. *Nature*, 626:8, 2024.

[8] Thomas Cherico Wanger. The lithium future—resources, recycling, and the environment. *Conservation Letters*, 4(3):202–206, 2011.

[9] Peter Greim, Asfaw Solomon, and Christian Breyer. Assessment of lithium criticality in the global energy transition and addressing policy gaps in transportation. *Nature communications*, 11(1):4570, 2020.

[10] International Energy Agency. World energy outlook 2021. Technical report, International Energy Agency, Paris, France, 2021.

[11] U.S. Department of Energy. Energy storage systems program, 2024.

[12] U.S. Geological Survey, National Minerals Information Center. Lithium statistics and information. Technical report, U.S. Geological Survey, 2021.

[13] U.S. Geological Survey. How much water is there on earth?, 2019.

[14] Yun Wang, Yiheng Pang, Hui Xu, Andrew Martinez, and Ken S Chen. Pem fuel cell and electrolysis cell technologies and hydrogen infrastructure development–a review. *Energy & Environmental Science*, 15(6):2288–2328, 2022.

[15] Yujing Guo, Gendi Li, Junbo Zhou, and Yong Liu. Comparison between hydrogen production by alkaline water electrolysis and hydrogen production by pem electrolysis. In *IOP Conference Series: Earth and Environmental Science*, volume 371, page 42022, 2019.

[16] Frano Barbir. *PEM Fuel Cells*, pages 27–51. Springer London, 2006.

[17] U.S. Department of Energy. Automotive fuel cell targets and status. Technical Report 1, Hydrogen and Fuel Cell Technologies Office, 2020.

[18] He Liu, Jiang Qin, Chengjie Li, Chenghao Li, and Peng Dong. Performance comparison and potential evaluation of energy systems with different fuel cells for electric aircraft. *Applied Thermal Engineering*, 242:122447, 2024.

[19] Chasen Tongsh, Siyuan Wu, Kui Jiao, Wenming Huo, Qing Du, Jae Wan Park, Jin Xuan, Huizhi Wang, Nigel P. Brandon, and Michael D. Guiver. Fuel cell stack redesign and component integration radically increase power density. *Joule*, 8(1):175–192, 2024.

[20] Jung-Ho Wee. Applications of proton exchange membrane fuel cell systems. *Renewable and sustainable energy reviews*, 11(8):1720–1738, 2007.

[21] Maximilian Bernt, Alexandra Hartig-Weiß, Mohammad Fathi Tovini, Hany A El-Sayed, Carina Schramm, Jonas Schröter, Christian Gebauer, and Hubert A Gasteiger. Current challenges in catalyst development for pem water electrolyzers. *Chemie Ingenieur Technik*, 92(1-2):31–39, 2020.

[22] Muhammad Tawalbeh, Suma Alarab, Amani Al-Othman, and Rana Muhammad Nauman Javed. The operating parameters, structural composition, and fuel sustainability aspects of pem fuel cells: A mini review. *Fuels*, 3(3):449–474, 2022.

[23] Angel P Manso, Florencio F Marzo, Jose Barranco, Xabier Garikano, and Garmendia Mujika. Influence of geometric parameters of the flow fields on the performance of a pem fuel cell. a review. *International journal of hydrogen energy*, 37(20):15256–15287, 2012.

[24] Stephen J Paddison and Hubert A Gasteiger. Pem fuel cells, materials and design development challenges. In *Fuel Cells: Selected Entries from the Encyclopedia of Sustainability Science and Technology*, pages 341–367. Springer, 2012.

[25] Huamin Zhang. Redox flow battery for energy storage. *ECS Transactions*, 28(22):1, 2010.

[26] Massimo Guarnieri, Paolo Mattavelli, Giovanni Petrone, and Giovanni Spagnuolo. Vanadium redox flow batteries: Potentials and challenges of an emerging storage technology. *IEEE Industrial Electronics Magazine*, 10(4):20–31, 2016.

[27] Thomas Puleston, Andreu Cecilia, Ramon Costa-Castelló, and Maria Serra. Vanadium redox flow batteries real-time state of charge and state of health estimation under electrolyte imbalance condition. *Journal of Energy Storage*, 68:107666, 2023.

[28] Kyle Lourenssen, James Williams, Faraz Ahmadpour, Ryan Clemmer, and Syeda Tasnim. Vanadium redox flow batteries: A comprehensive review. *Journal of Energy Storage*, 25:100844, 2019.

[29] László Eifert, Zenonas Jusys, Ray Banerjee, Juergen Behm, and Roswitha Zeis. Differential electrochemical mass spectrometry of carbon felt electrodes for vanadium redox flow batteries. *ACS Applied Energy Materials*, 1(12):6714–6718, 2018.

[30] Yue Liu, Tianlu Zhao, Wangwei Ju, and Siqi Shi. Materials discovery and design using machine learning. *Journal of Materiomics*, 3(3):159–177, 2017. High-throughput Experimental and Modeling Research toward Advanced Batteries.

[31] Khurram Shahzad, Andrei Ionut Mardare, and Achim Walter Hassel. Accelerating materials discovery: combinatorial synthesis, high-throughput characterization, and computational advances. *Science and Technology of Advanced Materials: Methods*, 4(1):2292486, 2024.

[32] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*, volume 25, pages 1097–1105, 2012.

[33] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[34] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016.

[35] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, volume 9351 of *Lecture Notes in Computer Science*, pages 234–241, 2015.

[36] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997.

[37] Alex Graves. Generating sequences with recurrent neural networks. *arXiv preprint arXiv:1308.0850*, 2013.

[38] John Nickolls, Ian Buck, Michael Garland, and Kevin Skadron. Scalable parallel programming with cuda. *ACM Queue*, 6(2):40–53, 2008.

[39] Klaus Schwab. *The Fourth Industrial Revolution.* Crown Business, New York, NY, 2016.

[40] Gissel Velarde. Artificial intelligence and its impact on the fourth industrial revolution: a review. *arXiv preprint arXiv:2011.03044*, 2020.

[41] Richard Van Noorden and Jeffrey M Perkel. Ai and science: what 1,600 researchers think. *Nature*, 621(7980):672–675, 2023.

[42] Catherine Wolfram, Orie Shelef, and Paul Gertler. How will energy demand develop in the developing world? *Journal of Economic Perspectives*, 26(1):119–138, 2012.

[43] Jerry L Holechek, Hatim ME Geli, Mohammed N Sawalhah, and Raul Valdez. A global assessment: can renewable energy replace fossil fuels by 2050? *Sustainability*, 14(8):4792, 2022.

[44] Tessaleno Devezas, Andrea Tick, Askar Sarygulov, and Polina Rukina. The slow pace of green transformation: Underlying factors and implications. *Energies*, 17(19), 2024.

[45] Heidar Jafarizadeh, Eliyad Yamini, Seyed Mohammad Zolfaghari, Farbod Esmaeilion, M El Haj Assad, and M Soltani. Navigating challenges in large-scale renewable energy storage: Barriers, solutions, and innovations. *Energy Reports*, 12:2179–2192, 2024.

[46] Yuqi Guo, Zhichao Hao, Shichong Zhao, Jiaqi Gong, and Fan Yang. Artificial intelligence in health care: bibliometric analysis. *Journal of medical Internet research*, 22(7):e18228, 2020.

[47] Alexander Windmann, Philipp Wittenberg, Marvin Schieseck, and Oliver Niggemann. Artificial intelligence in industry 4.0: A review of integration challenges for industrial systems. In *2024 IEEE 22nd International Conference on Industrial Informatics (INDIN)*, page 1–8. IEEE, August 2024.

[48] Claus Daniel, Jess C. Gehin, Kirsten Laurin-Kovitz, Bryan Morreale, Rick Stevens, and William Tumas. Ai for energy: Advanced research directions on ai for energy—report on the u.s. department of energy. Technical Report 1, Argonne National Laboratory, 2024.

[49] Simon P Stier, Christoph Kreisbeck, Holger Ihssen, Matthias Albert Popp, Jens Hauch, Kourosh Malek, Marine Reynaud, TPM Goumans, Johan Carlsson, Ilian Todorov, et al. Materials acceleration platforms (maps): accelerating materials research and development to meet urgent societal challenges. *Advanced Materials*, 36(45):2407791, 2024.

[50] Jerrit Wagner, Christian G Berger, Xiaoyan Du, Tobias Stubhan, Jens A Hauch, and Christoph J Brabec. The evolution of materials acceleration platforms: toward the laboratory of the future with amanda. *Journal of Materials Science*, 56:16422–16446, 2021.

[51] Milad Abolhasani and Eugenia Kumacheva. The rise of self-driving labs in chemical and materials sciences. *Nature Synthesis*, 2(6):483–492, 2023.

[52] Priynka Sharma, Maurizio Cirrincione, Ali Mohammadi, Giansalvo Cirrincione, and Rahul R Kumar. An overview of artificial intelligence-based techniques for pemfc system diagnosis. *IEEE Access*, 2024.

[53] Tetiana Zubatiuk and Olexandr Isayev. Development of multimodal machine learning potentials: toward a physics-aware artificial intelligence. *Accounts of Chemical Research*, 54(7):1575–1585, 2021.

[54] Qifeng Bai, Shuo Liu, Yanan Tian, Tingyang Xu, Antonio Jesús Banegas-Luna, Horacio Pérez-Sánchez, Junzhou Huang, Huanxiang Liu, and Xiaojun Yao. Application advances of deep learning methods for de novo drug design and molecular dynamics simulation. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 12(3):e1581, 2022.

[55] Qin Li, Nihang Fu, Sadman Sadeed Omee, and Jianjun Hu. Md-hit: Machine learning for material property prediction with dataset redundancy control. *npj Computational Materials*, 10(1):245, 2024.

[56] Grace Guinan, Addison Salvador, Michelle A Smeaton, Andrew Glaws, Hilary Egan, Brian C Wyatt, Babak Anasori, Kevin R Fiedler, Matthew J Olszta, and Steven R Spurgeon. Mind the gap: Bridging the divide between ai aspirations and the reality of autonomous characterization. *arXiv preprint arXiv:2502.18604*, 2025.

[57] Jun Ma, Yuting He, Feifei Li, Lin Han, Chenyu You, and Bo Wang. Segment anything in medical images. *Nature Communications*, 15(1):654, 2024.

[58] Uwe Schmidt, Martin Weigert, Coleman Broaddus, and Gene Myers. Cell detection with star-convex polygons. In *Medical image computing and computer assisted intervention–MICCAI 2018: 21st international conference, Granada, Spain, September 16-20, 2018, proceedings, part II 11*, pages 265–273. Springer, 2018.

[59] André Colliard-Granero, Mariah Batool, Jasna Jankovic, Jenia Jitsev, Michael H Eikerling, Kourosh Malek, and Mohammad J Eslamibidgoli. Deep learning for the automation of particle analysis in catalyst layers for polymer electrolyte fuel cells. *Nanoscale*, 14(1):10–18, 2022.

[60] Surender Kumar Sharma, Dalip Singh Verma, Latif Ullah Khan, Shalendra Kumar, and Sher Bahadar Khan. *Handbook of materials characterization*. Springer, 2018.

[61] André Colliard-Granero, Keusra Armel Gompou, Christian Rodenbücher, Kourosh Malek, Michael Eikerling, and Mohammad J Eslamibidgoli. Deep learning-enhanced characterization of bubble dynamics in proton exchange membrane water electrolyzers. *Physical Chemistry Chemical Physics*, 26:14529, 2024.

[62] Batuhan Yildirim and Jacqueline M. Cole. Bayesian particle instance segmentation for electron microscopy image quantification. *Journal of Chemical Information and Modeling*, 61(3):1136–1149, 2021. PMID: 33682402.

[63] Robin Girod, Timon Lazaridis, Hubert A Gasteiger, and Vasiliki Tileli. Three-dimensional nanoimaging of fuel cell catalyst layers. *Nature catalysis*, 6(5):383–391, 2023.

[64] Ying Da Wang, Quentin Meyer, Kunning Tang, James E McClure, Robin T White, Stephen T Kelly, Matthew M Crawford, Francesco Iacoviello, Dan JL Brett, Paul R Shearing, et al. Large-scale physically accurate modelling of real proton exchange membrane fuel cell with deep learning. *Nature communications*, 14(1):745, 2023.

[65] Kunning Tang, Quentin Meyer, Robin White, Ryan T Armstrong, Peyman Mostaghimi, Ying Da Wang, Shiyang Liu, Chuan Zhao, Klaus Regenauer-Lieb, and Patrick Kin Man Tung. Deep learning for full-feature x-ray microcomputed tomography segmentation of proton electron membrane fuel cells. *Computers & Chemical Engineering*, 161:107768, 2022.

[66] Khan Faraz, Jennifer Smith, Michael Lee, and David Jones. Deep learning detection of nanoparticles and multiple object tracking of their dynamic evolution during in situ etem studies. *Scientific Reports*, 12:2484, 2022.

[67] Ming Shen, Aarav Patel, and Rohit Kumar. A deep learning based automatic defect analysis framework for in situ tem ion irradiations. *arXiv preprint arXiv:2108.08882*, 2021.

[68] Wen Fu, Xin Zhao, Hui Li, and Ming Chen. Deep-learning-based prediction of nanoparticle phase transitions during in situ tem. *arXiv preprint arXiv:2205.11407*, 2022.

[69] Simon Müller, Brian Wood, Zongquan Cai, Qing Zhao, and David L. Wood. Accurate 3d representations of lithium-ion battery electrodes via deep-learning segmentation. *Nature Communications*, 12:2484, 2021.

[70] Engineering Advancement Group. Machine-learning assisted analysis of battery electrode by pfib-sem tomography. Technical report, EAG Laboratories, 2024.

[71] Jianguo He, Li Sun, and Peng Wang. Swin-unet for volumetric segmentation of porous media. In *ICCV Workshops*, 2023.

[72] Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *2016 fourth international conference on 3D vision (3DV)*, pages 565–571, 2016.

[73] Yu-Qi Yang, Yu-Xiao Guo, Jian-Yu Xiong, Yang Liu, Hao Pan, Peng-Shuai Wang, Xin Tong, and Baining Guo. Swin3d: A pretrained transformer backbone for 3d indoor scene understanding. *arXiv preprint arXiv:2304.06906*, 2023.

[74] Saining Xie, Ross Girshick, Piotr Dollár, Zhuowen Tu, and Kaiming He. Aggregated residual transformations for deep neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1492–1500, 2017.

[75] Yongchao Lu, Hong Wang, Lanting Zhang, Ning Yu, Siqi Shi, and Hang Su. Unleashing the power of ai in science-key considerations for materials data preparation. *Scientific Data*, 11(1):1039, 2024.

[76] Lauri Himanen, Amber Geurts, Adam Stuart Foster, and Patrick Rinke. Data-driven materials science: status, challenges, and perspectives. *Advanced Science*, 6(21):1900808, 2019.

[77] Mark D Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, Philip E Bourne, et al. The fair guiding principles for scientific data management and stewardship. *Scientific data*, 3(1):1–9, 2016.

[78] Maximiliaan Huisman, Mathias Hammer, Alex Rigano, Ulrike Boehm, James J Chambers, Nathalie Gaudreault, Alison J North, Jaime A Pimentel, Damir Sudar, Peter Bajcsy, et al. A perspective on microscopy metadata: data provenance and quality control. *arXiv preprint arXiv:1910.11370*, 2019.

[79] Alexander J Ropelewski, Megan A Rizzo, Jason R Swedlow, Jan Huisken, Pavel Osten, Neda Khanjani, Kurt Weiss, Vesselina Bakalov, Michelle Engle, Lauren Gridley, et al. Standard metadata for 3d microscopy. *Scientific data*, 9(1):449, 2022.

[80] Ben Blaiszik, Kyle Chard, Jim Pruyne, Rachana Ananthakrishnan, Steve Tuecke, and Ian Foster. The materials data facility: Data services to advance materials science research. *JOM*, 68(8):2045–2052, 2016.

[81] Francesco De Carlo, Doğa Gürsoy, Daniel J. Ching, K. Joost Batenburg, Wolfgang Ludwig, Lucia Mancini, and Mark Rivers. Tomobank: a tomographic data repository for computational x-ray science. *Measurement Science and Technology*, 29(3):034004, 2018.

[82] Andrii Iudin, Paul K. Korir, José Salavert-Torres, Gerard J. Kleywegt, and Ardan Patwardhan. Empiar: a public archive for raw electron microscopy image data. *Nature Methods*, 13:387–388, 2016.

[83] Leopold Talirz, Snehal Kumbhar, Elsa Passaro, Aliaksandr V. Yakutovich, Valeria Granata, Fernando Gargiulo, Marco Borelli, Martin Uhrin, Sebastiaan P. Huber, Spyros Zoupanos, Carl S. Adorf, Casper W. Andersen, Ole Schütt, Carlo A. Pignedoli, Daniele Passerone, Joost VandeVondele, Thomas C. Schulthess, Berend Smit, Giovanni Pizzi, and Nicola Marzari. Materials cloud, a platform for open computational science. *Scientific Data*, 7:299, 2020.

[84] William Kent. *Data and reality: a timeless perspective on perceiving and managing information*. Technics publications, 2012.

[85] Dan Gusfield. *Algorithms on Strings, Trees, and Sequences: Computer Science and Computational Biology*. Cambridge University Press, Cambridge, UK, 1997.

[86] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing*. Pearson, Upper Saddle River, NJ, 3rd edition, 2008.

[87] Anil K. Jain. *Fundamentals of Digital Image Processing*. Prentice Hall, Englewood Cliffs, NJ, 1989.

[88] Gregory K. Wallace. The jpeg still picture compression standard. *Communications of the ACM*, 34(4):30–44, 1991.

[89] Jiang Hsieh. *Computed Tomography: Principles, Design, Artifacts, and Recent Advances*. SPIE Press, 2003.

[90] Philip J. Withers. X-ray tomographic methods for materials science. *International Materials Reviews*, 53(3):129–181, 2008.

[91] Eric Maire and Philip J. Withers. Quantitative x-ray tomography. *International Materials Reviews*, 59(1):1–43, 2014.

[92] Davood Jafari et al. Time-resolved in situ visualization of gas evolution in li-ion batteries. *Advanced Energy Materials*, 8(12), 2018.

[93] Giovanni M Vanacore, Anthony WP Fitzpatrick, and Ahmed H Zewail. Four-dimensional electron microscopy: Ultrafast imaging, diffraction and spectroscopy in materials science and biology. *Nano Today*, 11(2):228–249, 2016.

[94] Jenn Riley. *Understanding Metadata: What is Metadata, and What is it For? A Primer*. National Information Standards Organization, Baltimore, MD, 2017.

[95] International Organization for Standardization. Information and documentation – guidelines for bibliographic references and citations to information resources. *ISO 690:2010*, 2010.

[96] International Organization for Standardization. Information and documentation – the dublin core metadata element set – part 1: Core elements. *ISO 15836-1:2017*, 2017.

[97] David Haynes. *Metadata for Information Management and Retrieval: Understanding Metadata and Its Use*. Facet Publishing, Bristol, UK, 2nd edition, 2018.

[98] Nicola Guarino. *Formal ontology in information systems: Proceedings of the first international conference (FOIS'98), June 6-8, Trento, Italy*, volume 46. IOS press, 1998.

[99] Gerhard Goldbeck, Emanuele Ghedini, Adham Hashibon, Georg J Schmitz, and Jesper Friis. A reference language and ontology for materials modelling and interop-

erability. In *National Agency for Finite Element Methods and Standards (NAFEMS World Congress) 2019*, 2019.

[100] Pierluigi Del Nostro, Gerhard Goldbeck, and Daniele Toti. Chameo: An ontology for the harmonisation of materials characterisation methodologies. *Applied Ontology*, 17(3):401–421, 2022.

[101] Simon Clark, Casper Welzel Andersen, Eibar Flores, Francesca Lønstad Bleken, and Jesper Friis. (digital presentation) a battery interface ontology for data interoperability and semantic knowledge representation. In *Electrochemical Society Meeting Abstracts 242*, number 7 in 1, pages 2582–2582, 2022.

[102] Claudia Draxl and Matthias Scheffler. The nomad laboratory: from data sharing to artificial intelligence. *Journal of Physics: Materials*, 2(3):036001, 2019.

[103] Simon Clark, Xavier Raynaud, Otello M. Roscioni, Francesca L. Bleken, and Jesper Friis. emmo-repo/domain-chemical-substance: 0.12.2-beta (0.12.2-beta), 2025. Accessed May 23, 2025.

[104] Korry Douglas and Susan Douglas. *PostgreSQL: a comprehensive guide to building, programming, and administering PostgresSQL databases*. SAMS publishing, 2003.

[105] Paul DuBois. *MySQL*. New riders publishing, 1999.

[106] Shannon Bradshaw, Eoin Brazil, and Kristina Chodorow. *MongoDB: the definitive guide: powerful and scalable data storage*. " O'Reilly Media, Inc.", 2019.

[107] Justin J Miller. Graph database applications and concepts with neo4j. In *Proceedings of the southern association for information systems conference, Atlanta, GA, USA*, volume 2324, pages 141–147, 2013.

[108] François Chollet. *Deep learning with Python*. Simon and Schuster, 2021.

[109] John Canny. A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence*, 6:679–698, 1986.

[110] Victor Francisco Rodriguez-Galiano, Bardan Ghimire, John Rogan, Mario Chica-Olmo, and Juan Pedro Rigol-Sanchez. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS journal of photogrammetry and remote sensing*, 67:93–104, 2012.

[111] Jianxin Wu. Introduction to convolutional neural networks. *National Key Lab for Novel Software Technology. Nanjing University. China*, 5(23):495, 2017.

[112] Aurélien Géron. *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems.* " O'Reilly Media, Inc.", 2022.

[113] Stuart J Russell and Peter Norvig. *Artificial intelligence: a modern approach*. pearson, 2016.

[114] Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. Language models are few-shot learners. *arXiv preprint arXiv:2005.14165*, 2020.

[115] Daya Guo, Dejian Yang, Haowei Zhang, Junxiao Song, Ruoyu Zhang, Runxin Xu, Qihao Zhu, Shirong Ma, Peiyi Wang, Xiao Bi, et al. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning. *arXiv preprint arXiv:2501.12948*, 2025.

[116] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation, 2015.

[117] Özgün Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger. 3d u-net: learning dense volumetric segmentation from sparse annotation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19*, pages 424–432, 2016.

[118] Idriss Sinapan, Christophe Lin-Kwong-Chon, Cédric Damour, Jean-Jacques Amangoua Kadjo, and Michel Benne. Oxygen bubble dynamics in pem water electrolyzers with a deep-learning-based approach. *Hydrogen*, 4(3):556–572, 2023.

[119] Jason Keonhag Lee and Aimy Bazylak. Bubbles: the good, the bad, and the ugly. *Joule*, 5(1):19–21, 2021.

[120] Jude O Majasan, Jason IS Cho, Ishanka Dedigama, Dimitrios Tsaoulidis, Paul Shearing, and Dan JL Brett. Two-phase flow behaviour and performance of polymer electrolyte membrane electrolysers: Electrochemical and optical characterisation. *international journal of hydrogen energy*, 43(33):15659–15672, 2018.

[121] Ishanka Dedigama, Pangiota Angeli, Nicholas van Dijk, Jason Millichamp, Dimitrios Tsaoulidis, Paul R Shearing, and Daniel JL Brett. Current density mapping and optical flow visualisation of a polymer electrolyte membrane water electrolyser. *Journal of Power Sources*, 265:97–103, 2014.

[122] Farid Aubras, J Deseure, J-JA Kadjo, I Dedigama, J Majasan, Brigitte Grondin-Perez, J-P Chabriat, and DJL Brett. Two-dimensional model of low-pressure pem electrolyser: Two-phase flow regime, electrochemical modelling and experimental validation. *International journal of hydrogen energy*, 42(42):26203–26216, 2017.

[123] Xin Su, Lijun Xu, and Bing Hu. Simulation of proton exchange membrane electrolyzer: Influence of bubble covering. *International Journal of Hydrogen Energy*, 47(46):20027–20039, 2022.

[124] Maximilian Maier, Quentin Meyer, Jude Majasan, Rhodri E Owen, James B Robinson, James Dodwell, Yunsong Wu, Luis Castanheira, Gareth Hinds, Paul R Shearing, et al. Diagnosing stagnant gas bubbles in a polymer electrolyte membrane water electrolyser using acoustic emission. *Frontiers in Energy Research*, 8:582919, 2020.

[125] AB Pandit, J Varley, RB Thorpe, and JF Davidson. Measurement of bubble size distribution: an acoustic technique. *Chemical Engineering Science*, 47(5):1079–1089, 1992.

[126] Gary G Yen and Haiming Lu. Acoustic emission data assisted process monitoring. *ISA transactions*, 41(3):273–282, 2002.

[127] Ishanka Dedigama, Daniel JL Brett, Thomas J Mason, Jason Millichamp, Paul R Shearing, and Katherine E Ayers. An electrochemical impedance spectroscopy study and two phase flow analysis of the anode of polymer electrolyte membrane water electrolyser. *ECS Transactions*, 68(3):117, 2015.

[128] Yadong Zhang, Jinbing Zhang, Yuemin Zhao, Xiangyu Zhang, Xuliang Yang, Enhui Zhou, Chenlong Duan, Guanghui Wang, and Liang Dong. Investigations on dynamics of bubble in a 2d vibrated fluidized bed using pressure drop signal and high-speed image analysis. *Chemical Engineering Journal*, 395:125129, 2020.

[129] Giorgio Besagni, Pietro Brazzale, Alberto Fiocca, and Fabio Inzoli. Estimation of bubble size distributions and shapes in two-phase bubble column using image analysis and optical probes. *Flow Measurement and Instrumentation*, 52:190–207, 2016.

[130] Mohammad J Eslamibidgoli, Fabian P Tipp, Jenia Jitsev, Jasna Jankovic, Michael H Eikerling, and Kourosh Malek. Convolutional neural networks for high throughput screening of catalyst layer inks for polymer electrolyte fuel cells. *RSC advances*, 11(51):32126–32134, 2021.

[131] André Colliard-Granero, Jenia Jitsev, Michael H Eikerling, Kourosh Malek, and Mohammad J Eslamibidgoli. Utile-gen: Automated image analysis in nanoscience using synthetic dataset generator and deep learning. *ACS Nanoscience Au*, page 398–407, 2023.

[132] Ali Malek, Mohammad Javad Eslamibidgoli, Mehrdad Mokhtari, Qianpu Wang, Michael H Eikerling, and Kourosh Malek. Virtual materials intelligence for design and discovery of advanced electrocatalysts. *ChemPhysChem*, 20(22):2946–2955, 2019.

[133] Shoukun Sun, Fei Xu, Lu Cai, Daniele Salvato, Fidelma Dilemma, Luca Capriotti, Min Xian, and Tiankai Yao. An efficient instance segmentation approach for extracting fission gas bubbles on u-10zr annular fuel, 2023.

[134] Chris M Anderson, Jacob Klein, Heygaan Rajakumar, Colin D Judge, and Laurent Karim Beland. Automated detection of helium bubbles in irradiated x-750. *Ultramicroscopy*, 217:113068, 2020.

[135] Yewon Kim and Hyungmin Park. Deep learning-based automated and universal bubble detection and mask extraction in complex two-phase flows. *Scientific reports*, 11(1):8940, 2021.

[136] Maxim Tkachenko, Mikhail Malyuk, Andrey Holmanyuk, and Nikolai Liubimov. Label Studio: Data labeling software, 2020-2022. Open source software available from https://github.com/heartexlabs/label-studio, Accessed May 23, 2025.

[137] Ozan Oktay, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori, Steven McDonagh, Nils Y Hammerla, Bernhard Kainz, Ben Glocker, and Daniel Rueckert. Attention u-net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999*, 2018.

[138] Gary Bradski. The opencv library. *Dr. Dobb's Journal: Software Tools for the Professional Programmer*, 25(11):120–123, 2000.

[139] Lucas von Chamier, Romain F Laine, Johanna Jukkala, Christoph Spahn, Daniel Krentzel, Elias Nehme, Martina Lerche, Sara Hernández-Pérez, Pieta K Mattila, Eleni Karinou, et al. Democratising deep learning for microscopy with zerocostdl4mic. *Nature communications*, 12(1):2276, 2021.

[140] Tianxi Yan, Yahong Liu, Dong Wei, Xiaofan Sun, and Qing Liu. Shape analysis of sand particles based on fourier descriptors. *Environmental Science and Pollution Research*, 30(22):62803–62814, 2023.

[141] Mark L Hentschel and Neil Page. Selection of descriptors for particle shape characterization. *Particle & Particle Systems Characterization*, 20:25 – 38, 03 2003.

[142] Yanfang Yu, Peng Yang, Huibo Meng, Huanchen Liu, Jiawei Zhang, and Henglei Yu. Bubble morphology analysis and pressure drop of gas–liquid two-phase flow inside a quarto static mixer. *Industrial & Engineering Chemistry Research*, 61(50):18574–18587, 2022.

[143] André Colliard-Granero, Kangjun Duan, Roswitha Zeis, Kourosh Malek, Michael Eikerling, and Mohammad J Eslamibidgoli. Advancing vanadium redox flow battery analysis: A deep learning approach for high-throughput 3d visualization and bubble quantification. *Digital Discoveries*, accepted, 2025.

[144] Zebo Huang and Anle Mu. Research and analysis of performance improvement of vanadium redox flow battery in microgrid: A technology review. *International Journal of Energy Research*, 45(10):14170–14193, 2021.

[145] Kyle Lourenssen, James Williams, Faraz Ahmadpour, Ryan Clemmer, and Syeda Tasnim. Vanadium redox flow batteries: A comprehensive review. *Journal of Energy Storage*, 25:100844, 2019.

[146] Ruijie Ye, Dirk Henkensmeier, Sang Jun Yoon, Zhifeng Huang, Dong Kyu Kim, Zhenjun Chang, Sangwon Kim, and Ruiyong Chen. Redox flow batteries for energy storage: a technology review. *Journal of electrochemical energy conversion and storage*, 15(1):010801, 2018.

[147] Alexandre Lucas and Stamatios Chondrogiannis. Smart grid energy storage controller for frequency regulation and peak shaving, using a vanadium redox flow battery. *International Journal of Electrical Power & Energy Systems*, 80:26–36, 2016.

[148] Wei-Chieh Huang, Qianzhi Zhang, and Fengqi You. Impacts of battery energy storage technologies and renewable integration on the energy transition in the new york state. *Advances in Applied Energy*, 9:100126, 2023.

[149] Kyeongmin Oh, Seongyeon Won, and Hyunchul Ju. Numerical study of the effects of carbon felt electrode compression in all-vanadium redox flow batteries. *Electrochimica acta*, 181:13–23, 2015.

[150] D Saygin, OB Tör, ME Cebeci, S Teimourzadeh, and P Godron. Increasing turkey's power system flexibility for grid integration of 50 *Energy Strategy Reviews*, 34:100625, 2021.

[151] Sampson Tetteh, Maryam Roza Yazdani, and Annukka Santasalo-Aarnio. Cost-effective electro-thermal energy storage to balance small scale renewable energy systems. *Journal of Energy Storage*, 41:102829, 2021.

[152] Tossaporn Jirabovornwisut and Amornchai Arpornwichanop. A review on the electrolyte imbalance in vanadium redox flow batteries. *International Journal of Hydrogen Energy*, 44(45):24485–24509, 2019.

[153] Xiao-Zi Yuan, Chaojie Song, Alison Platt, Nana Zhao, Haijiang Wang, Hui Li, Khalid Fatih, and Darren Jang. A review of all-vanadium redox flow battery durability: Degradation mechanisms and mitigation strategies. *International Journal of Energy Research*, 43(13):6599–6638, 2019.

[154] Burak Turker, Sebastian Arroyo Klein, Eva-Maria Hammer, Bettina Lenz, and Lidiya Komsiyska. Modeling a vanadium redox flow battery system for large scale applications. *Energy conversion and management*, 66:26–32, 2013.

[155] Maria Skyllas-Kazacos, Liuyue Cao, Michael Kazacos, Nadeem Kausar, and Asem Mousa. Vanadium electrolyte studies for the vanadium redox battery—a review. *ChemSusChem*, 9(13):1521–1543, 2016.

[156] Laszlo Eifert, Rupak Banerjee, Zenonas Jusys, and Roswitha Zeis. Characterization of carbon felt electrodes for vanadium redox flow batteries: impact of treatment methods. *Journal of The Electrochemical Society*, 165(11):A2577, 2018.

[157] Kerstin Köble, Alexey Ershov, Kangjun Duan, Monja Schilling, Alexander Rampf, Angelica Cecilia, Tomáš Faragó, Marcus Zuber, Tilo Baumbach, and Roswitha Zeis. Insights into the hydrogen evolution reaction in vanadium redox flow batteries: A synchrotron radiation based x-ray imaging study. *Journal of Energy Chemistry*, 91:132–144, 2024.

[158] Monja Schilling, Michael Braig, Kerstin Köble, and Roswitha Zeis. Investigating the v (iv)/v (v) electrode reaction in a vanadium redox flow battery–a distribution of relaxation times analysis. *Electrochimica Acta*, 430:141058, 2022.

[159] Ruediger Schweiss, Alexander Pritzl, and Christian Meiser. Parasitic hydrogen evolution at different carbon fiber electrodes in vanadium redox flow batteries. *Journal of The Electrochemical Society*, 163(9):A2089, 2016.

[160] Igor Derr, Michael Bruns, Joachim Langner, Abdulmonem Fetyan, Julia Melke, and Christina Roth. Degradation of all-vanadium redox flow batteries (vrfb) investigated by electrochemical impedance and x-ray photoelectron spectroscopy: Part 2 electrochemical degradation. *Journal of Power Sources*, 325:351–359, 2016.

[161] Yu-Jia Zhang, Qiang Ye, and Meng Ni. The impact of in-situ hydrogen evolution on the flow resistance of electrolyte flowing through the carbon felt electrode in a redox flow battery. *Journal of Power Sources*, 564:232837, 2023.

[162] Laszlo Eifert, Zenonas Jusys, Rolf Jürgen Behm, and Roswitha Zeis. Side reactions and stability of pre-treated carbon felt electrodes for vanadium redox flow batteries: A dems study. *Carbon*, 158:580–587, 2020.

[163] Jason Thomas Clement. Investigation of localized performance and gas evolution in all-vanadium redox flow batteries via in-situ distributed diagnostic techniques. *PhD diss.*, 2016.

[164] László Eifert, Nico Bevilacqua, Kerstin Köble, Kieran Fahy, Liusheng Xiao, Min Li, Kangjun Duan, Aimy Bazylak, Pang-Chieh Sui, and Roswitha Zeis. Synchrotron x-ray radiography and tomography of vanadium redox flow batteries—cell design, electrolyte flow geometry, and gas bubble formation. *ChemSusChem*, 13(12):3154–3165, 2020.

[165] Abdulmonem Fetyan, Gumaa A El-Nagar, Iver Lauermann, Maike Schnucklake, Jonathan Schneider, and Christina Roth. Detrimental role of hydrogen evolution and its temperature-dependent impact on the performance of vanadium redox flow batteries. *Journal of Energy Chemistry*, 32:57–62, 2019.

[166] Jihye Lee, Jules Tshishimbi Muya, Hoeil Chung, and Jinho Chang. Unraveling v (v)-v (iv)-v (iii)-v (ii) redox electrochemistry in highly concentrated mixed acidic media for a vanadium redox flow battery: origin of the parasitic hydrogen evolution reaction. *ACS applied materials & interfaces*, 11(45):42066–42077, 2019.

[167] Lei Wei, TS Zhao, Qian Xu, XL Zhou, and ZH Zhang. In-situ investigation of hydrogen evolution behavior in vanadium redox flow batteries. *Applied Energy*, 190:1112–1118, 2017.

[168] Giorgio Besagni, Pietro Brazzale, Alberto Fiocca, and Fabio Inzoli. Estimation of bubble size distributions and shapes in two-phase bubble column using image analysis and optical probes. *Flow Measurement and Instrumentation*, 52:190–207, 2016.

[169] Maximilian Maier, Quentin Meyer, Jude Majasan, Rhodri E Owen, James B Robinson, James Dodwell, Yunsong Wu, Luis Castanheira, Gareth Hinds, and Paul R Shearing. Diagnosing stagnant gas bubbles in a polymer electrolyte membrane water electrolyser using acoustic emission. *Frontiers in Energy Research*, 8:582919, 2020.

[170] AB Pandit, J Varley, RB Thorpe, and JF Davidson. Measurement of bubble size distribution: an acoustic technique. *Chemical Engineering Science*, 47(5):1079–1089, 1992.

[171] Jihye Lee, Jules Tshishimbi Muya, Hoeil Chung, and Jinho Chang. Unraveling v (v)-v (iv)-v (iii)-v (ii) redox electrochemistry in highly concentrated mixed acidic media for a vanadium redox flow battery: origin of the parasitic hydrogen evolution reaction. *ACS applied materials & interfaces*, 11(45):42066–42077, 2019.

[172] Kerstin Köble, László Eifert, Nico Bevilacqua, Kieran F Fahy, Aimy Bazylak, and Roswitha Zeis. Synchrotron x-ray radiography of vanadium redox flow batteries–time and spatial resolved electrolyte flow in porous carbon electrodes. *Journal of Power Sources*, 492:229660, 2021.

[173] Nico Bevilacqua, László Eifert, Rupak Banerjee, Kerstin Köble, Tomáš Faragó, Marcus Zuber, Aimy Bazylak, and Roswitha Zeis. Visualization of electrolyte flow in vanadium redox flow batteries using synchrotron x-ray radiography and tomography–impact of electrolyte species and electrode compression. *Journal of Power Sources*, 439:227071, 2019.

[174] Ashley P Black, Andrea Sorrentino, François Fauth, Ibraheem Yousef, Laura Simonelli, Carlos Frontera, Alexandre Ponrouch, Dino Tonti, and M Rosa Palacín. Synchrotron radiation based operando characterization of battery materials. *Chemical science*, 14(7):1641–1665, 2023.

[175] Purna C Ghimire, Arjun Bhattarai, Tuti M Lim, Nyunt Wai, Maria Skyllas-Kazacos, and Qingyu Yan. In-situ tools used in vanadium redox flow battery research. *Batteries*, 7(3):53, 2021.

[176] Yue Wen and Rhodri Jervis. New strategies for interrogation of redox flow batteries via synchrotron radiation. *Current Opinion in Chemical Engineering*, 37:100836, 2022.

[177] Kunning Tang, Quentin Meyer, Robin White, Ryan T Armstrong, Peyman Mostaghimi, Ying Da Wang, Shiyang Liu, Chuan Zhao, Klaus Regenauer-Lieb, and Patrick Kin Man Tung. Deep learning for full-feature x-ray microcomputed

tomography segmentation of proton electron membrane fuel cells. *Computers & Chemical Engineering*, 161:107768, 2022.

[178] Zhenhao Li, Ling Qin, Baisong Guo, Junping Yuan, Zhiguo Zhang, Wei Li, and Jiawei Mi. Characterization of the convoluted 3d intermetallic phases in a recycled al alloy by synchrotron x-ray tomography and machine learning. *Acta Metallurgica Sinica (English Letters)*, pages 1–9, 2021.

[179] Reed Kopp, Joshua Joseph, Xinchen Ni, Nicholas Roy, and Brian L Wardle. Deep learning unlocks x-ray microtomography segmentation of multiclass microdamage in heterogeneous materials. *Advanced Materials*, 34(11):2107817, 2022.

[180] Dena Kartouzian, Arezou Mohseninia, Henning Markötter, Joachim Scholta, and Ingo Manke. Neutron tomographic investigation of the effect of hydrophobicity gradients within mpl and mea on the spatial distribution and transport of liquid water in pemfcs. *ECS Transactions*, 85(13):927, 2018.

[181] Arezou Mohseninia, Dena Kartouzian, Henning Markötter, Utku U Ince, Ingo Manke, and Joachim Scholta. Neutron radiographic investigations on the effect of hydrophobicity gradients within mpl and mea on liquid water distribution and transport in pemfcs. *ECS Transactions*, 85(13):1013, 2018.

[182] Kerstin Köble, Monja Schilling, László Eifert, Nico Bevilacqua, Kieran F Fahy, Plamen Atanassov, Aimy Bazylak, and Roswitha Zeis. Revealing the multifaceted impacts of electrode modifications for vanadium redox flow battery electrodes. *ACS Applied Materials & Interfaces*, 15(40):46775–46789, 2023.

[183] Tomasz W Wysokinski, Dean Chapman, Gregg Adams, Michel Renier, Pekka Suortti, and William Thomlinson. Beamlines of the biomedical imaging and therapy facility at the canadian light source—part 1. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 582(1):73–76, 2007.

[184] Matthias Vogelgesang, Suren Chilingaryan, Tomy dos Santos Rolo, and Andreas Kopmann. Ufo: a scalable gpu-based image processing framework for on-line monitoring. In *2012 IEEE 14th International conference on high performance computing and communication & 2012 IEEE 9th international conference on embedded software and systems*, pages 824–829, 2012.

[185] Burak Turker, Sebastian Arroyo Klein, Eva-Maria Hammer, Bettina Lenz, and Lidiya Komsiyska. Modeling a vanadium redox flow battery system for large scale applications. *Energy conversion and management*, 66:26–32, 2013.

[186] Maria Skyllas-Kazacos, Liuyue Cao, Michael Kazacos, Nadeem Kausar, and Asem Mousa. Vanadium electrolyte studies for the vanadium redox battery—a review. *ChemSusChem*, 9(13):1521–1543, 2016.

[187] Ozan Oktay, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori, Steven McDonagh, Nils Y Hammerla, and Bernhard Kainz. Attention u-net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999*, 2018.

[188] Huimin Huang, Lanfen Lin, Ruofeng Tong, Hongjie Hu, Qiaowei Zhang, Yutaro Iwamoto, Xianhua Han, Yen-Wei Chen, and Jian Wu. Unet 3+: A full-scale connected unet for medical image segmentation. In *ICASSP 2020-2020 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pages 1055–1059. IEEE, 2020.

[189] Hu Cao, Yueyue Wang, Joy Chen, Dongsheng Jiang, Xiaopeng Zhang, Qi Tian, and Manning Wang. Swin-unet: Unet-like pure transformer for medical image segmentation. In *European conference on computer vision*, pages 205–218. Springer, 2022.

[190] Gary Bradski. The opencv library. *Dr. Dobb's Journal: Software Tools for the Professional Programmer*, 25(11):120–123, 2000.

[191] William J Schroeder, Lisa Sobierajski Avila, and William Hoffman. Visualizing with vtk: a tutorial. *IEEE Computer graphics and applications*, 20(5):20–27, 2000.

[192] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[193] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision*, pages 2980–2988, 2017.

[194] André Colliard-Granero, Salvatore Ranieri, Aimy Bazylak, Tobias Morawietz, K Andreas Friedrich, Jasna Jankovic, Michael H Eikerling, Kourosh Malek, and Mohammad J Eslamibidgoli. Utile-pore: Deep learning-enabled 3d analysis of porous materials in polymer electrolyte membrane-based energy devices. *Journal of The Electrochemical Society*, 172(7):074515, 2025.

[195] Yujing Guo, Gendi Li, Junbo Zhou, and Yong Liu. Comparison between hydrogen production by alkaline water electrolysis and hydrogen production by pem electrolysis. In *IOP Conference Series: Earth and Environmental Science*, volume 371, page 042022. IOP Publishing, 2019.

[196] Idriss Sinapan, Christophe Lin-Kwong-Chon, Cédric Damour, Jean-Jacques Amangoua Kadjo, and Michel Benne. Oxygen bubble dynamics in pem water electrolyzers with a deep-learning-based approach. *Hydrogen*, 4(3):556–572, 2023.

[197] Marcelo Carmo, David L Fritz, Jürgen Mergel, and Detlef Stolten. A comprehensive review on pem water electrolysis. *International journal of hydrogen energy*, 38(12):4901–4934, 2013.

[198] Michael Eikerling, Alexei A Kornyshev, and Andrei A Kulikovsky. Physical modeling of fuel cells and their components. *Encyclopedia of Electrochemistry: Online*, 2007.

[199] Michael Eikerling and Andrei Kulikovsky. *Polymer electrolyte fuel cells: physical principles of materials and operation*. Crc Press, 2014.

[200] Eva Wallnöfer-Ogris, Ilena Grimmer, Matthias Ranz, Martin Höglinger, Stefan Kartusch, Julius Rauh, Marie-Gabrielle Macherhammer, Bianca Grabner, and Alexander Trattner. A review on understanding and identifying degradation mechanisms in pem water electrolysis cells: Insights for stack application, development, and research. *International Journal of Hydrogen Energy*, 65:381–397, 2024.

[201] Frank A de Bruijn, VAT Dam, and GJM Janssen. Durability and degradation issues of pem fuel cell components. *Fuel cells*, 8(1):3–22, 2008.

[202] Hong Xu, Minna Bührer, Federica Marone, Thomas J Schmidt, Felix N Büchi, and Jens Eller. Effects of gas diffusion layer substrates on pefc water management: Part i. operando liquid water saturation and gas diffusion properties. *Journal of The Electrochemical Society*, 168(7):074505, 2021.

[203] N Ge, R Banerjee, D Muirhead, J Lee, H Liu, P Shrestha, AKC Wong, J Jankovic, M Tam, and D Susac. Membrane dehydration with increasing current density at high inlet gas relative humidity in polymer electrolyte membrane fuel cells. *Journal of Power Sources*, 422:163–174, 2019.

[204] Jian Zhao, Huiyuan Liu, and Xianguo Li. Structure, property, and performance of catalyst layers in proton exchange membrane fuel cells. *Electrochemical Energy Reviews*, 6(1):13, 2023.

[205] Yen-Chun Chen, Tim Dörenkamp, Christoph Csoklich, Anne Berger, Federica Marone, Jens Eller, Thomas J Schmidt, and Felix N Büchi. On the water transport mechanism through the microporous layers of operando polymer electrolyte fuel cells probed directly by x-ray tomographic microscopy. *Energy Advances*, 2(9):1447–1463, 2023.

[206] P Antonacci, S Chevalier, J Lee, N Ge, J Hinebaugh, R Yip, Y Tabuchi, T Kotaka, and A Bazylak. Balancing mass transport resistance and membrane resistance when tailoring microporous layer thickness for polymer electrolyte membrane fuel cells operating at high current densities. *Electrochimica Acta*, 188:888–897, 2016.

[207] J Lee, R Yip, P Antonacci, N Ge, T Kotaka, Y Tabuchi, and A Bazylak. Synchrotron investigation of microporous layer thickness on liquid water distribution in a pem fuel cell. *Journal of The Electrochemical Society*, 162(7):F669, 2015.

[208] MH Shojaeefard, GR Molaeimanesh, M Nazemian, and MR Moqaddari. A review on microstructure reconstruction of pem fuel cells porous electrodes for pore scale simulation. *International Journal of Hydrogen Energy*, 41(44):20276–20293, 2016.

[209] U Panchenko, Thomas Arlt, I Manke, M Müller, D Stolten, and Werner Lehnert. Synchrotron radiography for a proton exchange membrane (pem) electrolyzer. *Fuel Cells*, 20(3):300–306, 2020.

[210] Mohamed Ahmed-Maloum, Thomas David, Laure Guetaz, Arnaud Morin, Joël Pauchet, Michel Quintard, and Marc Prat. Characterizing pem fuel cell catalyst layer properties from high resolution three-dimensional digital images, part i: A numerical procedure for the ionomer distribution reconstruction. *International Journal of Hydrogen Energy*, 80:39–56, 2024.

[211] Steven Claes, Wouter Van De Walle, Muhammad Islahuddin, and Hans Janssen. The application of computed tomography for characterizing the pore structure of building materials. *Journal of Building Physics*, 43(4):254–276, 2020.

[212] Mehdi Mahdaviara, Mohammad Javad Shojaei, Javad Siavashi, Mohammad Sharifi, and Martin J Blunt. Deep learning for multiphase segmentation of x-ray images of gas diffusion layers. *Fuel*, 345:128180, 2023.

[213] KN Abrosimov, KM Gerke, IN Semenkov, and DV Korost. Otsu's algorithm in the segmentation of pore space in soils based on tomographic data. *Eurasian Soil Science*, 54(4):560–571, 2021.

[214] Iryna V Zenyuk, Dilworth Y Parkinson, Liam G Connolly, and Adam Z Weber. Gas-diffusion-layer structural properties under compression via x-ray tomography. *Journal of Power Sources*, 328:364–376, 2016.

[215] Pavel Iassonov, Thomas Gebrenegus, and Markus Tuller. Segmentation of x-ray computed tomography images of porous materials: A crucial step for characterization and quantitative analysis of pore structures. *Water resources research*, 45(9), 2009.

[216] H. Li, T. Qiao, and X. Ding. Effect of grayscale threshold on x-ray computed tomography reconstruction of gas diffusion layers in polymer electrolyte membrane fuel cells. *Heliyon*, 10(7):e29378, 2024.

[217] Taylr Cawte and Aimy Bazylak. A 3d convolutional neural network accurately predicts the permeability of gas diffusion layer materials directly from image data. *Current Opinion in Electrochemistry*, 35:101101, 2022.

[218] Xinxin Liu, Kexin Fan, Xinmeng Huang, Jiankai Ge, Yujie Liu, and Haisu Kang. Recent advances in artificial intelligence boosting materials design for electrochemical energy storage. *Chemical Engineering Journal*, page 151625, 2024.

[219] Alán Aspuru-Guzik and Kristin Persson. Materials acceleration platform: Accelerating advanced energy materials discovery by integrating high-throughput methods and artificial intelligence. *Mission Innovation*, 2018.

[220] Mariah Batool, Oluwafemi Sanumi, and Jasna Jankovic. Application of artificial intelligence in the materials science, with a special focus on fuel cells and electrolyzers. *Energy and AI*, 18:100424, 2024.

[221] Zhe Deng, Xing Lin, Zhenyu Huang, Jintao Meng, Yun Zhong, Guangting Ma, Yu Zhou, Yue Shen, Han Ding, and Yunhui Huang. Recent progress on advanced imaging techniques for lithium-ion batteries. *Advanced Energy Materials*, 11(2):2000806, 2021.

[222] Ralf F Ziesche, Thomas MM Heenan, Pooja Kumari, Jarrod Williams, Weiqun Li, Matthew E Curd, Timothy L Burnett, Ian Robinson, Dan JL Brett, and Matthias J Ehrhardt. Multi-dimensional characterization of battery materials. *Advanced Energy Materials*, 13(23):2300103, 2023.

[223] Andreas Pfrang, Stephan Didas, and Georgios Tsotridis. X-ray computed tomography of gas diffusion layers of pem fuel cells: Segmentation of the microporous layer. *Journal of power sources*, 235:81–86, 2013.

[224] Andrew D Shum, Christopher P Liu, Wei Han Lim, Dilworth Y Parkinson, and Iryna V Zenyuk. Using machine learning algorithms for water segmentation in gas diffusion layers of polymer electrolyte fuel cells. *Transport in Porous Media*, 144(3):715–737, 2022.

[225] Andreas Grießer, Rolf Westerteiger, Erik Glatt, Hans Hagen, and Andreas Wiegmann. Deep learning based segmentation of binder and fibers in gas diffusion layers. *Next Materials*, 6:100411, 2025.

[226] Caroline A. Schneider, Wayne S. Rasband, and Kevin W. Eliceiri. Nih image to imagej: 25 years of image analysis. *Nature Methods*, 9(7):671–675, 2012.

[227] Matthias Arzt, Joran Deschamps, Christopher Schmied, Tobias Pietzsch, Deborah Schmidt, Pavel Tomancak, Robert Haase, and Florian Jug. Labkit: labeling and segmentation toolkit for big image data. *Frontiers in computer science*, 4:777728, 2022.

[228] Jeff T Gostick, Zohaib A Khan, Thomas G Tranter, Matthew DR Kok, Mehrez Agnaou, Mohammadamin Sadeghi, and Rhodri Jervis. Porespy: A python toolkit for quantitative analysis of porous media images. *Journal of Open Source Software*, 4(37):1296, 2019.

[229] Esra Erdim, Ömer Akgiray, and İbrahim Demir. A revisit of pressure drop-flow rate correlations for packed beds of spheres. *Powder technology*, 283:488–504, 2015.

[230] Jeff Gostick, Mahmoudreza Aghighi, James Hinebaugh, Tom Tranter, Michael A Hoeh, Harold Day, Brennan Spellacy, Mostafa H Sharqawy, Aimy Bazylak, and Alan Burns. Openpnm: a pore network modeling package. *Computing in Science & Engineering*, 18(4):60–74, 2016.

[231] Viatcheslav Berejnov, Madhu Saha, Darija Susac, Juergen Stumper, Marcia West, and Adam P Hitchcock. Advances in structural characterization using soft x-ray scanning transmission microscopy (stxm): Mapping and measuring porosity in pem-fc catalyst layers. *ECS Transactions*, 80(8):241, 2017.

[232] Yurii V Yakovlev, Yevheniia V Lobko, Maryna Vorokhta, Jaroslava Nováková, Michal Mazur, Iva Matolínová, and Vladimir Matolín. Ionomer content effect on charge and gas transport in the cathode catalyst layer of proton-exchange membrane fuel cells. *Journal of Power Sources*, 490:229531, 2021.

[233] Tess Seip, Jason Keonhag Lee, Mingyuan Ge, Wah-Keat Lee, Nima Shaigan, Marius Dinu, Khalid Fatih, and Aimy Bazylak. Stochastic generation of electrolyzer anode catalyst layers. *Electrochimica Acta*, 462:142701, 2023.

[234] Utku U Ince, Henning Markötter, Nan Ge, Merle Klages, Jan Haußmann, Martin Göbel, Joachim Scholta, Aimy Bazylak, and Ingo Manke. 3d classification of polymer electrolyte membrane fuel cell materials from in-situ x-ray tomographic datasets. *International Journal of Hydrogen Energy*, 45(21):12161–12169, 2020.

[235] Martin Göbel, Michael Godehardt, and Katja Schladitz. Multi-scale structural analysis of gas diffusion layers. *Journal of Power Sources*, 355:8–17, 2017.

[236] Isaac C Okereke, Mohammed S Ismail, Derek B Ingham, Kevin Hughes, Lin Ma, and Mohamed Pourkashanian. Single-and double-sided coated gas diffusion layers used in polymer electrolyte fuel cells: a numerical study. *Energies*, 16(11):4363, 2023.

[237] Chujie Lin, Kuangwei Cheng, Shang Li, Siyi Ouyang, Yadong Wang, and Mu Pan. High-flux microporous layers with bimodal pore structures for proton exchange membrane fuel cells. *ACS omega*, 10(19):19966–19973, 2025.

[238] Adnan Ozden, Samaneh Shahgaldi, Jian Zhao, Xianguo Li, and Feridun Hamdullahpur. Assessment of graphene as an alternative microporous layer material for proton exchange membrane fuel cells. *Fuel*, 215:726–734, 2018.

[239] H Markötter, J Haußmann, R Alink, C Tötzke, Tobias Arlt, M Klages, Heinrich Riesemeier, J Scholta, D Gerteisen, and J Banhart. Influence of cracks in the microporous layer on the water distribution in a pem fuel cell investigated by synchrotron radiography. *Electrochemistry Communications*, 34:22–24, 2013.

[240] Celine H Chen, Kaustubh Khedekar, Andrea Zaffora, Monica Santamaria, Matthew Coats, Svitlana Pylypenko, Pablo A García-Salaberri, Jonathan Braaten, Plamen Atanassov, and Nobumichi Tamura. Effect of microporous layer cracks on catalyst durability of polymer electrolyte fuel cells for heavy-duty applications. *ACS Applied Energy Materials*, 7(14):5736–5744, 2024.

[241] Haolin Tang, Shenlong Wang, Mu Pan, and Runzhang Yuan. Porosity-graded microporous layers for polymer electrolyte membrane fuel cells. *Journal of Power Sources*, 166(1):41–46, 2007.

[242] Chongxue Zhao, Haihang Zhang, Zheng Huang, Meng Zhao, Haiming Chen, and Guangyi Lin. Two-stage microporous layers with gradient pore size structure for improving the performance of proton exchange membrane fuel cells. *Polymers*, 15(12):2740, 2023.

[243] André Colliard-Granero, Max Dreger, Ingo Manke, Roswitha Zeis, Joachim Pasel, Tobias Morawietz, Andreas Friederich, Jasna Jankovic, Kourosh Malek, Michael Eikerling, and Mohammad J Eslamibidgoli. Fair infrastructure for imaging and characterization data assets with large language model-aided ontology standardization in materials science. *Journal of Materials Informatics*, submitted, 2025.

[244] Lynn Yarmey and Karen S Baker. Towards standardization: A participatory framework for scientific standard-making. *International Journal of Digital Curation*, 8(1):157–172, 2013.

[245] Paul A Lapchak and John H Zhang. Data standardization and quality management. *Translational Stroke Research*, 9:4–8, 2018.

[246] Seth H Myers and Brett M Huhman. Enabling scientific collaboration and discovery through the use of data standardization. *IEEE Transactions on Plasma Science*, 43(5):1190–1193, 2015.

[247] James L. Spivack, James N. Cawse, Donald W. Whisenhunt, Bruce F. Johnson, Kirill V. Shalyaev, Jonathan Male, Eric J. Pressman, John Y. Ofori, Grigorii L. Soloveichik, Ben P. Patel, Timothy L. Chuck, David J. Smith, Tracey M. Jordan, Michael R. Brennan, Richard J. Kilmer, and Eric D. Williams. Combinatorial discovery of metal co-catalysts for the carbonylation of phenol. *Applied Catalysis A: General*, 254(1):5–25, 2003.

[248] Wilhelm F. Maier, Klaus Stöwe, and Simone Sieg. Combinatorial and high-throughput materials science. *Angewandte Chemie International Edition*, 46(32):6016–6067, 2007.

[249] Hideomi Koinuma and Ichiro Takeuchi. Combinatorial solid-state chemistry of inorganic materials. *Nature Materials*, 3(7):429–438, 2004.

[250] X. D. Xiang, Xiaodong Sun, Gabriel Briceño, Yulin Lou, Kai-An Wang, Hauyee Chang, William G. Wallace-Freedman, Sung-Wei Chen, and Peter G. Schultz. A combinatorial approach to materials discovery. *Science*, 268(5218):1738–1740, 1995.

[251] Lauri Himanen, Amber Geurts, Adam Stuart Foster, and Patrick Rinke. Data-driven materials science: status, challenges, and perspectives. *Advanced Science*, 6(21):1900808, 2019.

[252] Shelley Stall, Lynn Yarmey, Joel Cutcher-Gershenfeld, Brooks Hanson, Kerstin Lehnert, Brian Nosek, Mark Parsons, Erin Robinson, and Lesley Wyborn. Make scientific data fair. *Nature*, 570(7759):27–29, 2019.

[253] Yang Leng. *Materials characterization: introduction to microscopic and spectroscopic methods*. John Wiley & Sons, 2013.

[254] G Bahar Basim and Mohsen Khalili. Particle size analysis on wide size distribution powders; effect of sampling and characterization technique. *Advanced Powder Technology*, 26(1):200–207, 2015.

[255] Keith T Butler, Daniel W Davies, Hugh Cartwright, Olexandr Isayev, and Aron Walsh. Machine learning for molecular and materials science. *Nature*, 559(7715):547–555, 2018.

[256] Maarten De Jong, Wei Chen, Thomas Angsten, Anubhav Jain, Randy Notestine, Anthony Gamst, Marcel Sluiter, Chaitanya Krishna Ande, Sybrand Van Der Zwaag, and Jose J Plata. Charting the complete elastic properties of inorganic crystalline compounds. *Scientific data*, 2(1):1–13, 2015.

[257] Anubhav Jain, Shyue Ping Ong, Geoffroy Hautier, Wei Chen, William Davidson Richards, Stephen Dacek, Shreyas Cholia, Dan Gunter, David Skinner, and Gerbrand Ceder. Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL materials*, 1(1), 2013.

[258] Ben Blaiszik, Kyle Chard, Jim Pruyne, Rachana Ananthakrishnan, Steven Tuecke, and Ian Foster. The materials data facility: data services to advance materials science research. *Jom*, 68(8):2045–2052, 2016.

[259] Nick Romanos, Maritini Kalogerini, Elias P Koumoulos, AK Morozinis, Marco Sebastiani, and Costas Charitidis. Innovative data management in advanced characterization: Implications for materials design. *Materials Today Communications*, 20:100541, 2019.

[260] Ralf F Ziesche, Thomas MM Heenan, Pooja Kumari, Jarrod Williams, Weiqun Li, Matthew E Curd, Timothy L Burnett, Ian Robinson, Dan JL Brett, Matthias J Ehrhardt, et al. Multi-dimensional characterization of battery materials. *Advanced Energy Materials*, 13(23):2300103, 2023.

[261] Renate Hiesgen, Tobias Morawietz, Michael Handl, Martina Corasaniti, and K Andreas Friedrich. Atomic force microscopy on cross sections of fuel cell membranes, electrodes, and membrane electrode assemblies. *Electrochimica Acta*, 162:86–99, 2015.

[262] Nadime Francis, Alastair Green, Paolo Guagliardo, Leonid Libkin, Tobias Lindaaker, Victor Marsault, Stefan Plantikow, Mats Rydberg, Petra Selmer, and Andrés Taylor. Cypher: An evolving query language for property graphs. In *Proceedings of the 2018 international conference on management of data*, pages 1433–1445, 2018.

[263] Justin J Miller. Graph database applications and concepts with neo4j. In *Proceedings of the southern association for information systems conference, Atlanta, GA, USA*, volume 2324, pages 141–147, 2013.

[264] Max Dreger, Kourosh Malek, Mohammed J Eslamibidgoli, and Michael H Eikerling. Synergizing ontologies and graph databases for highly flexible materials-to-device workflow representations. *Journal of Materials Informatics*, 2023.

[265] Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C Berg, Wan-Yen Lo, et al. Segment anything. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 4015–4026, 2023.

[266] Xueyan Zou, Jianwei Yang, Hao Zhang, Feng Li, Linjie Li, Jianfeng Wang, Lijuan Wang, Jianfeng Gao, and Yong Jae Lee. Segment everything everywhere all at once. *Advances in neural information processing systems*, 36:19769–19782, 2023.

[267] Fabian Meyer, Violeta Karyofylli, André Colliard-Granero, Christoph König, Ali Javed, Michael Eikerling, Hans Kungl, Eva Jodat, André Karl, and Rüdiger-A Eichel. Evaluation of gas bubble coverage in a flow field by deep learning-based analysis: The impact of gas content on pem electrolyzer performance. *Journal of The Electrochemical Society*, 172(8):084501, 2025.

List of conference talks

- André Colliard-Granero, Michael Eikerling, Kourosh Malek, and Mohammad J Eslamibidgoli. Ontology Development for Efficient Data Management of Energy Materials. *34th Topical Meeting of the International Society of Electrochemistry (ISE)*, 2023. (Mar del Plata, Argentina)
- André Colliard-Granero, Michael Eikerling, Kourosh Malek, and Mohammad J Eslamibidgoli. UTILE: A Deep Learning-Driven Imaging Journey Across Dimensions. *Helmholtz Imaging Conference*, 2024. (Heidelberg, Germany)
- André Colliard-Granero, Michael Eikerling, Kourosh Malek, and Mohammad J Eslamibidgoli. Deep Learning-driven Autonomous Imaging Characterization of Polymer Electrolyte Membrane Fuel Cell Technologies.*Pacific Rim International Meeting on Electrochemical and Solid State Science (PRiME)*, 2024. (Honolulu, USA)

List of publications

Contributions to the following publications published/submitted so far were made in the time frame of the PhD Project.

- André Colliard-Granero, Jenia Jitsev, Michael H Eikerling, Kourosh Malek, and Mohammad J Eslamibidgoli. Utile-gen: Automated image analysis in nanoscience using synthetic dataset generator and deep learning. *ACS Nanoscience Au*, page 398–407, 2023.

- André Colliard-Granero, Keusra Armel Gompou, Christian Rodenbücher, Kourosh Malek, Michael Eikerling, and Mohammad J Eslamibidgoli. Deep learning-enhanced characterization of bubble dynamics in proton exchange membrane water electrolyzers. *Physical Chemistry Chemical Physics*, 26:14529, 2024.
- André Colliard-Granero, Kangjun Duan, Roswitha Zeis, Kourosh Malek, Michael Eikerling, and Mohammad J Eslamibidgoli. Advancing vanadium redox flow battery analysis: A deep learning approach for high-throughput 3d visualization and bubble quantification. *Digital Discoveries*, accepted, 2025.
- André Colliard-Granero, Salvatore Ranieri, Aimy Bazylak, Tobias Morawietz, K Andreas Friedrich, Jasna Jankovic, Michael H Eikerling, Kourosh Malek, and Mohammad J Eslamibidgoli. Utile-pore: Deep learning-enabled 3d analysis of porous materials in polymer electrolyte membrane-based energy devices. *Journal of The Electrochemical Society*, 172(7):074515, 2025.
- André Colliard-Granero, Max Dreger, Ingo Manke, Roswitha Zeis, Joachim Pasel, Tobias Morawietz, Andreas Friederich, Jasna Jankovic, Kourosh Malek, Michael Eikerling, and Mohammad J Eslamibidgoli. Fair infrastructure for imaging and characterization data assets with large language model-aided ontology standardization in materials science. *Journal of Materials Informatics*, submitted, 2025.
- Fabian Meyer, Violeta Karyofylli, André Colliard-Granero, Christoph König, Ali Javed, Michael Eikerling, Hans Kungl, Eva Jodat, André Karl, and Rüdiger-A Eichel. Evaluation of gas bubble coverage in a flow field by deep learning-based analysis: The impact of gas content on pem electrolyzer performance. *Journal of The Electrochemical Society*, 172(8):084501, 2025.

Band / Volume 696
Nanocrystalline Silicon Carbide in Transparent Passivating Contact Solar Cells
A. Eberst (2026), xiii, 225 pp
ISBN: 978-3-95806-891-9

Band / Volume 697
Theory of Electronic and Ionic Perturbations at Supported Electrocatalyst Nanoparticles
Y. Zhang (2026), XI, 131 pp
ISBN: 978-3-95806-896-4

Band / Volume 698
Aufbau und Einsatz eines on-board Messsystems zur Untersuchung der Abgaszusammensetzung von Fahrzeugen betrieben mit konventionellen und alternativen Kraftstoffen
V. Polinowski (2026), VIII, 269 pp
ISBN: 978-3-95806-897-1

Band / Volume 699
Ab Initio-based large-scale Atomistic Simulations of Cathode Materials for Secondary Batteries: From Computational Methodologies to Applications towards improved Structural and Chemical Stability
K. Köster (2026), x, 204, x pp
ISBN: 978-3-95806-898-8

Band / Volume 700
Materials Design, Processing and Application of Proton-Conducting Oxides for Electrochemical Energy Conversion
Y. Zeng (2026), vii, 117 pp
ISBN: 978-3-95806-899-5

Band / Volume 701
Accelerating the discovery of alkaline-stable anion exchange membrane materials via computational exploration
F. P. Tipp (2026), xii, 135 pp
ISBN: 978-3-95806-900-8

Schriften des Forschungszentrums Jülich
Reihe Energie & Umwelt / Energy & Environment

Band / Volume 702
Methods for Investigating the Structure-Performance Correlation in Membrane Electrode Assemblies
N. Utsch (2026), XIII, 138, LXVII pp
ISBN: 978-3-95806-901-5

Band / Volume 703
Hot-Spot Formation in Cu(In,Ga)Se_2 Thin Film Solar Cells
S. Nofal (2026), xvii, 131 pp
ISBN: 978-3-95806-903-9

Band / Volume 704
Machine Learning Guidance of Manufacturing Process of Solid Oxide Cells
T. Le Dinh (2026), xxi, 122 pp
ISBN: 978-3-95806-904-6

Band / Volume 705
Drivers of spatiotemporal variability of European terrestrial ecosystem processes
C. Poppe Téran (2026), 254 pp
ISBN: 978-3-95806-905-3

Band / Volume 706
Characterization of Nitrogen Reduction Reaction Catalysts for the electrochemical Ammonia Synthesis
D. Sachse (2026), xi, 235, 6 pp
ISBN: 978-3-95806-907-7

Band / Volume 707
Autonomous Image Analysis to Accelerate the Discovery and Integration of Energy Materials
A. Colliard Granero (2026), xiv, 156 pp
ISBN: 978-3-95806-910-7 (Print)
ISBN: 978-3-95806-911-4 (E-Book)

Weitere ***Schriften des Verlags im Forschungszentrum Jülich*** unter
http://wwwzb1.fz-juelich.de/verlagextern1/index.asp

Energie & Umwelt / Energy & Environment
Band / Volume 707
ISBN 978-3-95806-911-4

Mitglied der Helmholtz-Gemeinschaft

JÜLICH
Forschungszentrum