AI Meets Chemistry: Unlocking New Frontiers in Molecular Design and Reaction Prediction

Amit Kumar Mishra[1], **Yonglan Kan**[2,*]

[1]Aalborg university Denmark
[2]Han Bao Lu Co. Ltd.
Corresponding author: Yonglan Kan.
E-mail: yl_kan@163.com.

https://doi.org/10.63619/ijais.v1i1.005

Manuscript received December 23, 2025; revised February 19, 2025. published March 17, 2025.

Abstract: Artificial intelligence (AI) is revolutionising chemical research by significantly enhancing automated experimental processes, reaction prediction, and molecular design. Despite these advances, problems with data quality, computational resource limitations, and model interpretability persist. In order to further speed up chemical discoveries and advances, prospects include creating hybrid AI models, quantum AI, and multimodal frameworks. Using artificial intelligence (AI) to significantly improve automated experimental procedures, reaction prediction, and molecular design is revolutionising chemical research. Creating hybrid AI models, quantum AI and multimodal frameworks is a potential future avenue to accelerate chemical discoveries and further advances. Chemical research is being revolutionised by artificial intelligence (AI), which is significantly enhancing automated experimental procedures, reaction prediction, and molecular design. Recent advances in generative AI methods, such as diffusion models, GANs, and variational autoencoders (VAEs) that aid in creating unique molecular structures, are the main focus of this review effort. To increase the precision of reaction predictions, transformer-based designs and graph neural networks (GNNs) are being investigated. Some challenges remain, including low-quality data, a lack of processing capacity, and concerns over the model's interpretability. The creation of hybrid AI models, quantum AI, and multimodal frameworks, among other exciting study topics, could accelerate future developments in chemistry.

Keywords: Artificial Intelligence, Molecular Design, Reaction Prediction, Automated Experimentation, Generative Models, Graph Neural Networks, Quantum AI

1. Introduction

1.1. The Rise of AI in Chemistry

Artificial intelligence (AI) has rapidly emerged as a transformative technology across various scientific disciplines, profoundly reshaping traditional practices in chemistry. Historically, computational chemistry has relied heavily on techniques such as quantum mechanical calculations, molecular dynamics simulations, and density functional theory (DFT) to explore molecular structures, reaction mechanisms, and material properties [1]. While these traditional computational methods offer precise theoretical insights, they encounter significant challenges, including limited scalability, high computational costs, and difficulties in accurately modeling complex, real-world chemical systems. Consequently, these limitations restrict extensive explorations of chemical space and limit practical applicability to relatively small or simplified systems [2], [3].

Recent advancements in machine learning (ML) and deep learning (DL) have introduced powerful, data-driven alternatives that overcome some of these inherent limitations. Large datasets may be used by

AI techniques to find hidden patterns, predict molecule attributes effectively, and expedite the chemical discovery process. Rapid exploration of chemical spaces that were previously thought to be computationally prohibitive is now possible because to ML and DL approaches, which have also dramatically improved predicted accuracy and decreased computing costs [4], [5].

DeepMind's AlphaFold, which has transformed protein structure prediction, serves as an illustration of the revolutionary influence of AI. AlphaFold has demonstrated the usefulness of artificial intelligence in chemistry by dramatically speeding up structure-based drug development by correctly predicting protein structures from the sequences of amino acids.

1.2. Challenges in Molecular Design and Reaction Prediction

Despite substantial advances, practical deployment of AI in molecular design and reaction prediction continues to face critical challenges. Molecular design often necessitates optimizing specific chemical properties or biological activities, tasks that inherently involve navigating vast and complex chemical spaces. Additionally, reaction prediction remains particularly challenging due to the intrinsic complexity of chemical reaction mechanisms and reaction dynamics, which demand precise and highly interpretable AI-driven models. The accuracy of AI predictions strongly depends on the quality, comprehensiveness, and representativeness of chemical datasets, currently limited by experimental inconsistencies and incomplete chemical databases [6], [7].

For example, the development of new pharmaceuticals often requires the optimization of multiple properties, such as bioavailability, toxicity, and efficacy. Navigating this multi-dimensional optimization landscape is a significant challenge that AI can help address, but it requires high-quality, comprehensive datasets to be effective.

1.3. Potential Breakthroughs Through AI

Generative AI approaches such as variational autoencoders (VAEs), generative adversarial networks (GANs), and diffusion models have shown exceptional capabilities for molecular generation. These models effectively address combinatorial complexities inherent in exploring chemical space, generating novel and chemically meaningful molecular structures with desired properties. These innovative methods promise significant breakthroughs by accelerating discovery processes and facilitating the rapid identification of promising chemical candidates [8], [9], [10].

Moreover, hybrid AI models integrating data-driven methodologies with domain-specific chemical knowledge represent an increasingly critical area of research. Such approaches aim to mitigate data limitations and enhance model interpretability, combining the predictive power of AI with chemical intuition and theoretical principles. Significant attention is also being directed toward algorithmic innovations that reduce computational burdens, such as active learning strategies and optimized computational frameworks [11], [12].

For instance, the integration of AI with quantum computing, known as quantum AI, holds the potential to revolutionize molecular simulations and reaction mechanism elucidation. Quantum AI can simulate complex chemical phenomena that are currently beyond the reach of classical computational methods, offering unprecedented insights into chemical processes.

1.4. Objectives and Contributions of This Review

In this review, we provide a comprehensive analysis of recent breakthroughs and state-of-the-art applications of AI in chemistry, with an explicit focus on molecular design and chemical reaction prediction. We critically evaluate the effectiveness, strengths, and limitations of existing methodologies, highlight prominent success stories, and discuss ongoing challenges and promising future directions. Through this synthesis, we aim to clarify the current state of AI-driven chemistry research and highlight avenues for future exploration, thereby supporting continued innovation and interdisciplinary collaboration.

2. AI Breakthroughs in Molecular Design

2.1. Generative AI for Molecular Generation

Generative artificial intelligence (AI) methodologies have significantly reshaped molecular design by enabling efficient exploration of chemical space through novel molecular structure generation. Variational Autoencoders (VAEs) are prominent examples, utilizing probabilistic encoding-decoding architectures to represent molecular data in continuous latent spaces. This approach enables efficient optimization of molecular properties, such as bioactivity, solubility, and stability, while ensuring chemical validity [8], [13]. Generative Adversarial Networks (GANs) further enhance this capability through adversarial training, facilitating targeted molecular property optimization and diversity enhancement, exemplified by methods like Objective-Reinforced GANs (ORGAN) [9].

For instance, VAEs have been successfully applied to generate novel drug-like molecules with optimized properties, significantly accelerating the drug discovery process. Similarly, GANs have been used to create diverse molecular libraries, enhancing the exploration of chemical space and identifying potential candidates for further development.

2.2. Diffusion Models and Advanced Generative Techniques

Diffusion models have recently emerged as powerful generative frameworks, leveraging iterative refinement processes to convert random noise into structured molecular entities. These methods exhibit remarkable performance, surpassing traditional generative models in capturing complex molecular features and producing diverse chemical compounds, representing a state-of-the-art paradigm in molecular generation [10], [14].

An example of the practical application of diffusion models is their use in generating complex organic molecules with specific functional groups, which are crucial for various industrial applications. These models have demonstrated superior performance in generating chemically valid and diverse structures compared to traditional methods.

2.3. Reinforcement Learning for Molecule Optimization

Variational reinforcement learning integrates generative AI with reinforcement learning (RL) techniques, offering robust tools for property-driven molecular optimization. By combining generative flexibility with the decision-making capabilities of RL, these approaches systematically guide molecular synthesis toward predefined optimal properties, significantly enhancing molecule discovery efficiency and precision [15], [16].

Reinforcement learning has been applied to optimize the synthesis pathways of complex molecules, reducing the number of steps and improving overall yield. This approach has proven particularly effective in optimizing catalytic processes, where precise control over reaction conditions is essential.

2.4. Graph-based Models for Property Prediction

Advancements in machine learning-driven property prediction, particularly through graph neural networks (GNNs), have revolutionized traditional Quantitative Structure-Activity Relationship (QSAR) approaches. GNNs treat molecules as graphs, automating feature extraction directly from molecular structures. This methodology significantly improves predictive accuracy, allowing chemists to more effectively predict complex molecular behaviors and biological activities [17], [18].

For instance, GNNs have been used to predict the toxicity and bioactivity of new chemical compounds, providing valuable insights during the early stages of drug development. These models have outperformed traditional QSAR methods, offering more accurate and reliable predictions.

2.5. AI-driven Drug Discovery Applications

AI-driven breakthroughs, notably DeepMind's AlphaFold and BioGPT, illustrate the practical success of AI in drug discovery. AlphaFold employs deep neural networks to accurately predict protein structures from amino acid sequences, significantly accelerating structure-based drug discovery. Similarly, BioGPT, a generative pre-trained transformer tailored to biochemical literature, streamlines information extraction and hypothesis generation, substantially enhancing early-stage drug discovery processes [19], [20].

For example, AlphaFold has been instrumental in identifying potential drug targets by providing accurate protein structure predictions, which are critical for understanding disease mechanisms. BioGPT has facilitated the extraction of relevant information from vast biochemical literature, aiding researchers in generating new hypotheses and accelerating the drug discovery process.

These AI methodologies have propelled molecular design into a new era, providing chemists with sophisticated computational tools capable of rapidly, systematically, and innovatively exploring chemical spaces.

3. AI Breakthroughs in Chemical Reaction Prediction

3.1. Transformer and Graph-based Reaction Prediction Models

AI has notably advanced chemical reaction prediction through Transformer-based and graph-based models. The Molecular Transformer utilizes self-attention mechanisms, treating chemical reactions as sequential data, achieving significant improvements in reaction outcome predictions over traditional rule-based methods [21], [22]. Graph-based models complement these sequence approaches by explicitly representing molecular structures as graphs, thereby capturing complex molecular interactions and efficiently predicting reaction pathways [23].

The Molecular Transformer has been successfully applied to predict the outcomes of organic reactions, significantly reducing the time and resources required for experimental validation. Graph-based models have superior performance in predicting reaction mechanisms, providing chemists valuable insights into complex molecular interactions.

3.2. AI for Reaction Mechanism Understanding and Synthesis Planning

Machine learning algorithms significantly contribute to the classification of reactions, elucidation of chemical reaction mechanisms, and prediction of synthetic routes. Techniques combining symbolic AI with deep learning facilitate advanced retrosynthetic analysis, substantially reducing resource demands associated with synthesizing complex molecules [24], [7].

AI-driven retrosynthetic analysis has been used to design efficient synthetic routes for pharmaceuticals, minimizing the number of steps and improving overall yield. This approach has proven particularly effective in the synthesis of complex organic compounds, where traditional methods often fall short.

3.3. AI-assisted Catalyst Design

Machine learning has also become pivotal in catalyst discovery and optimization. Integrating AI-driven surrogate models with density functional theory (DFT) enables efficient computational screening and property optimization of catalytic materials. Active learning further enhances this process, strategically guiding computational resources and experimental efforts toward promising catalytic candidates [25], [12].

An example of AI-assisted catalyst design is using machine learning models to predict the activity and stability of new catalytic materials, accelerating the discovery of efficient catalysts for industrial processes. Active learning algorithms have been employed to optimize reaction conditions, significantly reducing the time and cost of experimental trials.

3.4. Advanced Reaction Pathway Exploration

Advanced AI algorithms, including Monte Carlo Tree Search (MCTS) and Generative Flow Networks (GFlowNet), significantly enhance reaction pathway discovery and optimization. MCTS effectively balances exploration and exploitation to uncover viable reaction pathways systematically. GFlowNet integrates generative modeling and reinforcement learning, probabilistically sampling reaction pathways and efficiently identifying optimal synthetic routes within complex reaction spaces [26], [27].

MCTS has been applied to explore reaction pathways for synthesizing complex organic molecules, providing chemists with a systematic approach to identify viable synthetic routes. GFlowNet has demonstrated superior performance in optimizing reaction pathways, offering a robust tool for navigating complex chemical spaces.

Together, these innovative AI approaches significantly expand chemists' capabilities in reaction prediction, offering robust, scalable, and accurate tools for addressing critical challenges in chemical synthesis and analysis.

4. AI-Enhanced Automated Chemical Experimentation

4.1. Robotic Laboratories and Automated Synthesis

Recent advances in artificial intelligence (AI) have significantly transformed automated chemical experimentation, primarily through the integration of robotic laboratory systems. These robotic platforms perform chemical synthesis tasks with high precision and reproducibility, effectively minimizing human error and enhancing laboratory efficiency. AI integration enables dynamic decision-making and process optimization, where intelligent algorithms control robotics to autonomously execute complex chemical syntheses. Bayesian optimization methods have become particularly influential, systematically guiding these automated systems toward optimal reaction conditions by intelligently balancing exploration of new chemical spaces with the exploitation of known successful experimental setups [28], [29], [12].

Robotic laboratory systems have been deployed to conduct high-throughput screening of chemical reactions, significantly accelerating the discovery of new compounds. These systems can autonomously adjust reaction parameters in real time, ensuring optimal conditions are maintained throughout the synthesis process. AI-driven optimization techniques have led to the identification of novel reaction pathways and efficient synthesization of complex molecules.

4.2. Closed-loop Autonomous Discovery

AI-driven closed-loop discovery platforms represent an essential breakthrough in chemical experimentation. Such autonomous systems integrate real-time data collection, analysis, and feedback mechanisms into an iterative experimental loop, dynamically adjusting conditions to rapidly achieve optimized outcomes. These frameworks autonomously adjust experimental parameters in response to continuous data input, quickly identifying optimal conditions or novel synthetic routes. The closed-loop approach significantly enhances discovery efficiency across diverse chemical fields, including nanoparticle synthesis, catalytic system optimization, and novel material exploration, demonstrating a profound impact on reducing discovery timelines and resource consumption [7], [30].

Closed-loop systems have been instrumental in optimizing the synthesis of nanoparticles, where precise control over reaction conditions is crucial. By continuously monitoring and adjusting parameters, these systems ensure the production of high-quality nanoparticles with desired properties. The integration of AI in these platforms has also facilitated the rapid exploration of catalytic systems, leading to the discovery of highly efficient catalysts for various chemical reactions.

4.3. Active Learning for Data-driven Experimental Optimization

Data-driven experimental optimization through active learning methodologies further enhances the effectiveness of automated chemical experimentation. Active learning algorithms strategically select experiments based on predicted uncertainties and information gain, identifying conditions that yield the highest potential improvement in predictive model performance. This approach substantially reduces the number of necessary experiments, minimizing resource use and accelerating the optimization process in diverse chemical applications, such as catalyst discovery, formulation design, and reaction parameter optimization. By systematically selecting the most informative experimental conditions, active learning significantly boosts experimental efficiency and predictive model accuracy [12], [31].

Active learning has been applied to optimize reaction parameters in catalytic processes, where selecting optimal conditions is critical for achieving high yields. By focusing on the most informative experiments, researchers can quickly identify the best reaction conditions, reducing the time and resources required for experimental trials. This approach has also been used in formulation design, where optimizing multiple parameters is necessary to achieve the desired product properties.

Integrating AI methodologies—including robotic laboratories, autonomous closed-loop discovery, and active learning—redefines chemical experimentation, promising faster, more accurate, and resource-efficient chemical innovations.

5. Challenges and Future Perspectives

5.1. Data Quality and Integration Challenges

A significant challenge in the practical application of artificial intelligence (AI) in chemistry is ensuring high-quality data integration. Chemical datasets commonly exhibit inconsistencies, incompleteness, and biases arising from variability in experimental conditions, measurement inaccuracies, and heterogeneous reporting standards. Such deficiencies severely impact the robustness and accuracy of AI models, limiting their effectiveness in real-world chemical discovery. Addressing these challenges requires hybrid AI models that effectively integrate chemical domain knowledge with data-driven techniques, thereby enhancing predictive reliability and interpretability. Efforts to standardize chemical data and improve data quality through rigorous data curation protocols are equally essential to overcome these limitations [3], [6].

For instance, the development of standardized protocols for data collection and reporting can significantly improve the quality and consistency of chemical datasets. Hybrid AI models that combine data-driven approaches with chemical expertise can enhance the interpretability and reliability of predictions, making them more applicable to real-world scenarios.

5.2. Computational Efficiency and Explainability

Another critical barrier is the computational intensity associated with state-of-the-art AI methodologies, particularly deep learning models. Techniques such as graph neural networks (GNNs) and transformer-based models demand substantial computational resources for training and deployment, limiting accessibility and scalability, particularly for smaller laboratories and institutions. Moreover, the opaque nature of complex AI models, often referred to as the ”black box” problem, hinders their widespread acceptance in chemical research. Developing computationally efficient algorithms, alongside explainable AI frameworks, is therefore pivotal. These advancements will not only improve accessibility but also enhance trust and facilitate broader adoption by clearly elucidating model predictions [2], [32].

Efforts to develop more computationally efficient algorithms can reduce the resource demands associated with training and deploying AI models, making them more accessible to smaller research groups. Explainable AI frameworks that provide insights into model predictions can enhance trust and facilitate broader adoption in the chemical research community.

5.3. Emerging Research Directions: Quantum AI and Multimodal Models

Future developments in AI for chemistry hold substantial promise, particularly in emerging areas such as quantum AI and multimodal modeling frameworks. Quantum AI combines quantum computing's computational strengths with AI algorithms, offering powerful capabilities for simulating complex chemical phenomena that currently surpass classical computational limits. This synergy may significantly accelerate molecular simulations, reaction mechanism elucidation, and catalyst design processes [33], [34]. Additionally, multimodal AI models, integrating diverse chemical data sources—including experimental outcomes, spectroscopic information, molecular structures, and textual databases—are set to dramatically expand AI capabilities in chemistry. Such integrative models promise more comprehensive and robust representations, improving prediction accuracy and fostering deeper insights into chemical phenomena [5], [7].

Quantum AI can revolutionize molecular simulations and reaction mechanism elucidation by leveraging the computational strengths of quantum computing. Multimodal AI models integrating diverse data sources can provide more comprehensive and accurate predictions, enhancing our understanding of complex chemical phenomena.

5.4. Future Outlook

The continued advancement of AI methodologies, supported by improvements in data integration, computational efficiency, and model interpretability, promises to reshape the chemical sciences profoundly. The integration of emerging technologies such as quantum computing and multimodal data approaches will further accelerate discoveries, reduce experimental resource demands, and enhance predictive capabilities, paving the way toward unprecedented scientific innovations in chemistry.

The integration of quantum computing and multimodal data approaches can significantly accelerate chemical discoveries and reduce resource demands, paving the way for groundbreaking innovations in the field. Continued advancements in AI methodologies will enhance predictive capabilities and foster deeper insights into chemical phenomena.

6. Conclusion

6.1. Summary of AI Advances in Chemistry

Artificial intelligence (AI) has dramatically transformed chemistry by significantly advancing molecular design, chemical reaction prediction, and automated experimentation. Generative AI models, such as variational autoencoders (VAEs), generative adversarial networks (GANs), and diffusion models, have enabled efficient exploration and discovery of novel molecular structures, streamlining chemical space navigation and reducing development timelines. Moreover, advancements in graph neural networks (GNNs) and transformer-based models have significantly improved prediction accuracy, enhancing our understanding of complex chemical reactions and mechanisms.

Integrating AI methodologies in robotic laboratories and closed-loop discovery platforms has revolutionized automated chemical experimentation, ensuring high precision and reproducibility in chemical synthesis. Active learning algorithms have further optimized experimental processes, reducing resource use and accelerating discovery timelines.

6.2. Addressing Existing Challenges

Despite these successes, persistent challenges remain, notably in the areas of data quality, computational resource limitations, and AI model interpretability. Addressing these issues through hybrid AI approaches that blend chemical domain knowledge with data-driven insights is critical. Innovations aimed at computational efficiency and the development of explainable AI are essential to overcome existing barriers and ensure widespread adoption within the chemical research community.

Efforts to standardize chemical data and improve data quality through rigorous data curation protocols are essential to enhance the robustness and accuracy of AI models. Developing computationally efficient algorithms and explainable AI frameworks will improve accessibility and trust, facilitating broader adoption of AI methodologies in chemical research.

6.3. Future Impact and Opportunities

Emerging research areas such as quantum AI and multimodal AI frameworks hold great promise for future advancements. These novel approaches could dramatically enhance computational capabilities and improve prediction comprehensiveness, potentially revolutionizing chemical research and discovery processes. Ultimately, AI technologies' continuous integration and evolution promise to accelerate scientific discovery and foster more profound, impactful chemical innovations across various disciplines.

Quantum AI can significantly accelerate molecular simulations and reaction mechanism elucidation, offering unprecedented insights into chemical processes. Multimodal AI models integrating diverse data sources will enhance prediction accuracy and foster a deeper understanding of complex chemical phenomena.

The integration of AI methodologies in practical applications, such as drug discovery and catalyst design, has already demonstrated substantial benefits. Continued advancements in these areas will further enhance the efficiency and precision of chemical research, paving the way for groundbreaking innovations.

References

[1] F. Jensen, "Introduction to computational chemistry," *John Wiley & Sons,*.

[2] M. R. Islam, M. S. Islam, S. Majumder, and S. Fathi Hafshejani, "Breast cancer prediction: A fusion of genetic algorithm, chemical reaction optimization, and machine learning techniques," *Applied Computational Intelligence & Soft Computing*, vol. 2024, 2024.

[3] J. A. Keith, V. V. Vassilev-Galindo, B. Cheng, S. Chmiela, M. Gastegger, K.-R. Müller, and A. Tkatchenko, "Combining machine learning and computational chemistry for predictive insights into chemical systems." 2021.

[4] Butler, Keith, T., Davies, Daniel, W., Cartwright, Hugh, Isayev, and Olexandr, "Machine learning for molecular and materials science," *Nature*, 2018.

[5] N. Ye, J. An, and J. Yu, "Deep-learning-enhanced noma transceiver design for massive mtc: Challenges, state of the art, and future directions," *IEEE wireless communications*, no. 28-4, 2021.

[6] A. Nikouline, J. Feng, F. Rudzicz, A. Nathens, and B. Nolan, "Machine learning in the prediction of massive transfusion in trauma: a retrospective analysis as a proof-of-concept," *European Journal of Trauma and Emergency Surgery*, vol. 50, no. 3, pp. 1073–1081, 2024.

[7] C. W. Coley, N. S. Eyke, and K. F. Jensen, "Autonomous discovery in the chemical sciences part ii: Outlook," 2020.

[8] R. Gómez-Bombarelli, J. N. Wei, D. Duvenaud, J. M. Hernández-Lobato, B. Sánchez-Lengeling, D. Sheberla, J. Aguilera-Iparraguirre, T. D. Hirzel, R. P. Adams, and A. Aspuru-Guzik, "Automatic chemical design using a data-driven continuous representation of molecules," *ACS Central Science*, vol. 4, no. 2, pp. 268–276, 2018.

[9] G. L. Guimaraes, B. Sanchez-Lengeling, C. Outeiral, P. L. C. Farias, and A. Aspuru-Guzik, "Objective-reinforced generative adversarial networks (organ) for sequence generation models," 2017.

[10] E. Hoogeboom, V. G. Satorras, C. Vignac, and M. Welling, "Equivariant diffusion for molecule generation in 3d," 2022.

[11] C. Chen, D. T. Nguyen, S. J. Lee, N. A. Baker, A. S. Karakoti, L. Lauw, C. Owen, K. T. Mueller, B. A. Bilodeau, and V. Murugesan, "Accelerating computational materials discovery with machine learning and cloud high-performance computing: from large-scale screening to experimental validation," *Journal of the American Chemical Society*, vol. 146, no. 29, p. 10, 2024.

[12] E. Claes, T. Heck, K. Coddens, M. Sonnaert, J. Schrooten, and J. Verwaeren, "Bayesian cell therapy process optimization," *Biotechnology and Bioengineering*, vol. 121, no. 5, p. 14, 2024.

[13] M. Simonovsky and N. Komodakis, "Graphvae: Towards generation of small graphs using variational autoencoders," *Springer, Cham*, 2018.

[14] M. Xu, L. Yu, Y. Song, C. Shi, S. Ermon, and J. Tang, "Geodiff: a geometric diffusion model for molecular conformation generation," 2022.

[15] Z. Zhou, S. Kearnes, L. Li, R. N. Zare, and P. Riley, "Optimization of molecules via deep reinforcement learning," 2018.

[16] D. Neil, M. Segler, L. Guasch, M. Ahmed, D. Plumbley, M. Sellwood, and N. Brown, "Exploring deep recurrent models with reinforcement learning for molecule design," 2018.

[17] J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl, "Neural message passing for quantum chemistry," 2017.

[18] K. Yang, K. Swanson, W. Jin, C. W. Coley, and R. Barzilay, "Analyzing learned molecular representations for property prediction," *Journal of Chemical Information and Modeling*, vol. 59, no. 8, 2019.

[19] J. Jumper, R. Evans, A. Pritzel, T. Green, and D. Hassabis, "Highly accurate protein structure prediction with alphafold," *Nature*, pp. 1–11, 2021.

[20] S. Hussain, U. Naseem, M. Ali, D. B. Avendao Avalos, S. Cardona-Huerta, B. A. Bosques Palomo, and J. G. Tamez-Pea, "Tecrr: a benchmark dataset of radiological reports for bi-rads classification with machine learning, deep learning, and large language model baselines," *BMC Medical Informatics and Decision Making*, vol. 24, no. 1, pp. 1–10, 2024.

[21] P. Schwaller, T. Laino, T. Gaudin, P. Bolgar, and A. A. Lee, "Molecular transformer: A model for uncertainty-calibrated chemical reaction prediction," *ACS Central Science*, vol. 5, no. 9, 2019.

[22] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," *arXiv*, 2017.

[23] J. Kirman, A. Johnston, D. A. Kuntz, M. Askerka, Y. Gao, P. Todorovi, D. Ma, G. G. Privé, and E. H. Sargent, "Machine-learning-accelerated perovskite crystallization," *Matter*, vol. 2, no. 4, 2020.

[24] Z. Yu, J. Wang, X. Yang, and J. Ma, "Superpixel-based style transfer method for single-temporal remote sensing image identification in forest type groups," *Remote Sensing*, vol. 15, no. 15, p. 3875, 2023.

[25] Z. W. Ulissi, M. T. Tang, J. Xiao, X. Liu, D. A. Torelli, M. Karamad, K. Cummins, C. Hahn, N. S. Lewis, and T. F. Jaramillo, "Machine-learning methods enable exhaustive searches for active bimetallic facets and reveal active site motifs for co2 reduction," *ACS Catalysis*, p. acscatal.7b01648, 2017.

[26] E. Bengio, M. Jain, M. Korablyov, D. Precup, and Y. Bengio, "Flow network based generative models for non-iterative diverse candidate generation," 2021.

[27] Z. Yu and C. S. Chan, "Yuan: Yielding unblemished aesthetics through a unified network for visual imperfections removal in generated images," *arXiv preprint arXiv:2501.08505*, 2025.

[28] B. Burger, P. M. Maffettone, V. V. Gusev, C. M. Aitchison, Y. Bai, X. Wang, X. Li, B. M. Alston, B. Li, and R. Clowes, "A mobile robotic chemist," *Nature*.

[29] P. Patel, "Ai chemist performs complex experiments based on plain text prompts," *Chemical & Engineering News*, vol. 102, no. 1, p. 1, 2024.

[30] L. M. Roch, F. Hse, C. Kreisbeck, T. Tamayo-Mendoza, L. P. E. Yunker, J. E. Hein, and A. Aspuru-Guzik, "Chemos: An orchestration software to democratize autonomous discovery," *PLOS ONE*, vol. 15, 2020.

[31] T. Lookman, P. V. Balachandran, D. Xue, and R. Yuan, "Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design," *Computational Materials Science (English)*, no. 1, p. 17, 2019.

[32] J. Hernandez-Orallo, "Explainable ai: interpreting, explaining and visualizing deep learning," *Computing reviews*, no. 1, p. 62, 2021.

[33] Y. Cao, J. Romero, J. P. Olson, M. Degroote, and A. Aspuru-Guzik, "Quantum chemistry in the age of quantum computing," *Chemical Reviews*, vol. 119, no. 19, 2019.

[34] B. Bauer, S. Bravyi, M. Motta, and K. L. Chan, "Quantum algorithms for quantum chemistry and quantum materials science," 2020.