PAPER

Check for updates

Accelerated screening of colloidal nanocrystals using artificial neural network-assisted autonomous flow reactor technology†

Ajit Vikram, [ID] [a] Ken Brudnak,[a] Arwa Zahid,[a] Moonsub Shim [ID] [b] and
Paul J. A. Kenis [ID] *[a]

Colloidal semiconductor nanocrystals with tunable optical and electronic properties are opening up exciting opportunities for high-performance optoelectronics, photovoltaics, and bioimaging applications. Identifying the optimal synthesis conditions and screening of synthesis recipes in search of efficient synthesis pathways to obtain nanocrystals with desired optoelectronic properties, however, remains one of the major bottlenecks for accelerated discovery of colloidal nanocrystals. Conventional strategies, often guided by limited understanding of the underlying mechanisms remain expensive in both time and resources, thus significantly impeding the overall discovery process. In response, an autonomous experimentation platform is presented as a viable approach for accelerated synthesis screening and optimization of colloidal nanocrystals. Using a machine-learning-based predictive synthesis approach, integrated with automated flow reactor and inline spectroscopy, indium phosphide nanocrystals are autonomously synthesized. Their polydispersity for different target absorption wavelengths across the visible spectrum is simultaneously optimized during the autonomous experimentation, while utilizing minimal self-driven experiments (less than 50 experiments within 2 days). Starting with no-prior-knowledge of the synthesis, an ensemble neural network is trained through autonomous experiments to accurately predict the reaction outcome across the entire synthesis parameter space. The predicted parameter space map also provides new nucleation-growth kinetic insights to achieve high monodispersity in size of colloidal nanocrystals.

Introduction

Colloidal semiconductor nanomaterials hold potential to control optical and electronic properties with unprecedented precision through tuning of their size, shape, and morphology, thus enabling exciting opportunities for high-performance electronics, optoelectronics, and photovoltaics.[1–5] Quantum dot (QD) synthesis methodologies have progressed substantially in the past two decades. Various synthetic approaches have been developed for the synthesis of QDs, ranging from liquid-phase methods to vapor phase epitaxial growth.[6] Among the various compositions, cadmium based colloidal QDs have received the most attention, resulting in monodispersed nanocrystals with photoluminescence quantum yields close to

unity.[7–11] Despite their attractive properties, the use of Cd-based QDs for applications is largely limited due to the inherent toxicity of heavy metals. Consequently, significant efforts have focused on the discovery of efficient synthesis strategies for heavy-metal-free alternatives such as InP, ZnTe, and ZnSe-based nanocrystals.[6,12] While substantial improvement in the optical properties of these heavy-metal-free compositions has been made over the past decade, a major bottleneck hindering their discovery and optimization remains due to the trial-and-error-based methodologies used for synthesis design and optimization. This trial-and error approach impedes the overall discovery process involving the identification of the target materials and the subsequent development of synthesis pathways to achieve optimal optical properties.

Conventional flask-based batch reactor platforms for synthesis of colloidal QDs often require a large number of experiments to explore vast reaction space before a synthesis recipe is accepted or abandoned. Although, the thermodynamic and kinetic understanding of the nucleation-growth processes can provide helpful insights, they are often very complex and poorly understood to provide a robust framework for optimization of semiconductor nanocrystal synthesis parameters. As a

[a]*Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA. E-mail: kenis@illinois.edu*
[b]*Department of Materials Science and Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA*
†Electronic supplementary information (ESI) available: Detailed experimental methods, and supplementary figures to support the results. See DOI: 10.1039/d1nr05497j

result, the screening and optimization process is often guided by human intuition based on limited understanding of the underlying nucleation-growth mechanism based on prior and often disparate literature. This limited understanding is related to the complexity of the parameter space. Moreover, the synthesis of colloidal nanocrystals is highly sensitive in nature, where compositionally identical solutions may yield entirely different average size and polydispersity, depending on conditions employed in the synthesis.[13–15] Something as simple as decreasing the band gap by increasing QD size can cause deviations in other desired properties such as an increasing linewidth due to size broadening. Hence, several synthetic parameters need to be varied simultaneously to achieve a desired outcome, thus complicating the exploration of the parameter space and the development of new synthesis recipes. This exploration of parameter space requires unrealistically large number of experiments due to the trial-and-error nature of the conventional experimentation strategies.

These challenges in the exploration of parameter space necessitate a statistically driven approach that focuses on maximizing the utilization of synthesis insights from minimal experiments to accelerate the overall synthesis screening and materials discovery process in a resource-efficient fashion. Conventional statistical approaches utilizing optimization algorithms such as the Nelder–Mead Simplex algorithm[16] and the stable noisy optimization by branch and fit (SNOBFIT) algorithm[17,18] are based on a black-box optimization approach and often fail to predict the parameter space map unless a large number of experimental data is provided. Several of these bottlenecks are analogous to a combination of those reported in the synthesis of complex organic molecules (*e.g.*, selectivity *vs.* yield),[19–21] and polymer synthesis (*e.g.*, polydispersity/structural diversity *vs.* molecular weight).[22,23] In recent years, progress in the area of Machine Learning (ML) techniques such as artificial neural networks, reinforcement learning, and convolutional neural networks, has fueled their adaptation for chemical space exploration in the areas of organic chemistry and polymers.[24–27] These ML-based approaches, however, are still in its infancy for solution-based synthesis of inorganic nanomaterials such as semiconductor nanocrystals.

In the field of colloidal QDs, recent report by Sargent *et al.* demonstrated the application of ML models to experimental data collected over 6 years in their laboratory (~2300 experiments) to identify the chemical parameter space that can yield improved monodispersity in size of PbS QDs.[28] In an alternative approach, Zhu *et al.* developed an automated self-optimizing reactor utilizing reinforcement learning to optimize two key reaction parameters (temperature and time) to target CdSe QDs of specific size and monodispersity, highlighting the importance of automation and data science in the area of nanomaterial discovery.[29] More recently, Abolhasani *et al.* reported an autonomous flow synthesis platform for self-optimization of a target set of properties (size, polydispersity, and quantum yield) for perovskite quantum dots within 30 hours using a training set of ~275 autonomous experiments.[30–32] In particular, the integration of microfluidic

platforms with online optimization algorithms such as SNOBFIT[33] and Kriging Interpolation[34,35] for synthesis of colloidal nanocrystals underscores the promise of automation and data science for accelerating the optimization of synthesis parameter space. Additionally, utilization of combinatorial experimentation using multiple microfluidic platforms in parallel can enable rapid collection of experimental datasets that can then be used with ML models for rapid optimization and exploration of the synthesis parameters.[36,37] More recent advancements focused on synthesis of colloidal nanocrystals such as InP using multi-stage tubular flow reactors and micro-stirred tank reactors further offer more precise control of reaction conditions and process parameters, compared to conventional flask-based strategies and are also more suitable for online automation, characterization, and integration with ML models. These successful implementations of ML-based self-optimizing synthesis platforms underscore the promise of this approach to achieve target optical properties of nanomaterials with minimal user intervention. These approaches, while promising, still do not provide comprehensive understanding of the parameter space, as these feedback-based algorithms are designed specifically to achieve a target size or polydispersity but not both simultaneously. Furthermore, a large number of experiments (each set of parameters requiring experimental time on scale of tens of hours) is needed to achieve meaningful insights into the parameter space, thus making these feedback-based optimization approaches inefficient. For instance, the application of Kriging interpolation modeling for synthesis of CdSe and CdSeTe was limited to a small parameter space and resulted in low accuracy of the predictive model and thus significantly underestimates (R^2 < 0.50) the experimental outcomes (in terms of FWHM and PL intensity).[35] Alternative strategies utilizing data collected from combinatorial automated experimentation and trained using multi-layered neural network models offer improved accuracy of the model.[36,37] However, the improvement in accuracy is achieved at a cost of very large number of post-synthesis experimental data (>3000 datasets) to train ensemble neural network models for accurate mapping of the CdS and CdSe synthesis parameter space.[36,37] Hence, alternative ML-based strategies that rely on a minimal number of experiments to accurately model the synthesis parameter space are desired to provide comprehensive insights into the parameter space in addition to self-optimization for the desired optical properties. In other words, a method that can operate autonomously in a resource efficient fashion is desired, which is the purpose of this work.

In this work, we present a closed-loop autonomous flow reactor platform that utilizes a real-time ML-based feedback algorithm to efficiently map the synthesis parameter space without prior knowledge of the synthesis chemistries. It can predict the reaction outcome in a 5-dimensional input (synthesis conditions) and 2-dimensional output (size and polydispersity) parameter space, along with the uncertainty in predictions, while executing a minimum number of experiments in a closed-loop iterative automated framework. The ML-based algorithmic framework presented in this work can also be extended to efficiently learn and predict other nanocrystal pro-

perties (beyond size and polydispersity) such as crystal phase and structure of the synthesized QDs, by providing the appropriate input–output parameters from the experimental dataset. Following the standard practice across different fields including organic chemistry, polymers, and nanomaterials, we designate this self-optimizing and screening platform as an autonomous technology.[30,38,39] To develop and validate this autonomous synthesis technology, we use InP QDs, an example of a promising heavy-metal-free composition. A large fraction of its synthesis parameter space remains unexplored, thus making it an ideal candidate to demonstrate the merits of the autonomous platform. Kinetic insights gained from the ML algorithm, for the first time enable spatial separation of InP nucleation and growth, which in turn provides unprecedented control over the band gap and absorption FWHM of the nanocrystals using a flow synthesis approach.

Results and discussion

Autonomous platform components

Automated synthesis platforms often help to accelerate the exploration of synthesis parameter space for the design of nanocrystals. However, the analysis of the underlying reaction chemistries and the subsequent determination of the next set of experiments still limit their effectiveness. A closed-loop autonomous platform to perform self-driven experimental selection, execution, and analysis would significantly accelerate the overall screening and discovery process. A schematic of the autonomous flow reactor platform developed in this work is shown in Fig. 1a. The platform is comprised of three key components: (i) automated flow synthesis unit with a dual-stage flow reactor configuration, (ii) automated inline spectroscopy unit for analyzing the reaction mixture that is
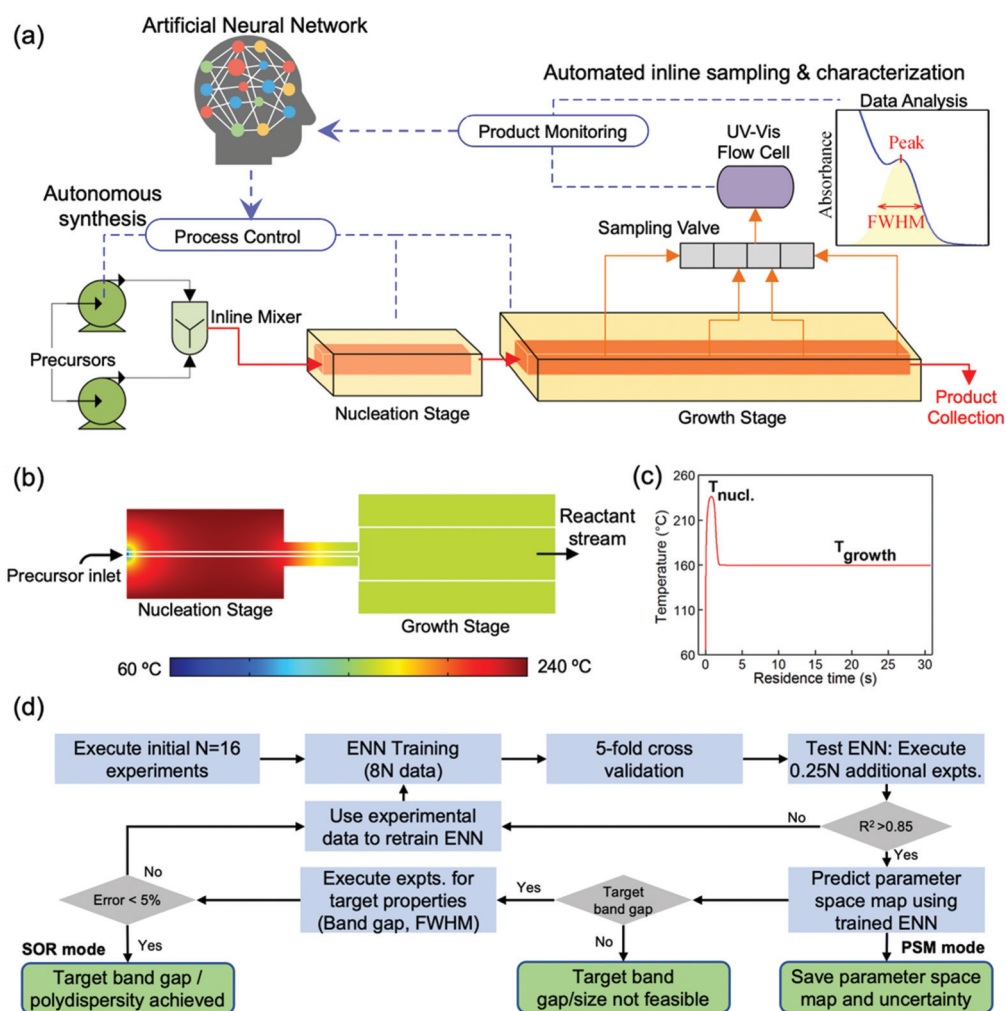

Fig. 1 (a) Schematic representation of the autonomous flow reactor platform showing key stages: precursor mixing stage, core nucleation reactor, core growth reactor, automated sampling of reaction mixture from the reaction channel, inline UV-Vis spectroscopy for real-time reaction monitoring, and artificial neural network module for training and predicting the next set of experiments. (b) Heat transfer model using COMSOL shows rapid and uniform heating of the precursors across the cross nucleation and growth reactor. (c) Temperature profile of the reaction mixture for first 30 seconds of the reaction. (d) Flow chart of the autonomous experimentation: training, validation, and testing of the neural network in an iterative fashion, followed by predicting the parameter space map and execution of next best experiments for specified target band gap and polydispersity.

sampled at different points within the reactor channel to correlate the optical properties with their structural properties, and (iii) an automated decision-making unit based on an artificial neural network model that trains on real-time experimental data and uses it to predict the next set of experiments – all in a closed loop fashion to enable autonomous synthesis of nanocrystals.

The automated flow synthesis unit is comprised of a precursor formulation module (a combination of two syringe pumps, and two peristaltic pumps), a mixing module (inline flow mixer comprised of SMX mixers), a nucleation reactor (miniature flow reactor heated with cartridge heaters), and a growth reactor with a helical static mixer insert (comprising of 8 sampling ports to sample the reaction mixture). The flow rate of the precursors is controlled to vary the precursor concentration and residence time in the reactors. All process parameters including the flow rates and reactor temperatures are automated using a custom-developed Matlab script (see ESI† for the details on flow reactor platform). A two-stage flow reactor configuration is used in this work for to enable separate control over the nucleation and growth reactions. Prior work in the flow synthesis of QDs utilizes a single-stage configuration, that limits the nucleation and growth conditions to be identical, similar to the conventional synthesis using batch reactors.[29] Fig. 1b and c shows the temperature profile inside the nucleation stage prior to entering the growth stage. The ability to rapidly (in less than 5 s) heat the precursors to a desired nucleation temperature prior to entering the growth stage provides an additional knob to control the synthesis of colloidal nanocrystals.[40,41]

Automated sampling and analysis unit: Along the growth reactor channel, a sampling valve (a switch valve with ten inputs and one output) is used to continuously sample reaction mixtures from the reactor channel. For each experiment, the reaction mixtures, corresponding to different reaction times are sampled from eight different positions within the reactor channel. The sampled reaction mixture then passes through an inline UV/Vis absorption flow cell for spectral analysis. The spectral data are fitted with a mixture of Gaussian functions to assign the absorption peak (that correlates with band gap and size of the nanocrystals) and the full-width-at-half-maxima (FWHM) of the absorption peak (that correlates with the polydispersity in size). The sampling, spectra acquisition, and spectral analysis are automated to achieve real-time monitoring of the reaction mixture. The absorption peak position and FWHM data of the reaction mixture is simultaneously utilized by an artificial neural network (ANN)-based ML algorithm[42] to determine and execute the next set of experimental conditions. Since InP QDs do not offer significant photoluminescence without post-synthetic treatment, we limited inline analysis to UV-Vis flow cell only. Alternatively, for fluorescent nanocrystals such as InP/ZnSe, inline photoluminescence flow cell can also be integrated with the automated flow reactor platform. The decision-making framework of the autonomous experimentation is discussed in the next section.

Automated decision-making unit: The autonomous experimentation platform operates in a fully automated closed loop fashion: from executing initial experiments for training and validating the decision-making neural network model, to testing its prediction accuracy. Once the model is trained, validated, and tested through self-driven experiments, the neural network is used to map parameter space and to execute subsequent autonomous experiments for a specified target band gap and polydispersity.

Fig. 1d shows the flow chart of the autonomous experimentation with different steps. First, the autonomous platform executes an initial set of partially randomized experiments (N = 16). The performance of the ANN model as a function of these randomized initial experiments is discussed in the ESI.† For each set of reaction conditions (a five-dimensional input), the peak position and FWHM obtained from the inline analysis of the reaction mixture is then used by the ensemble neural network (ENN) based ML model as 'ground truth' value for training and validation. Subsequently, the model performs a k-fold cross validation (k = 5) on the entire experimental data to evaluate the quality of the trained ENN model. An additional $0.25N$ (N = 16) set of randomized experiments is performed to test the ENN predictions using a coefficient of determination approach. Here, instead of using the initial set of experimental data to test the prediction accuracy of the ML model (as done in prior related studies), the autonomous platform performs an independent set of experiments to ensure that the ENN model performs well on dissimilar experimental data, not encountered during the initial training. If the model performs poorly (R^2 < 0.90), the data from the testing experiments are then used as additional training data. This training-to-testing operation is performed in an iterative fashion until a good predictive ENN model (R^2 > 0.90 on independent testing data) is achieved (Fig. S1†).

Once the model has been tested, the ENN model operates sequentially using two primary modes: parameter space mapping (PSM), followed by self-optimization of reaction (SOR). In PSM mode, the trained ENN model is used as the function estimator to predict the outcome (band gap and polydispersity) of the entire parameter space (a 5-dimensional input space comprising of synthesis conditions, with 11 levels for each input), along with the uncertainty in the predictions. This high-density parameter space map then provides a basis for kinetic insight into the role of different synthesis parameters as well as exploration of specific regions of interest (depending on uncertainty in prediction). In SOR mode, a grid-search method is used to identify and execute three optimal experiments for a specified target band gap and minimal polydispersity based on the predicted parameter space map. These experimental data are continuously used by the ENN model to learn and thus to improve its prediction accuracy as more and more experiments are performed.

Architecture of ensemble neural network

The ENN model is the most critical decision-making module of the autonomous experimentation platform reported here.

Paper

Nanoscale

This section discusses the architecture and performance of the ENN model, designed specifically to achieve high accuracy in prediction, while utilizing a small set of experimental data for colloidal synthesis of QDs.

ANNs have been implemented as regression models for predicting reaction outcomes for chemical synthesis previously.[43–45] Most of these works utilized a framework that consists of an input layer (reaction conditions), followed by a series of hidden layers (weights and biases), and an output layer (reaction outcomes). Through the training iteration, weights and biases are assigned to the neurons in the hidden layer to develop a function that correlates the input and the output layer. However, one of the drawbacks of such an architecture is the generalizability of the model for new set of experimental data. In our initial investigation, we found that for a small set of training data, such ANN architectures have high variance and tend to overfit the data. Even if they perform well on a given set of initial experimental data, they tend to overfit and perform poorly on a dissimilar set of experimental data. Thus, these conventional ANN architectures are inefficient for application to autonomous experimentation

approach that relies on minimal experimental inputs to gain maximum insights.

To enhance the prediction efficiency, we designed an ENN architecture that uses an ensemble learning approach. This ENN architecture comprises of multiple (up to 25) non-identical neural networks (each with a different architecture) that are trained in parallel, such that the outputs of each network can be combined to achieve a final prediction with reduced variance and higher generalizability (Fig. 2a). The input layer comprising of five synthesis parameters (nucleation temperature, growth temperature, solvent amount, zinc addition, and reaction time) are fed to each neural network in the ensemble to predict the band gap as the output parameter. In the next step, the output from the first ensemble (predicted band gap) is used as an additional input node for the prediction of polydispersity using a similar ENN architecture connected in series. This approach of using the predicted band gap as the input layer of ENN for predicting polydispersity drastically improves the overall accuracy of the model (Fig. S1†). The median of all outputs in both ensembles are used as the final predicted band gap and polydispersity, to avoid any outliers in
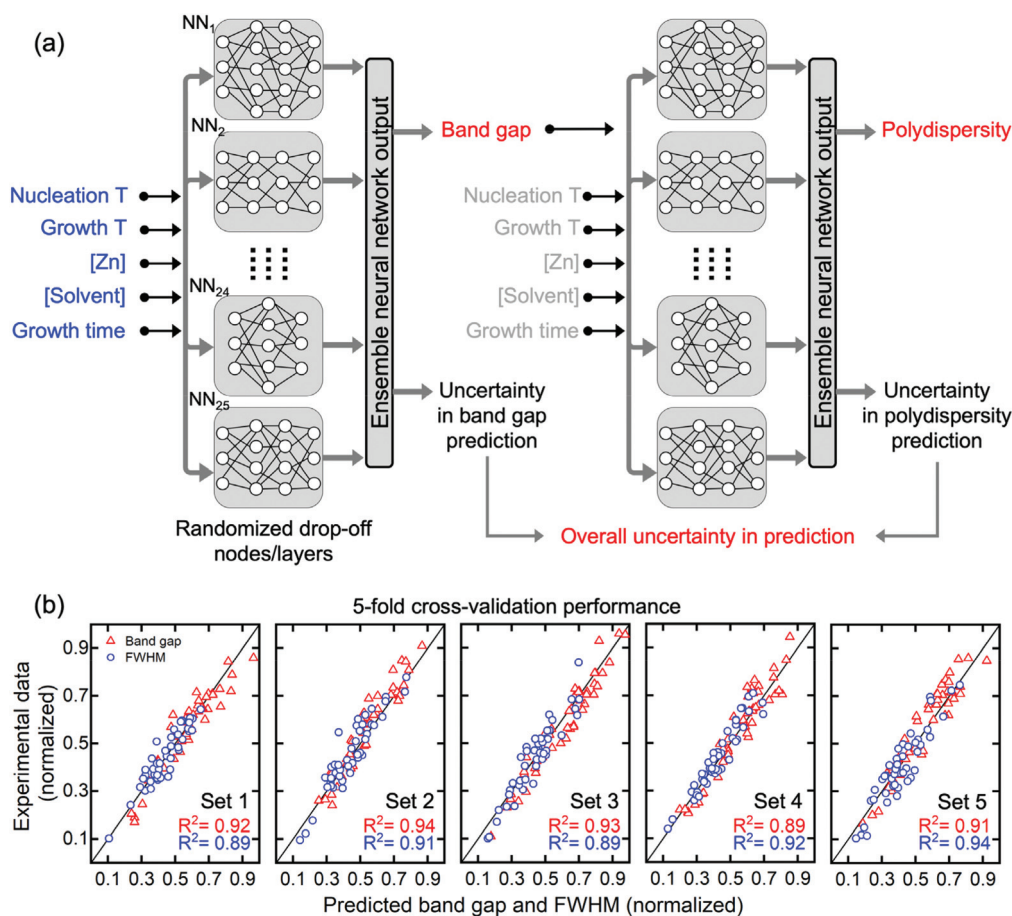

Fig. 2 (a) Ensemble neural network architecture consisting of 25 parallel neural networks with five synthesis parameters as the input layer, hidden layers with variable size. The predicted output (band gap) from the first ensemble of networks is used as an additional input to a subsequent ensemble of neural network for predicting the polydispersity. (b) 5-Fold cross validation performance based on predicted band gap and polydispersity plotted against the experimental data at same reaction conditions.

the prediction. Additionally, the uncertainty in the prediction can be estimated from the standard deviations in the predictions from each network in the ensemble. This uncertainty in prediction can guide the identification of regions that warrant further exploration.

Moreover, the cascade based ENN architecture that autonomously learns the synthesis parameter space to predict the outcome of the entire synthesis parameter space, provides the generalizability of this unique approach for applications across the domain of colloidal nanocrystals. In contrast, all prior work in the area of AI-assisted nanomaterial design ranging from real-time optimization[18,29] to deep neural network-based methods for understanding the parameter space[28,46] either require a very large number of experimental data to build an accurate predictive model or lack generalizability across different class of colloidal nanocrystals. In this work, we addressed these challenges through developing a cascade based ENN architecture that leverages band gap (size) predictions to accurately estimate the FWHM (polydispersity) across the entire parameter space.

To develop a robust network that avoids overfitting or underfitting of the training data, a dropout regularization technique is used to randomly ignore or 'drop out' a fraction of neurons in each of the neural network layers during both training and evaluation of the model, thus forcing neurons in each layer to probabilistically take less or more responsibility for the inputs. The dropout regularization enables the individual layers and networks in the ENN to be distinct from one another. More importantly, it avoids overfitting by not assigning high weightage to a specific neuron during training or evaluation.

To estimate the performance of the ENN architecture, a k-fold cross validation ($k = 5$) is performed. K-Fold cross-validation splits the training data into k sets and uses them to create multiple training-testing sets, such that the model is trained on $k − 1$ sets and tested on the remaining k^{th} set. This process is repeated k times, each with a different training-testing set of data. The ENN model predictions show high accuracy ($R^2 > 0.89$) in prediction of both band gap and polydispersity with low variance ($s_{R^2} < 0.02$) across all different k splits, confirming that the model avoids overfitting and is less biased towards any specific set of data as it performs well even on the previously unseen experimental data (Fig. 2b).

Implementation of autonomous platform for PSM

Utilizing the aforementioned ENN model as the function estimator, the autonomous platform executes partially randomized initial experiments, and uses the ENN model for PSM such that the reaction outcomes across the entire chemical space are predicted, along with the uncertainty in all predictions. As explained above, the autonomous platform starts with 16 initial self-driven experiments and undergoes multiple iterations of training–validation–testing cycle (as visualized in Fig. 1d). After a number of such iterations (utilizing a total of 4 full iterations), an accurate predictive model is developed. The autonomous platform employed for InP nanocrystals tested here required a total of 28 self-driven experiments performed through 44 hours of continuous operation. Moreover, the autonomous platform develops this accurate model without any prior knowledge of the synthesis chemistry.

Fig. 3a shows the predicted absorption peak wavelength (band gap), FWHM (polydispersity), and the overall certainty in the predictions for all possible 161 051 combinations of the synthesis parameters (11 levels for each of the 5 synthesis parameters: nucleation temperature, growth temperature, solvent
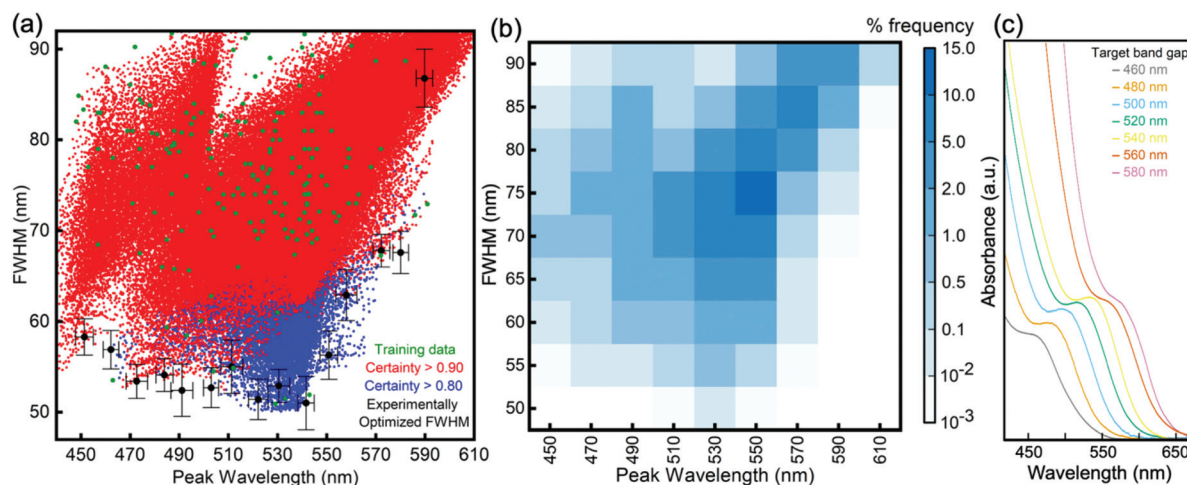

Fig. 3 (a) Ensemble neural network-predicted absorbance peak wavelength (band gap) and FWHM (polydispersity) across the entire parameter space. The data points in red and blue correspond to predictions with certainty higher than 0.90 and 0.80 respectively. Experimental band gap and polydispersity data corresponding to the optimal experiments performed by the autonomous platform for specified target wavelengths from 450 to 590 nm (at an interval of 10 nm each) are shown in black. The error bar corresponds to 3 different predictions for each specified target wavelength. (b) Heat map showing the output frequency based on all combination of synthesis conditions across the entire parameter space for specific peak wavelength and FWHM. (c) Absorption spectra of several InP sizes with best monodispersity, synthesized using the autonomous flow reactor platform.

amount, zinc addition, and reaction time), along with the experimental data used for training the ENN model. The certainty of the majority of the predictions across the entire parameter space is higher than 0.90, confirming the robustness of the ENN model to efficiently learn from a small set of initial training data. Fig. S2† shows a plot of the evolution of certainty in the predictions upon executing different iterations of self-driven experiments by the autonomous platform.

More importantly, the PSM plot (Fig. 3a) reveals the minimum polydispersity that can be expected for InP nanocrystals exhibiting different possible band gaps across the entire visible spectrum. These minima in FWHM suggest that the synthesis chemistry (using indium chloride, zinc chloride, oleylamine and aminophosphine-based starting reagents) is well-suited for synthesis of InP QDs with a band gap between 2.24 eV (550 nm) and 2.63 eV (470 nm), because the minimum possible FWHM remains less than 55 nm in this range. For the synthesis of nanocrystals with a smaller or larger band gap, the minimum possible FWHM is predicted to be significantly higher, suggesting higher polydispersity in the nanocrystal size in this region. These observations from the PSM plot align with the reports in the prior literature. For instance, the feasible InP QD bandgap reported in the prior literature using aminophosphine-based chemistry lies between 460 nm and 570 nm with the minimal polydispersity (absorbance FWHM) reaching up to 53 nm for particles exhibiting first absorption excitonic peak at 530 nm.[13,33,47] The absence of feasible nanocrystals exhibiting FWHM below 50 nm using aminophosphine precursors can be attributed to the rapid precursor conversion kinetics of the active phosphorous species, that eventually limits the monomer supply rate and thus prevents the required control over the tunability of the band gap and polydispersity.[13] To achieve InP nanocrystals exhibiting FWHM below 50 nm and expand the feasible band gap region, alternate solvent-precursor combinations can be explored. For instance, the implementation of silylphosphine based phosphorous precursors using seed-mediated synthetic approach for synthesis of InP QDs have yielded nanocrystals exhibiting narrower absorption FWHM and broader band gap tunability across the visible spectrum.[48] Similarly, adjusting the alkyl group of the aminophosphine precursors can allow further control over the InP nucleation and growth reactions, thus yielding nanocrystals with broader band gap tunability.[49]

Implementation of autonomous platform for SOR

In the SOR mode, the autonomous platform utilizes the predicted PSM dataset to determine and execute the next set of experiments that would yield the best polydispersity (minimum FWHM) for a target band gap. The experimental results from these self-optimizing experiments are shown as black circles in Fig. 3a. For specified target wavelengths in the range of 450 nm to 590 nm, each at an interval of 10 nm, the autonomous platform performs several self-driven experiments with the goal to achieve the smallest FWHM across all target

wavelengths. For each target wavelength, three predicted experiments are performed to assess the robustness of the autonomous platform. The standard deviations of these experimental yields are shown with the error bar associated with both peak wavelength and FWHM. As evident from Fig. 3a, for nearly all target wavelengths (<580 nm), the self-executed experiments yielded nanocrystals with peak wavelength and FWHM close to the target wavelength and predicted optimal FWHM. The absorption spectra of the synthesized InP QDs for different target band gaps with the best monodispersity are shown in Fig. 3c. The ability of the autonomous platform to efficiently synthesize InP QDs of specified target band gap with optimal polydispersity in size, without any prior knowledge of the reaction space is attributed to the robust ENN-based ML model that efficiently learns the reaction space through self-driven iterative experiments to eventually predict the reaction outcomes with high certainty (>80%).

These results highlight the merits of AI-assisted autonomous experimentation for intelligent sequential selection of the next best experiment by efficiently learning synthesis parameter space. The question now remains: how impactful and efficient is this AI-guided approach, compared to user-selected exploration of synthesis parameter space? To answer this question, we looked at the frequency map of AI-derived predictions (Fig. 3b). This plot bins the frequency of different synthesis conditions in terms of predicted peak wavelength (bins of 20 nm, *x*-axis) and FWHM (bins of 5 nm, *y*-axis). Most synthesis conditions (>60%) are predicted to yield nanocrystals with band gap between 2.13 eV (580 nm) to 2.47 eV (500 nm) and high polydispersity in size (FWHM > 70 nm). Notably, using this analysis, less than 1% of the entire parameter space is expected to result in nanocrystals with optimal polydispersity for a target band gap across the entire visible spectrum. Such low density of synthesis parameter space in the optimal polydispersity region suggests that the probability of identifying optimal synthesis conditions using conventional trial-and-error-based synthesis planning approaches (few experiments) is unlikely. In other words, a conventional synthesis approach can only be effectively used to map a limited region of synthesis parameter space, thus making the exploration of complex parameter space expensive in time and resources. Although, similar polydispersity and size of the colloidal QDs can also be synthesized by an experienced human researcher (who starts with significant prior knowledge of similar or dissimilar systems) after several trials, the mapping of the entire synthesis parameter space using conventional experimentation approach will require a drastically large number of experiments (>1000), given the complex relation between different synthesis parameters. In contrast, utilizing the AI-guided approach that makes use of the autonomous platform, maps the entire synthesis parameter space by learning from 28 self-driven experiments in less than 44 hours of continuous operation. This highlights the essential role of intelligent autonomous platforms to minimize the exploration time and resources to accelerate the overall screening and discovery process of colloidal nanocrystals.

Effect of zinc precursor addition on band gap and polydispersity

Further analysis of the predicted parameter space map revealed the critical role of zinc precursors in controlling the size and monodispersity of the InP QDs. Fig. 4a and b show the effect of zinc precursor concentration on the absorption peak position and FWHM of InP QDs for different nucleation and growth temperatures, as predicted by the ENN model. An increase in the zinc precursor concentration ($ZnCl_2$) is expected to yield InP QDs of smaller size, as evident from the blue-shift in absorption peak. The size of the nanocrystals is more sensitive to the zinc precursor concentration at higher growth temperatures (240 °C). Similarly, the FWHM decreases upon an increase in zinc precursor concentration, suggesting that higher concentration yields QDs with improved monodispersity in size. Fig. S3 and S4† shows a five-dimensional visualization of synthesis parameter space consisting of predicted reaction outcomes for different combinations of all five synthesis inputs.

Next, we look at how the trends predicted by the ENN-based algorithm compare with experimental results reported previously. Tessier *et al.* previously reported better control over size and improved excitonic features through addition of high concentration of zinc salts (such that [Zn] : [In] = 5) to the aminophosphine-based synthesis chemistry.[33] The mechanistic role of zinc in the InP nucleation and growth process, however, remains elusive. One hypothesis for the observed improvement in monodispersity upon zinc addition is that zinc passivates InP traps and defects, thus stabilizing the QD surface that leads to slower growth and better size-uniformity. Recent work by Buffard *et al.*, however, revealed that only a very small fraction of zinc is incorporated in the InP NCs and hence, is unlikely to localize and passivate surface traps.[47] A more likely possibility is that the InP precursor conversion reaction requires activation by a zinc halide precursor to form the active precursor species, thus altering the nucleation and growth rates. Density functional theory calculations suggest that the Lewis acid nature of zinc compensates for the high electronegativity of lone pairs on the aminophosphine nitrogen, thus stabilizing the InP precursor conversion reaction.[50] Subsequent research needs to focus on developing a better mechanistic understanding of the role of zinc precursor in the mechanism of InP precursor conversion reaction. These observations are in alignment with the trends observed in the predictions by the ML-based model. Despite no prior knowledge of the synthesis chemistry, the ML-based model appropriately captures the trends observed in prior work regarding the effect of zinc precursors on InP nanocrystal size and polydispersity. Additionally, the trends suggested from the ENN predictions suggest that the FWHM of the synthesized InP QDs are more sensitive to the presence of Zinc salts at lower growth temperatures (160 °C, Fig. 4b). Nearly all prior studies on understanding the role of Zinc salts on the growth kinetics and size tunability utilized higher growth temperatures (>180 °C) and thus lack the critical interplay between nucleation-growth temperature and zinc chloride concentration and its impact on the growth kinetics. A possible explanation for improved FWHM at higher zinc salt concentration could be that the addition of zinc chloride (that acts as a Lewis acid) stabilizes the InP monomers formed during the precursor conversion kinetics. Moreover, the lower growth temperature condition allows for maintaining slow precursor conversion kinetics and thus allows continuous supply of the monomers for a prolonged duration, compared to the reactions in absence of zinc chloride. Future studies should focus on exploring the role of zinc salt addition on understanding the intrinsic precursor conversion kinetics, possibly using a phosphorous NMR study of the reaction kinetics at controlled nucleation and growth temperatures.
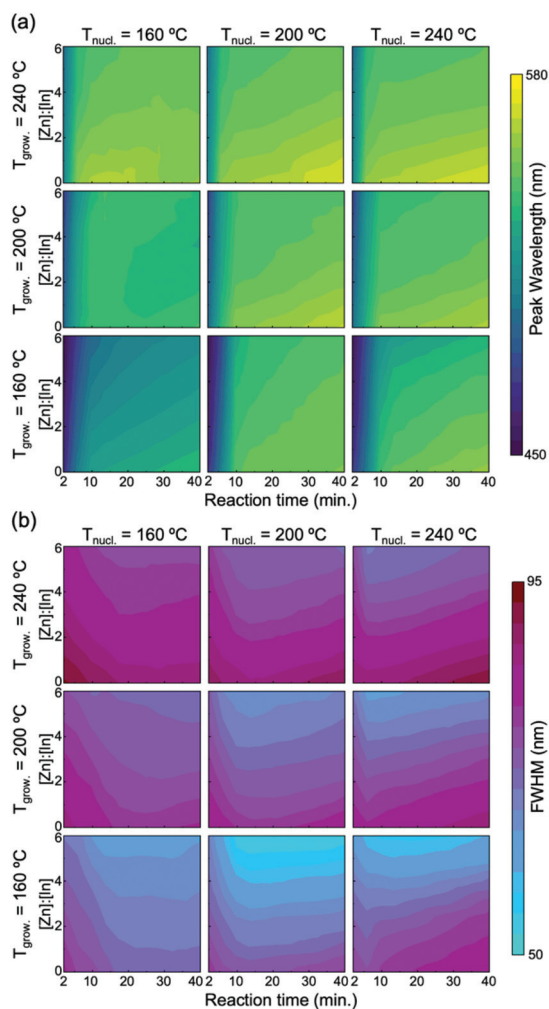

Fig. 4 (a) Absorbance peak wavelength and (b) FWHM corresponding to different [Zn] : [In] ratio for varying reaction time, nucleation temperature, and growth temperatures based on the parameter space mapping predictions.

Kinetic insights into InP nucleation and growth

In addition to a strong dependence on zinc precursor concentration, the predictions by the ENN model also indicate a

strong dependence of QD size and polydispersity on nucleation and growth temperatures. The role of nucleation temperature as a separate synthesis knob for InP synthesis, however, remains unexplored. Nearly all previously reported syntheses of InP QDs are performed at identical nucleation and growth temperatures. The ability of the dual-stage reactor configuration reported here to rapidly heat the reaction mixtures through miniature flow channels prior to entering the growth stage in a dual-stage flow configuration, however, can be leveraged to achieve independent control over nucleation and growth conditions (Fig. 1). This reactor configuration is used by the ENN model to learn the parameter space through autonomous experimentation. The predictions from the ENN model (Fig. 4b) suggest that moderately high nucleation temperature (200 °C) in combination with a lower growth temperature (160 °C) is expected to yield InP QDs with improved monodispersity (lower FWHM) in size. Fig. 5a shows the absorption spectra of InP QDs, obtained from the experiments performed at optimal nucleation and growth temperature for different growth times. The excitonic peak features sharpen (size-focusing before 20 minutes of growth, followed by a slight broadening from 30 minutes to 40 minutes) with an increase in both growth time and QD size, thus suggesting a size-focused growth mechanism, depending on nucleation-growth temperature and reaction time. The range of nucleation-growth temperatures explored by the autonomous platform was limited by the boiling point and reactivity of the reagents. Therefore, for synthesis of InP QDs using aminophosphine precursors, the lower and upper limit of this range is maintained at 160 °C and 240 °C respectively. Further exploration, outside of this temperature range can also be performed

experimentally. However, the predictions from the ENN model suggest that the FWHM increases drastically at higher growth temperatures (Fig. 4).

Experimental data on nanocrystal size, concentration, and polydispersity of InP nanocrystals synthesized at different combinations of nucleation and growth temperatures (Fig. 5b–d) provided further insights into the role of nucleation temperature as yet another synthesis knob. An increase in nucleation temperature (above the growth temperature, 160 °C) yields InP nanocrystals of smaller diameters (Fig. 5b). This decrease in size of the nanocrystals can be rationalized by the higher rate of nucleation at elevated temperatures. An increased rate of nucleation leads to depletion of the monomer reservoir in the early stage of the growth reaction and as a result, the estimated concentration of nanocrystals in the reaction mixture is higher for reactions at elevated nucleation temperature (Fig. 5c). However, at nucleation temperatures above 200 °C, the concentration of nanocrystals steadily drops after initial growth. This drop in the nanocrystal concentration suggests that with depletion of the monomer reservoir, the growth process is largely limited by ripening or redissolution of nanocrystals. At the lower nucleation temperature of 160 °C (same as the growth temperature), however, the rate of nucleation is significantly slower, which prevents the burst nucleation process needed to separate nucleation from growth. As a result, the nanocrystals continue to nucleate throughout the reaction and hence the concentration increases steadily with growth time. At moderately high nucleation temperatures (160 °C < $T_{\text{nucleation}}$ < 220 °C), the nanocrystal concentration saturates after an initial increase, suggesting that no additional nuclei are formed, leaving growth as the predominant process. These
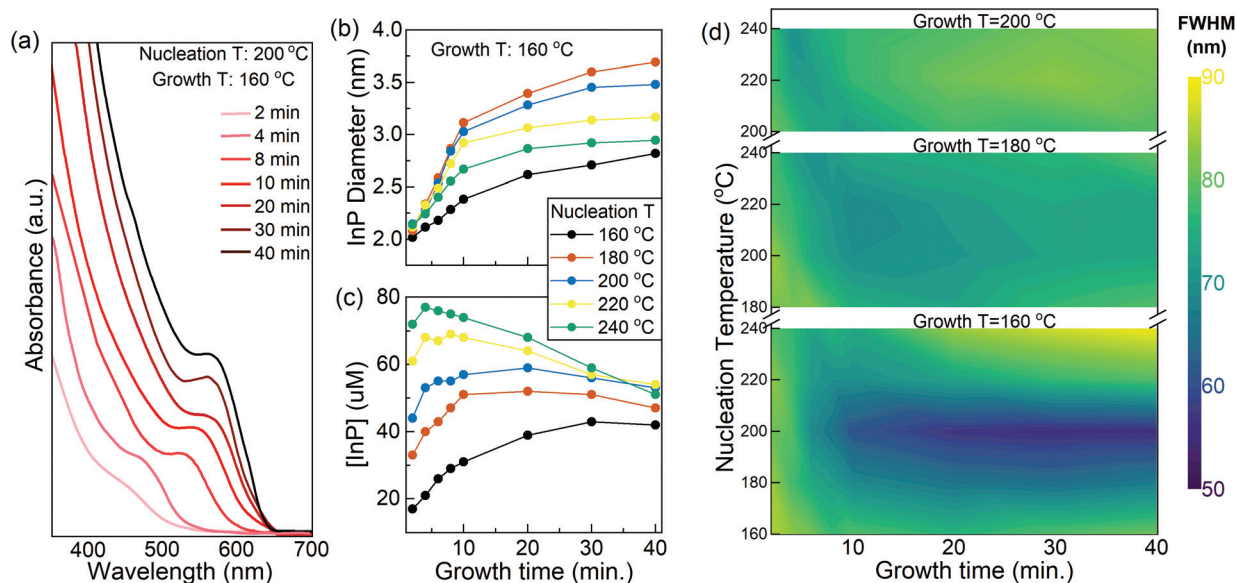

Fig. 5 Experimental data: (a) UV-Vis spectra of InP nanocrystals at optimal nucleation and growth temperature. Change in (b) diameter and (c) concentration of InP nanocrystals at different nucleation temperature throughout the growth process. (d) FWHM map at different nucleation and growth temperature across the growth time. The corresponding polydispersity map in terms of FWHM/bandgap (in energy units) fraction is shown in Fig. S5 (ESI†).

trends are also in agreement with the observed change in polydispersity in size. Fig. 5d shows a map of polydispersity (quantified by FWHM) as a function of nucleation temperature, growth temperature, and growth time. The surface plot is obtained by interpolating between the experimental data collected at different nucleation temperatures, growth temperatures, and growth time. InP QDs with minimum polydispersity (FWHM < 55 nm) are synthesized at a moderately high nucleation temperature of 200 °C and a low growth temperature of 160 °C. Fig. S5† shows the polydispersity surface plot corresponding to the energy units (FWHM (eV)/bandgap (eV)). Under these optimal conditions, the FWHM drops (Fig. 5d) and the concentration of nanocrystals in the reaction mixtures saturates (Fig. 5c) upon an increase in growth time – a strong indication of spatial separation of nucleation and growth. The ability to independently tune the nucleation and growth conditions enables us to drive the reaction to a size-focusing regime. Higher nucleation temperatures (>200 °C) yielded QDs exhibiting higher polydispersity in size (FWHM > 80 nm) for longer growth times. This increase in FWHM (Fig. 5d), coupled with the steady drop in concentration (Fig. 5c) of InP nanocrystals with growth time, further confirms that nanocrystal growth under these synthesis conditions is limited by redissolution and ripening. Owen *et al.* recently reported the dependence of reaction temperature on the nanocrystal growth kinetics for different derivatives of aminophosphine precursors using a hot injection-based batch synthesis method.[49] Using a nucleation factor as the measure of nucleation time scale, lower nucleation temperature was reported to be responsible for slow and continuous nucleation, while higher reaction temperatures yielded in more of a "burst" type nucleation. These results are similar to the observations in our work for identical nucleation and growth temperatures in both the stages. Additionally, controlling higher nucleation temperature (200 °C) for a short timescale, followed by prolonged growth at lower temperature (160 °C) in a flow reactor configuration mimics the "burst" type nucleation model and thus enables spatial separation of nucleation and growth.

These results underscore the critical the role of nucleation temperature as a versatile synthesis knob to drive the pathway for InP growth in three different growth regimes: (i) continuous nucleation along with growth, (ii) spatial separation of nucleation from growth, and (iii) ripening or redissolution of nanocrystals. Spatial separation of nucleation and growth under optimal synthesis conditions is unique to the dual-stage flow reactor configuration used in this work and is not feasible with the similar degree of control when using a traditional hot injection-based batch synthesis approach. Although varying the amount of room-temperature reagents injected to the hot solvent using the hot-injection based synthesis approach can in principle separate the nucleation and growth processes, precise control over the temperature drop during the hot-injection process remains challenging in conventional batch based synthesis strategies due to limited control over the process conditions. Moreover, the ability of the autonomous experimentation platform to learn synthesis parameter space and to provide new and previously unexplored insights into the nucleation-growth kinetics underscores the merits of these AI-based strategies for accelerated screening of the synthesis chemistry and as an efficient tool for understanding the underlying reaction kinetics.

Conclusions

We have demonstrated a robust and efficient AI-assisted flow synthesis strategy that enables autonomous synthesis of colloidal nanocrystals without any prior knowledge of the synthesis parameter space. The autonomous platform performs (i) initial experiments to explore the parameter space, (ii) trains an ensemble neural network (ENN) model to learn the parameter space, (iii) rapidly maps parameter space with high prediction certainty to provide synthesis insights, and (iv) executes the next best experiment to synthesize nanocrystals of a specified target band gap and polydispersity. We demonstrated these capabilities by synthesizing InP nanocrystals exhibiting optimal polydispersity in size for specified target band gaps across the visible spectrum. Specifically, using a closed-loop iterative framework, it executed a minimal number of self-driven iterative experiments (28 experiments, 224 experimental data) through continuous operation (44 hours) to learn the entire synthesis parameter space and from that an accurate ENN model is developed for predicting the reaction outcomes for more than 100 000 different combinations of synthesis conditions. To put this in perspective, if compared to a combinatorial search (assuming a design of experiments methods with at least three levels for each factor) over the entire synthesis parameter space of semiconductor nanocrystals, the AI-guided approach reported here improves the experimental efficiency by at least ten-folds. Moreover, application of the unique sequential ensemble neural networks to other class of nanocrystals such as perovskites can drastically minimize the number of experiments required to achieve the desired properties by at least four-folds, compared to the current state-of-the-art AI-driven experimentation methods that require more than 150 experiments under autonomous operation mode.[30,31] Other data science driven methods such as Universal Kriging-based statistical interpolation approach has also been previously implemented for synthesis of CdSe and CdSeTe QDs. Although, the model performed well in predicting the qualitative trends in peak positions and FWHM, it failed to accurately predict these features (R^2 on testing data < 0.75), despite exploration of a small two-dimensional parameter space (comprising of residence time, and precursor ratio). Application of such methods to develop an accurate predictive model for more complex multi-dimensional parameter space such as InP (reported in this work) would thus not be feasible without requiring a large number of experimental data (>1000) for training. In contrast, the implementation of ENN-based approach (reported in this work) allows the model to accurately learn a multi-dimensional parameter space ($R^2 > 0.90$ on testing data, Fig. 2), through minimal number of self-driven iterative experiments (28 experiments, 224 experimental data).

In addition to being more efficient, the second major advantage of this ENN approach over prior ML-based approaches is that the predictions from the ENN model, despite no prior knowledge, not only capture the trends observed in prior work (*e.g.*, the effect of zinc concentration), but also provides new kinetic insights that have not been reported in the literature. For instance, the trends observed in the predicted reaction outcomes enabled us to identify regions of the parameter space that allow spatial separation of nucleation and growth to achieve size-focused growth of monodispersed InP nanocrystals with the desired band gap. These findings demonstrate the efficient learning capability of the ENN model for autonomous exploration of the chemical space and synthesis of the nanocrystals exhibiting optimal band gap and FWHM. The material properties of the InP nanocrystals synthesized autonomously by the AI-guided platform (starting with no prior knowledge of the chemical space) are similar to the best results reported in the prior literature on aminophosphine-based InP nanocrystals (synthesized using the wealth of knowledge and expertise acquired over several years of exploration of a specific synthesis chemistry), thus validating the application of the ENN model reported here for autonomous synthesis of colloidal nanocrystals. The application of such ENN-based models for synthesis of colloidal nanocrystals can thus have major advantage over conventional experimental approach in terms of both (i) accelerated exploration of synthesis parameter space and (ii) providing key data-driven insights into the synthesis chemistry.

The ability to rapidly learn synthesis parameter space for a given set of starting materials through self-driven experiments and the ability to unravel new kinetic insights, highlights the key merits of the AI-guided autonomous synthesis approach for accelerating the screening and discovery of colloidal nanocrystals reported here. Moreover, the generalizability and robustness suggest that our ENN-based approach can be readily adapted and applied to broader classes of nanocrystals such as perovskites and IR-emitting QDs. Moreover, the autonomous experimentation framework presented in this work requires minimal expertise in the data-science and thus can be easily adapted by the researchers from the broader nanomaterials field in their own automated experiments as well. Further implementation and adoption of such AI-based models and of the autonomous framework presented in his work to different class of materials and synthesis planning strategies will facilitate the development of efficient end-to-end autonomous discovery platforms. This AI methodology and autonomous experimentation framework reduces the use of resources in terms of time and chemicals while providing maximum synthetic insights that accelerate material discovery.

Data availability

The experimental and ML-predicted data on process parameters, and the corresponding bandgaps and FWHM dataset generated and analyzed during the study reported in this work is publicly available at https://github.com/kenisgroupml/Quantum-Dot-ENN. The autonomous experimentation code, including automation of different equipment and ensemble neural network integration is not publicly available due to inclusion of third-party code but is available upon request from the corresponding author.

Author contributions

A. V. conceived the project, developed the AI and automation codes, designed and conducted the experiments, analyzed the data and wrote the manuscript. K. B. contributed to the AI code development. A. Z. performed the synthesis experiments and contributed to the process automation. M. S. conceived the project and provided critical feedback on the experiments and the manuscript. P. J. A. K. conceived the project, directed the research, and wrote the manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Notes and references

1 J. Y. Kim, O. Voznyy, D. Zhitomirsky and E. H. Sargent, *Adv. Mater.*, 2013, **25**, 4986–5010.
2 D. V. Talapin, J. S. Lee, M. V. Kovalenko and E. V. Shevchenko, *Chem. Rev.*, 2010, **110**, 389–458.
3 P. V. Kamat, *J. Phys. Chem. C*, 2008, **112**, 18737–18753.
4 Y. Wang, E. L. Runnerstrom and D. J. Milliron, *Annu. Rev. Chem. Biomol. Eng.*, 2016, **7**, 283–304.
5 C. J. Stolle, T. B. Harvey and B. A. Korgel, *Curr. Opin. Chem. Eng.*, 2013, **2**, 160–167.
6 Y. Pu, F. H. Cai, D. Wang, J. X. Wang and J. F. Chen, *Ind. Eng. Chem. Res.*, 2018, **57**, 1790–1802.
7 Y. H. Won, O. Cho, T. Kim, D. Y. Chung, T. Kim, H. Chung, H. Jang, J. Lee, D. Kim and E. Jang, *Nature*, 2019, **575**, 634–638.
8 P. Mushonga, M. O. Onani, A. M. Madiehe and M. Meyer, *J. Nanomater.*, 2012, **2012**, 869284.
9 A. Thomas, P. V. Nair and K. G. Thomas, *J. Phys. Chem. C*, 2014, **118**, 3838–3845.
10 B. M. Cossairt, *Chem. Mater.*, 2016, **28**, 7181–7189.
11 A. M. Derfus, W. C. W. Chan and S. N. Bhatia, *Nano Lett.*, 2004, **4**, 11–18.
12 G. H. Jia, Y. P. Pang, J. J. Ning, U. Banin and B. T. Ji, *Adv. Mater.*, 2019, **31**, 1900781.

13 A. Vikram, A. Zahid, S. S. Bhargava, L. P. Keating, A. Sutrisno, A. Khare, P. Trefonas, M. Shim and P. J. A. Kenis, *Chem. Mater.*, 2020, **32**, 3577–3584.

14 P. Reiss, M. Carriere, C. Lincheneau, L. Vaure and S. Tamang, *Chem. Rev.*, 2016, **116**, 10731–10819.

15 J. Chang and E. R. Waclawik, *RSC Adv.*, 2014, **4**, 23505–23527.

16 J. P. McMullen and K. F. Jensen, *Org. Process Res. Dev.*, 2010, **14**, 1169–1176.

17 W. Huyer and A. Neumaier, *ACM Trans. Math. Software*, 2008, **35**, 9.

18 S. Krishnadasan, R. J. Brown, A. J. deMello and J. C. deMello, *Lab Chip*, 2007, **7**, 1434–1441.

19 J. M. Granda, L. Donina, V. Dragone, D. L. Long and L. Cronin, *Nature*, 2018, **559**, 377–381.

20 A. F. Zahrt, J. J. Henle, B. T. Rose, Y. Wang, W. T. Darrow and S. E. Denmark, *Science*, 2019, **363**, eaau5631.

21 Z. Zhou, X. Li and R. N. Zare, *ACS Cent. Sci.*, 2017, **3**, 1337–1344.

22 J. Bajorath, S. Kearnes, W. P. Walters, N. A. Meanwell, G. I. Georg and S. M. Wang, *J. Med. Chem.*, 2020, **63**, 8651–8652.

23 P. Schneider, W. P. Walters, A. T. Plowright, N. Sieroka, J. Listgarten, R. A. Goodnow, J. Fisher, J. M. Jansen, J. S. Duca, T. S. Rush, M. Zentgraf, J. E. Hill, E. Krutoholow, M. Kohler, J. Blaney, K. Funatsu, C. Luebkemann and G. Schneider, *Nat. Rev. Drug Discovery*, 2020, **19**, 353–364.

24 H. Y. Gao, T. J. Struble, C. W. Coley, Y. R. Wang, W. H. Green and K. F. Jensen, *ACS Cent. Sci.*, 2018, **4**, 1465–1476.

25 S. Mohapatra, N. Hartrampf, M. Poskus, A. Loas, R. Gomez-Bombarelli and B. L. Pentelute, *ACS Cent. Sci.*, 2020, **6**, 2277–2286.

26 C. Empel and R. M. Koenigs, *Angew. Chem., Int. Ed.*, 2019, **58**, 17114–17116.

27 P. Schwaller, D. Probst, A. C. Vaucher, V. H. Nair, D. Kreutter, T. Laino and J. L. Reymond, *Nat. Mach. Intell.*, 2021, **3**, 144–152.

28 O. Voznyy, L. Levina, J. Z. Fan, M. Askerka, A. Jain, M. J. Choi, O. Ouellette, P. Todorovic, L. K. Sagar and E. H. Sargent, *ACS Nano*, 2019, **13**, 11122–11128.

29 J. G. Li, Y. X. Tu, R. L. Liu, Y. H. Lu and X. Zhu, *Adv. Sci.*, 2020, **7**, 2001116.

30 A. A. Volk, R. W. Epps and M. Abolhasani, *Adv. Mater.*, 2021, **33**, 2004495.

31 R. W. Epps, M. S. Bowen, A. A. Volk, K. Abdel-Latif, S. Y. Han, K. G. Reyes, A. Amassian and M. Abolhasani, *Adv. Mater.*, 2020, **32**, 2001626.

32 R. W. Epps, A. A. Volk, K. G. Reyes and M. Abolhasani, *Chem. Sci.*, 2021, **12**, 6025–6036.

33 M. D. Tessier, D. Dupont, K. De Nolf, J. De Roo and Z. Hens, *Chem. Mater.*, 2015, **27**, 4893–4898.

34 L. Bezinge, R. M. Maceiczyk, I. Lignos, M. V. Kovalenko and A. J. Demello, *ACS Appl. Mater. Interfaces*, 2018, **10**, 18869–18878.

35 R. M. Maceiczyk and A. J. deMello, *J. Phys. Chem. C*, 2014, **118**, 20026–20033.

36 K. Watanabe, Y. Orimoto, K. Nagano, K. Yamashita, M. Uehara, H. Nakamura, T. Furuya and H. Maeda, *Chem. Eng. Sci.*, 2012, **75**, 292–297.

37 Y. Orimoto, K. Watanabe, K. Yamashita, M. Uehara, H. Nakamura, T. Furuya and H. Maeda, *J. Phys. Chem. C*, 2012, **116**, 17885–17896.

38 M. Rubens, J. H. Vrijsen, J. Laun and T. Junkers, *Angew. Chem., Int. Ed.*, 2019, **58**, 3183–3187.

39 D. Cortes-Borda, E. Wimmer, B. Gouilleux, E. Barre, N. Oger, L. Goularnaly, L. Peault, B. Charrier, C. Truchet, P. Giraudeau, M. Rodriguez-Zubiri, E. Le Grognec and F. X. Felpin, *J. Org. Chem.*, 2018, **83**, 14286–14299.

40 A. Vikram, A. Zahid, S. S. Bhargava, H. Jang, A. Sutrisno, A. Khare, P. Trefonas, M. Shim and P. J. A. Kenis, *ACS Appl. Nano Mater.*, 2020, **3**, 12325–12333.

41 A. Vikram, V. Kumar, U. Ramesh, K. Balakrishnan, N. Oh, K. Deshpande, T. Ewers, P. Trefonas, M. Shim and P. J. A. Kenis, *ChemNanoMat*, 2018, **4**, 943–953.

42 L. White, R. Togneri, W. Liu and M. Bennamoun, *Stud. Comput. Intell.*, 2019, **783**, 1–21.

43 J. M. Granda, L. Donina, V. Dragone, D. L. Long and L. Cronin, *Nature*, 2018, **559**, 377–381.

44 G. Schneider, *Nat. Rev. Drug Discovery*, 2018, **17**, 97–113.

45 B. A. Rizkin, A. S. Shkolnik, N. J. Ferraro and R. L. Hartman, *Nat. Mach. Intell.*, 2020, **2**, 200–209.

46 F. Mekki-Berrada, Z. K. Ren, T. Huang, W. K. Wong, F. Zheng, J. X. Xie, I. P. S. Tian, S. Jayavelu, Z. Mahfoud, D. Bash, K. Hippalgaonkar, S. Khan, T. Buonassisi, Q. X. Li and X. N. Wang, *npj Comput. Mater.*, 2021, **7**, 55.

47 A. Buffard, S. Dreyfuss, B. Nadal, H. Heuclin, X. Z. Xu, G. Patriarche, N. Mezailles and B. Dubertret, *Chem. Mater.*, 2016, **28**, 5925–5934.

48 P. Ramasamy, K. J. Ko, J. W. Kang and J. S. Lee, *Chem. Mater.*, 2018, **30**, 3643–3647.

49 B. M. McMurtry, K. Qan, J. K. Teglasi, A. K. Swarnakar, J. De Roo and J. S. Owen, *Chem. Mater.*, 2020, **32**, 4358–4368.

50 G. Laufersky, S. Bradley, E. Frecaut, M. Lein and T. Nann, *Nanoscale*, 2018, **10**, 8752–8762.