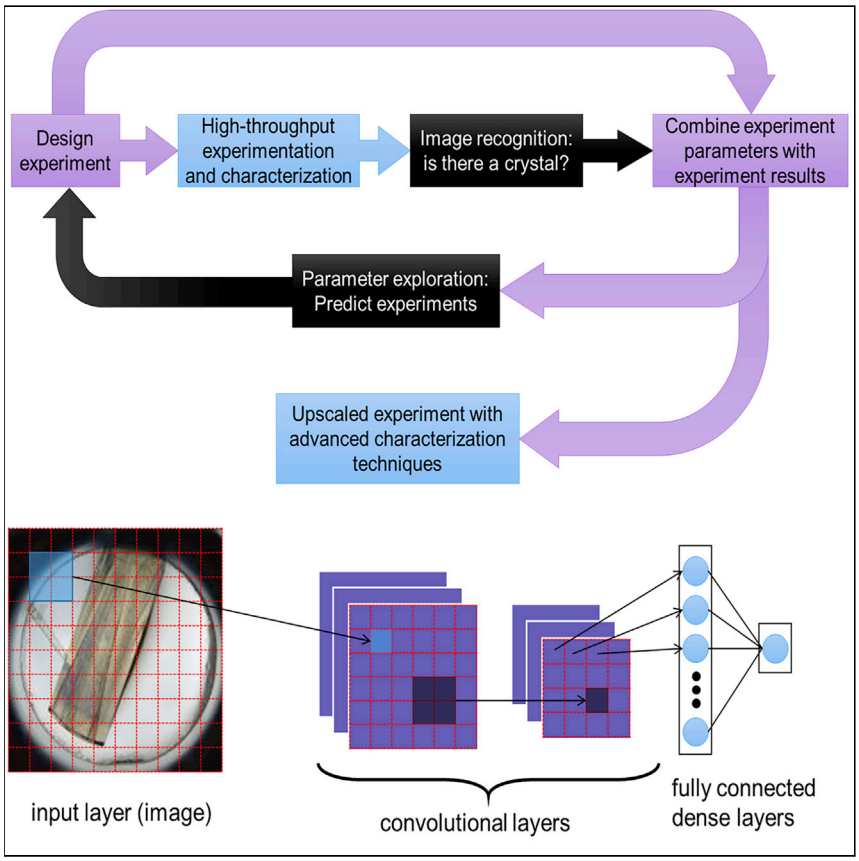
Matter

CellPress

Article

Machine-Learning-Accelerated Perovskite Crystallization



Here, we report a high-throughput experimental framework for the discovery of new perovskite single crystals. We use machine learning (ML) to guide the sequence of ever-improved robotic synthetic trials. We perform high-throughput syntheses of perovskite single crystals and characterize the outcomes with convolutional neural network-based image recognition. We then use an ML model to predict the optimal conditions for the synthesis of a new perovskite single crystal; as a result, we report the first synthesis of $(3\text{-PLA})_2PbCl_4$.

Jeffrey Kirman, Andrew Johnston, Douglas A. Kuntz, ..., Dongxin Ma, Gilbert G. Privé, Edward H. Sargent

ted.sargent@utoronto.ca

HIGHLIGHTS

A framework for the high-throughput synthesis of perovskite single crystals

Automated characterization with ML image-recognition techniques

Maximized rate-of-learning enabled by ML guidance

Discovery and synthesis of a new, blue-emitting perovskite-like material

2 **Benchmark**
First qualification/assessment of material properties and/or performance

Kirman et al., Matter 2, 938–947
April 1, 2020
https://doi.org/10.1016/j.matt.2020.02.012

Matter

CellPress

Article

Machine-Learning-Accelerated Perovskite Crystallization

Jeffrey Kirman,[1,4] Andrew Johnston,[1,4] Douglas A. Kuntz,[2] Mikhail Askerka,[1] Yuan Gao,[1] Petar Todorović,[1] Dongxin Ma,[1] Gilbert G. Privé,[2,3] and Edward H. Sargent[1,5,*]

SUMMARY

Perovskites have seen significant research interest in the last decade. As ternary and quaternary compounds, their chemical space is exceptionally large, yet perovskite development has been limited to a restricted set of chemical constituents often discovered through trial and error. Here, we report a high-throughput experimental framework for the discovery of new perovskite single crystals. We use machine learning (ML) to guide the sequence of ever-improved robotic synthetic trials. We perform high-throughput syntheses of perovskite single crystals with a protein crystallization robot and characterize the outcomes with the aid of convolutional neural network-based image recognition. We then use an ML model to predict the optimal conditions for the synthesis of a new perovskite single crystal, enabling us to report the first synthesis of $(3\text{-PLA})_2PbCl_4$.This material exhibits strong blue emission, illustrating the applicability of the method in identifying new optoelectronic materials.

INTRODUCTION

Among strategies used to improve performance in perovskite optoelectronic devices, compositional engineering has been particularly effective;[1–3] however, of the full chemical space available, only a relatively few elements and chemicals have been employed. Despite rapid improvements in performance, the best perovskite devices are still reliant on lead, and suffer from questions regarding long-term operational stability.[4–8] There thus remains significant interest in developing new, environmentally benign, stable perovskite materials. Density functional theory (DFT) has been used to predict new perovskite compounds in recent years,[9–11] but the number of compounds explored is orders of magnitude below what the chemical space permits.[12] In addition, it is often difficult to synthesize theoretically stable materials: as an example, chloride-based perovskites may be used as blue emitters but are underexplored in comparison with the iodide and bromide counterparts. This is in part due to the sensitivity of chloride-perovskites to synthesis conditions.[13]

High-throughput experimentation (HTE) has been applied to perovskites for the evaluation of the stability of wide-band-gap materials,[14] and more recently the discovery of new perovskite compounds.[15] HTE rapidly screens synthetic conditions but is often limited by the rate at which the experiments can be properly classified.

Machine learning (ML) has emerged as an attractive technique to augment and accelerate DFT calculations[16,17] and has seen increasing use in the prediction of new perovskites and their properties.[12,15,18–20] In general, there are two approaches to using ML to predict materials properties: in the first, the database with which the

Progress and Potential

Metal halide perovskites have seen significant interest for their applications in optoelectronic devices. Perovskites are tertiary and quaternary compounds with an exceptionally large chemical space, yet perovskite development has largely been limited to a restricted set of chemical constituents. Here, we report a high-throughput experimental framework for the discovery of new perovskite single crystals. We use machine learning to both characterize our products and guide the next cycle of experiments. As a result, we synthesize single crystals of a new perovskite, which we show has strong blue emission. This work demonstrates the efficacy of combined high-throughput experimentation and machine learning for accelerated materials discovery.

ML models are trained is generated by computations; in the second, experiments are used to generate the database.

Here, we present the high-throughput synthesis and ML-aided analysis of perovskite single crystals for optoelectronic applications. Single crystals of perovskites have been shown to possess long diffusion lengths and low trap-state densities[21] and have shown improved stability in optoelectronic devices.[22,23] Single crystals also enable the study of fundamental properties of materials without contributions from grain boundaries or thin-film formation dynamics and are thus desirable when screening new materials. The vapor-assisted antisolvent method of synthesizing perovskite single crystals is relatively fast and can be executed at low temperatures;[21] however, this method is sensitive to the concentration of precursors, the combinations of solvents chosen, and the type of antisolvent used. When attempting to synthesize a new material, it is necessary to use a fine mesh of synthetic combinations: for a single new perovskite compound, we estimate ~560 experiments are necessary to broadly explore the synthetic space (20 concentrations, four different antisolvents, and seven combinations of three solvents). This number grows exponentially as different precursors are considered. Growth of new perovskite single crystals is thus well suited for acceleration with HTE and ML.

Using a protein drop setter,[24] we prepared 96 independent crystallization conditions in a matter of minutes, and imaged, autonomously, the crystals at different stages of growth. With a dataset of 7,000 images, we trained a convolutional neural network (CNN) to recognize whether crystals had been grown. This classifier was then used on a broader dataset of 25,000 images. This combined high-throughput and ML classification enabled the rapid exploration of an experimental space for a given chemical space. Using the results from this exploration, we then trained an ML model to guide future experiments for crystal growth of a new perovskite: the ML model takes a given chemical space and returns the experimental parameters most likely to result in a successful crystal growth. We used this model to synthesize a new perovskite-like material with a direct band gap at 4 eV that emits at blue wavelengths. The synthetic conditions that resulted in successful crystal growth were restrictive, and it is only with the finely tuned experimental net from HTE combined with the accelerated characterization that we were able to discover the optimal conditions.

RESULTS AND DISCUSSION

High-Throughput Experimentation

We used the antisolvent vapor-assisted crystallization method[21] to grow single crystals, as this method yields high-quality single crystals in a relatively short time and can be carried out at room temperature. We prepared solutions containing the perovskite precursors and then sealed these in an environment containing an antisolvent, the vapor of which slowly diffuses into the precursor solution and induces crystallization (Figure 1A). A schematic of the trays used by the drop setter is shown in Figure 1B. Each tray contains 96 containers, and each container contains a large well and three smaller drops. The antisolvent is added to the well, and the precursor solution is added to the drops. Notably, the precursor solution can be combinatorially drawn from up to eight different starting sources: this allowed us to readily combine different precursor solutions with different solvents, allowing a fast and systematic exploration of different concentrations in different solvents, an important consideration.[21] This experimental design is based on a commonly used method for protein crystallization in structural biology.

[1]Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada

[2]Princess Margaret Cancer Centre, University Health Network, Toronto, ON M5G 1L7, Canada

[3]Department of Medical Biophysics, University of Toronto, Toronto, ON, Canada

[4]These authors contributed equally

[5]Lead Contact

*Correspondence: ted.sargent@utoronto.ca

https://doi.org/10.1016/j.matt.2020.02.012

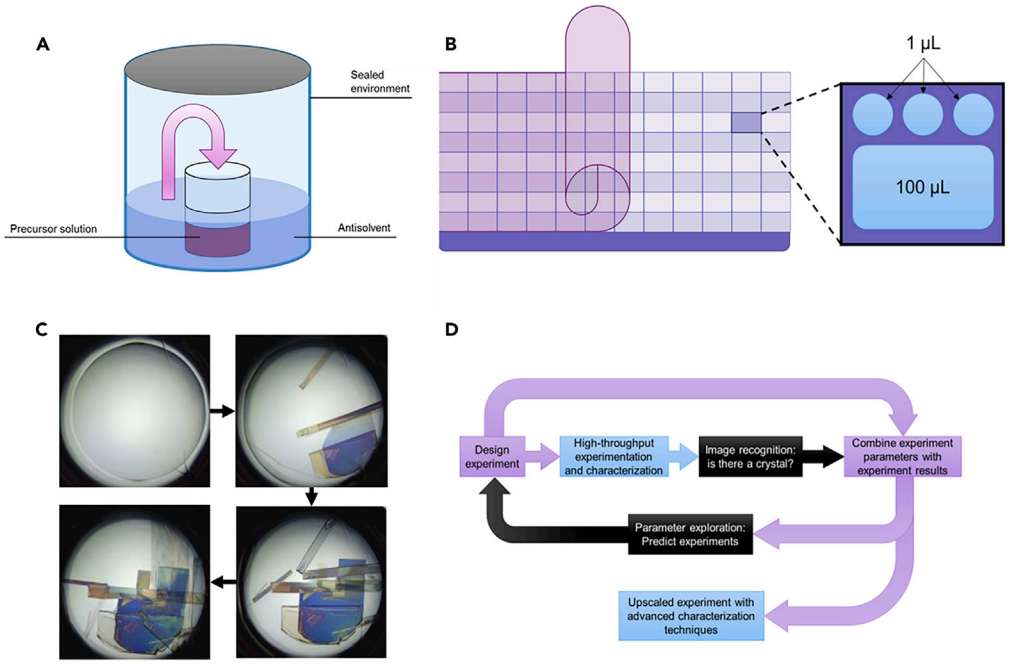


Figure 1. The High-Throughput Experimentation and Classification Process

(A) The antisolvent vapor-assisted crystallization method.

(B) Diagram of the 96-cell trays used for crystallization experiments.

(C) A 3-h time lapse of PEAPbBr crystal formation.

(D) The full ML pipeline for accelerated material discovery.

We focused our efforts on two-dimensional (2D) perovskites, because these materials have recently shown excellent photoluminescent quantum yields[25] and promise regarding improved stability in comparison with their bulk counterparts;[6,26,27] these materials pose an additional challenge because of the high dimensionality of the combinatorial chemical space in comparison with their three-dimensional (3D) counterparts. We focused initially on a known perovskite, phenethylammonium lead bromide, $(C_8H_{12}N)_2PbBr_4$ (PEAPbBr). We sought, as proof of principle, to replicate the growth of single crystals of PEAPbBr that we have previously reported,[25] and thus prepared several experiments with different conditions for this material.

Crystallization times varied from a few hours to a few days depending on the combination of experimental parameters. For PEAPbBr, crystallization typically occurred within a matter of hours. We obtained high-quality crystals and imaged them autonomously (Figure 1C). In addition to determining whether a crystal formed, we wanted to probe the photoluminescence of the crystals. We installed a UV light-emitting diode in the imaging system and imaged the crystal while illuminating it with a UV source; the PEAPbBr crystals showed strong luminescence (Figure S1), and using a color (RBG) imaging array, we estimated the central emission wavelength and relative brightness of different crystals (Figure S2).

After demonstrating that the protein crystallization robot and imaging system could successfully be employed for perovskite single-crystal growth, we developed an experimental pipeline for the discovery and synthesis of new perovskites. We first chose a desired chemical space to explore and designed an initial set of parameters to investigate. We employed HTE with the robot and then used a CNN to classify each experiment as a success (i.e., a crystal grew) or a failure (i.e., no crystal). Using

the parameters that correspond to each successful or unsuccessful experiment, we used a k-nearest-neighbors regression model to explore the next set of experiments that were likely to yield successful crystal growth, which were then carried out by the robot. The pipeline is illustrated in Figure 1D.

Image Recognition for Crystal Classification

The ML thrusts of this pipeline are two-pronged: the first involves training a CNN that can recognize crystals from non-crystals, and the second aims to train an ML model that relates experimental parameters of crystal growth to the likelihood of success of the crystal growth. Because we had both successful and unsuccessful experiments with PEAPbBr crystals, we used these experiments to train the neural net.

Seven thousand images of PEAPbBr experiments were collected on 576 starting conditions from six trays. There were three types of results from these experiments: no crystallization; a polycrystalline precipitate was formed; or a perovskite single crystal was formed.

These images were labeled into two categories: bad crystals or no crystals (i.e., wells with no precipitate, wells with polycrystalline materials, or wells with small poorly formed crystals, Figures S1A–S1C) and good crystals (i.e., large crystals even in the presence of some polycrystalline product, Figures S1D and S1E). Before preprocessing each image, we further categorized the dataset by container and split the data into train, validation, and test sets, ensuring that each container was only present within one of those sets. This accounted for minimal differences in acquired images from the same container at different times; otherwise the performance of a trained network could appear artificially high due to the validation and test sets containing data from the same containers used to train the network (i.e., an almost duplicate image from a different time step in the same container would be present in the validation and test sets).

We augmented the data before training, as this has been shown to improve image recognition of protein crystals with neural networks (NNs).[28] We augmented random sets of images chosen from the initial dataset by adjusting the brightness values by ± 32 (out of 255), adjusting the saturation values from 50% to 150%, adjusting the hue values by ± 70 (out of 180), and rotating the image by either 90°, 180°, or 270°. We kept both the augmented and initial images for training: by applying augmentations proportionally between positive and negative samples, we obtained a 1:1 ratio of images with a crystal and images without a crystal (i.e., we generated more augmented images for the class that had fewer initial samples). Augmented images were kept in the same set (training, validation, or testing) as their parent image.

After the dataset had been generated, we calculated the spatial Fourier transform (FT) for each image. We reasoned that crystals would possess a certain length scale that would be readily detected in Fourier space. We passed the FT of each image as an additional input to the CNN to ensure that any spatial frequency information was readily available as an input (Figure 2A). We trained a separate channel for each of the RBG and FT images for each data point, reasoning that as they have a different feature space, the architecture that maximizes the probability of a correct prediction would also be different. The CNN architecture is also depicted in Figure 2A.

An extensive hyperparameter search was conducted on both a single-branch architecture (RGB) and a dual-branch architecture (RGB + FT) CNN. Using identical

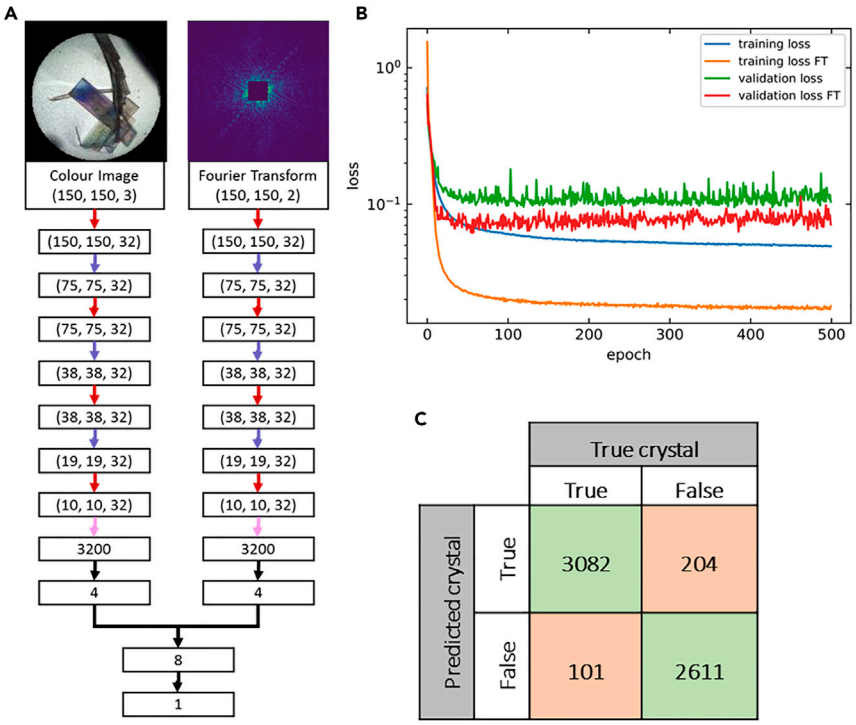


Figure 2. Development of the CNN for Crystal Recognition

(A) The CNN architecture for crystal image recognition split into two branches for color images and their corresponding Fourier transform.

(B) Results of training the best CNN versus a single-branch architecture, only taking a color image as input.

(C) Confusion matrix of the best CNN against a test dataset.

training, validation, and test datasets, the best RGB + FT network was ~5% more accurate than that without the FT (Figure 2B) and was thus chosen as the crystal recognition algorithm. The optimized NN recognizes crystal formation with 95% accuracy, and its overall performance is summarized in the confusion matrix for the test set in Figure 2C. The complete results of the training along with comparisons between both types of networks can be seen in Figures S2–S4.

The addition of the FT as a separate channel significantly improved the performance of the model. When only the RGB values were used to train an NN, the hyperparameters did not converge during optimization (Figure S2). Fourier space information assists in obtaining a reliable model and enabled us to apply the model to a new set of crystals.

Predictive ML Model

Once we successfully trained a CNN to recognize crystals, we sought to address the second thrust of the ML algorithm: parameter space exploration to guide subsequent experiments. We also sought to expand the chemical space to new perovskite materials. We continued to focus on precursors of 2D perovskites but wished to synthesize a new material that had not previously been made: we chose to use 3-picolylammonium (3-PLA) as the bulky ligand in the 2D perovskite structure, as we reasoned that the amine-substituted benzyl ring would have chemical properties different from those of the traditional PEA ligand.

Table 1. Input Features for Parameter Exploration

Feature	Constituents/Description
DFT features	the potential, Fermi energy, HOMO, HOMO-1, LUMO, LUMO-1, total energy, and polarizations for each of the A-ligand, A, B, and X sites of the perovskite
Concentrations	one feature each for the A-ligand, A, B, and X sites of the perovskite
Solvent ratios	the DMSO/DMF/GBL ratio
Volume of liquid	the volume of antisolvent in each of the 100-μL wells and the volume of precursor solution in each of the 1-μL well
Time	a time for when the crystal is characterized
Crystallization	whether or not there is a crystal present in the well at a given time
Visible photoluminescence	whether or not there is photoluminescence in the visible spectrum at a given time

HOMO, highest occupied molecular orbital; LUMO, lowest unoccupied molecular orbital.

We explored a chemical space with different precursors, solvents, and antisolvents, as shown in Table S1. The experimental parameters have significant impact on the crystallization of perovskites: for PEAPbBr crystals grown with same antisolvent (isopropanol [IPA]), strikingly different products are obtained for different dimethylformamide (DMF)/dimethyl sulfoxide (DMSO) ratios (Figure S6). Changing the concentration greatly affects the quality of the resulting crystal. In the containers in which the precursors were dissolved only in DMF, crystals formed at higher concentrations are larger but impure, with crystallites forming throughout. At low concentrations not enough crystallization occurs, and this results in formation of small patches of crystals. The large gradient in crystal quality as a function of synthetic parameters highlights the need to use HTE and ML to probe the experimental space intelligently.

When we began to synthesize $(3\text{-PLA})_2PbX_4$ (PLAPbX; X = Cl, Br, I), we found it difficult to obtain crystals. There are optimal concentrations at which crystals form rather than a continuous trend. Crystals grown with IPA are obtained at concentrations of 0.5 M and 0.25 M, but not at 0.33 M (Figure S7). Our initial experiments had a very low success rate (~1%), and the crystals that did grow often took ~72 h. As this rate and time contrasted with those of PEAPbBr crystals, we aimed to accelerate the crystal discovery process with a predictive ML model. We note that our initial attempts at synthesizing these crystals in bulk were unsuccessful.

We fingerprinted each experiment by using a combination of parameters (Table 1). The parameters included DFT features for each A-site, B-site, and X-site component, concentrations of each component, solvent ratios (including γ-butyrolactone [GBL]), volume of the precursor solution added, and the time since the experiment began. The DFT features chosen are those readily available from the Vienna Ab initio Simulation Package and are a subset of those employed when using Voronoi tessellation.[12,29] Using properties from standard DFT accelerated the workflow, avoiding reliance on expensive calculations. We compared different ML models[30] and found that an optimized k-nearest-neighbors classifier had the best accuracy (88%, Table S2). The model was trained on 300 experimental data points. Only distinct experiments were used (i.e., different time steps were not used as separate data points). The model hyperparameters were optimized using Tpot.[31]

Using the ML regressor, we mapped the likelihood of crystallization onto the experimental space of both PEAPbBr and PLAPbX. We note that the ML model was the same for both PEAPbBr and PLAPbX; different chemical spaces input to the model yielded different experimental spaces to explore. This eliminates the

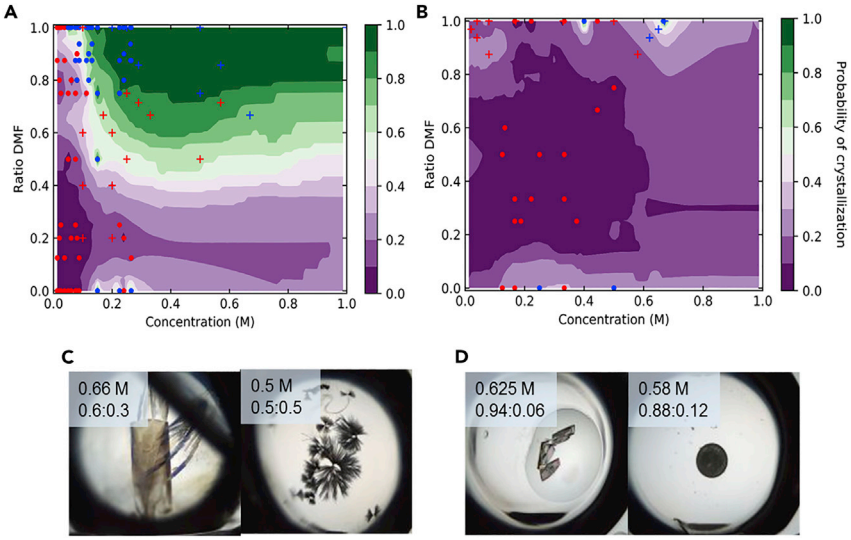


Figure 3. ML Exploration of the Experimental Space for New Perovskite-like Compounds

(A and B) Results of the parameter exploration for the (A) $(PEA)_2PbBr_4$ chemical space and (B) $(3\text{-}PLA)_2PbBr_4$ chemical space crystallized using IPA and only a mixture of DMSO and DMF as a solvent. The blue and red dots indicate crystallization and no crystallization, respectively, and were used in the training of the algorithm, whereas the crosses are from a new experiment.

(C and D) Images of a successful (left) and unsuccessful (right) outcome for the (C) $(PEA)_2PbBr_4$ and (D) $(3\text{-}PLA)_2PbCl_4$ experiments.

need to train a new model to explore new chemical spaces. Subsets of this chemical space are shown in Figure 3. Figure 3A shows the likelihood of crystallization for PEAPbBr with IPA as the antisolvent, shown as a function of DMSO and DMF ratios. The blue dots show experiments that were successful whereas the red dots show experiments that were unsuccessful. The crosses indicate experiments that were not included in the training set of the regressor.

Figure 3B shows the mapping of PLAPbBr crystals' experimental space; we note that prior to running this algorithm we had fewer than five successful experiments with PLAPbBr in more than 288 attempts. However, with ML-aided experimental design we were able to double the number of successful experiments with just one experimental cycle (~12 successful experiments in 96 attempts). The predictions indicated a very small chemical space in which crystallization was likely to occur, and the suggested experiments were conducted in the chemical space of higher probability.

We were also able to synthesize PLAPbCl crystals (Figure S7). After running the ML algorithm, we obtained high-quality crystals from the robotic synthesis; we therefore sought to upscale the synthesis for further analysis. Figure 4 shows the results of the synthesis. The new perovskite material has a direct band gap at 4 eV and has strong photoluminescence intensity centered at ~2.6 eV. We measured the single-crystal X-ray diffraction (XRD) of the newly synthesized material: interestingly, the predicted 2D structure was not formed but, instead, we obtained a perovskite-like structure in which the layered Pb-Cl framework is alternately intercalated and bonded with the 3-PLA ligand (Figure S8). This new structure appears to be more readily grown using chlorine compared with using bromine. This contrasts with traditional 3D and 2D perovskites, in which the Br^- and I^- variants are favored over the Cl^- variants. We obtained powder XRD of the crystals and found that the pattern matches the simulated XRD pattern from the crystallographic information file.

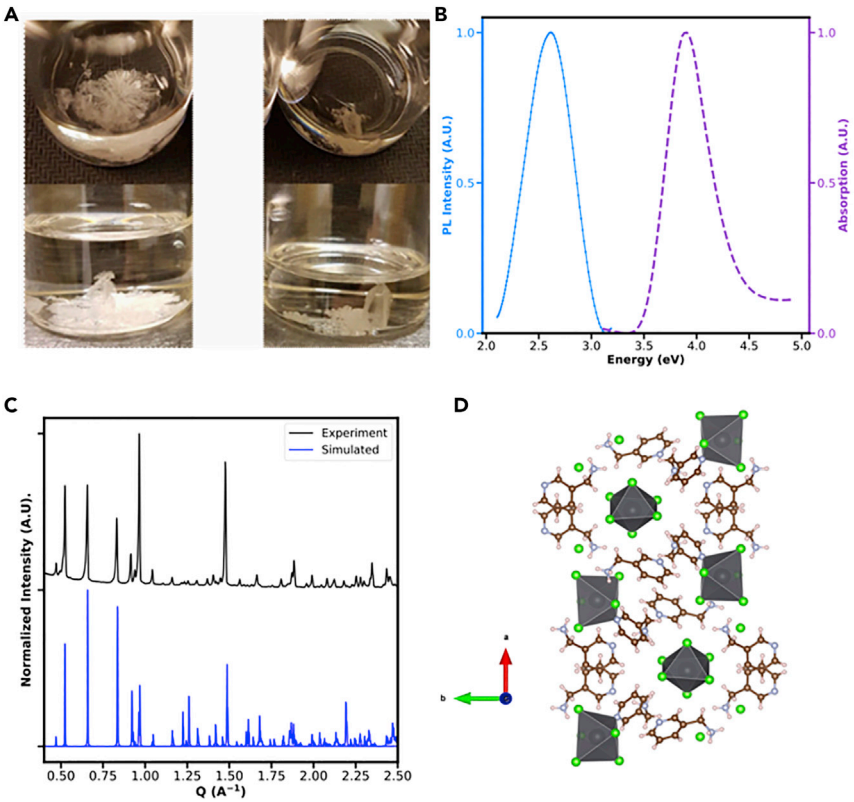


Figure 4. Bulk Synthesis of PLAPbCl Crystals

(A) Images of the synthesized crystals.

(B) Absorption and PL intensity of the single crystal, showing a direct bandgap at ~4 eV.

(C) Powder XRD of the obtained crystals, in comparison with the simulated diffraction pattern extracted from the single-crystal XRD. The experimental powder diffraction and simulated patterns are in good agreement.

(D) The refined crystal structure of the new perovskite-like material viewed along the c axis.

Conclusions

We developed a framework for high-throughput synthesis and characterization of perovskite single crystals using the antisolvent method. We first used a known synthesis to train a CNN for crystal recognition to allow autonomous characterization of the outcomes of the robotic experiments. High-throughput experiments were carried out by a facile method using a protein crystallization robot, and images were captured autonomously using a commercial robot. We used the high-throughput capabilities to generate a dataset relating the experimental and chemical parameter space to crystal growth and used this dataset to train a tree-based ML regressor. Using the ML regressor, we synthesized crystals of a new perovskite-like material, $(3\text{-PLA})_2PbCl_4$. We then successfully performed a scaled-up synthesis of the $(3\text{-PLA})_2PbCl_4$ perovskite and found that it has a direct band gap at 4 eV with blue emission. This work presents a new route to accelerate the discovery of new perovskites.

EXPERIMENTAL PROCEDURES

CNN

Image augmentation was done using the OpenCV Python library.[32] We implemented both approaches in the TensorFlow-[33]powered python module Keras.[34] Previous implementation of both methodologies to atomistic systems was based

on a very widely used VGGNet[35] architecture that was successfully used in image classification of the 1000-class ImageNet 2012 database. Following that approach, we optimized the hyperparameters of our VGGNet-like network, namely: the size of the 3D convolutional filters, the number of filters, the presence of a reducing convolutional layer, regularization, dropout, and type of activation function. These parameters were optimized using kopt[36] (forked from the original hyperopt)[37] Python library.

ML Regressor

We represented each experiment with a feature vector as outlined in Table 1. We optimized our ML algorithms using a genetic-algorithm optimization of the architecture of a random forest-based regressor as implemented in the tpot[31] Python module. We used five generations with a population size of 20 architectures for all representations.

DATA AND CODE AVAILABILITY

All experimental data and source code are available upon request.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.matt.2020.02.012.

ACKNOWLEDGMENTS

E.H.S. and all co-authors from the Department of Electrical and Computer Engineering at the University of Toronto acknowledge the financial support from the Ontario Research Foundation—Research Excellence Program and the Natural Sciences and Engineering Research Council of Canada (NSERC). The authors gratefully acknowledge Dr. Isaac Tamblyn's fruitful discussion and advice. This work was also supported by an NSERC grant to G.G.P.

AUTHOR CONTRIBUTIONS

J.K. and E.H.S. conceived the idea. A.J., J.K., D.M., and D.A.K. carried out the tray preparation and curated the data. J.K. developed the CNN. P.T. assisted with the development of the CNN. A.J. and J.K. developed the ML regressor. P.T. assisted with the development of the CNN. Y.G. and A.J. synthesized the bulk crystals. A.J. carried out bulk crystal characterization. M.A., G.G.P., and E.H.S. supervised the project.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: October 23, 2019
Revised: January 24, 2020
Accepted: February 14, 2020
Published: March 10, 2020

REFERENCES

1. Jeon, N.J., Noh, J.H., Yang, W.S., Kim, Y.C., Ryu, S., Seo, J., and Seok, S.I. (2015). Compositional engineering of perovskite materials for high-performance solar cells. Nature *517*, 476–480.

2. Yang, W.S., Park, B.-W., Jung, E.H., Jeon, N.J., Kim, Y.C., Lee, D.U., Shin, S.S., Seo, J., Kim, E.K., Noh, J.H., and Seok, S.I. (2017). Iodide management in formamidinium-lead-halide–based perovskite layers for efficient solar cells. Science *356*, 1376–1379.

3. Saliba, M., Matsui, T., Seo, J.-Y., Domanski, K., Correa-Baena, J.-P., Nazeeruddin, M.K., Zakeeruddin, S.M., Tress, W., Abate, A., Hagfeldt, A., et al. (2016). Cesium-containing triple cation perovskite solar cells: improved stability, reproducibility and high efficiency. Energy Environ. Sci. *9*, 1989–1997.

4. Giustino, F., and Snaith, H.J. (2016). Toward lead-free perovskite solar cells. ACS Energy Lett. *1*, 1233–1240.

5. Khalfin, S., and Bekenstein, Y. (2019). Advances in lead-free double perovskite nanocrystals, engineering band-gaps and enhancing stability through composition tunabilty. Nanoscale *11*, 8665–8679.

6. Quan, L.N., Yuan, M., Comin, R., Voznyy, O., Beauregard, E.M., Hoogland, S., Buin, A., Kirmani, A.R., Zhao, K., Amassian, A., et al. (2016). Ligand-stabilized reduced-dimensionality perovskites. J. Am. Chem. Soc. *138*, 2649–2655.

7. Wang, R., Mujahid, M., Duan, Y., Wang, Z.-K., Xue, J., and Yang, Y. (2019). A review of perovskites solar cell stability. Adv. Funct. Mater. *29*, 1808843.

8. Abate, A. (2017). Perovskite solar cells go lead free. Joule *1*, 659–664.

9. Jain, A., Voznyy, O., and Sargent, E.H. (2017). High-throughput screening of lead-free perovskite-like materials for optoelectronic applications. J. Phys. Chem. C *121*, 7183–7187.

10. Nakajima, T., and Sawada, K. (2017). Discovery of Pb-free perovskite solar cells via high-throughput simulation on the K computer. J. Phys. Chem. Lett. *8*, 4826–4831.

11. Zhao, X.-G., Yang, D., Ren, J.-C., Sun, Y., Xiao, Z., and Zhang, L. (2019). Rational design of halide double perovskites for optoelectronic applications. Joule *2*, 1662–1673.

12. Askerka, M., Li, Z., Lempen, M., Liu, Y., Johnston, A., Saidaminov, M.I., Zajacz, Z., and Sargent, E.H. (2019). Learning-in-Templates enables accelerated discovery and synthesis of new stable double perovskites. J. Am. Chem. Soc. *141*, 3682–3690.

13. Unger, E.L., Bowring, A.R., Tassone, C.J., Pool, V.L., Gold-Parker, A., Cheacharoen, R., Stone, K.H., Hoke, E.T., Toney, M.F., and McGehee, M.D. (2014). Chloride in lead chloride-derived organo-metal halides for perovskite-absorber solar cells. Chem. Mater. *26*, 7158–7165.

14. Chen, S., Hou, Y., Chen, H., Tang, X., Langner, S., Li, N., Stubhan, T., Levchuk, I., Gu, E., Osvet, A., and Brabec, C.J. (2018). Exploring the stability of novel wide bandgap perovskites by a robot based high throughput approach. Adv. Energy Mater. *8*, 1701543.

15. Sun, S., Hartono, N.T.P., Ren, Z., Oviedo, F., Buscemi, A., Layurova, M., Chen, D.X., Ogunfunmi, T., and Buonassisi, T. (2019). Accelerated development of perovskite-inspired materials via high-throughput synthesis and machine-learning diagnosis. Joule *3*, 1437–1451.

16. Pilania, G., Wang, C., Jiang, X., Rajasekaran, S., and Ramprasad, R. (2013). Accelerating materials property predictions using machine learning. Sci. Rep. *3*, 2810.

17. Brockherde, F., Vogt, L., Li, L., Tuckerman, M.E., Burke, K., and Muller, K.R. (2017). Bypassing the Kohn-Sham equations with machine learning. Nat. Commun. *8*, https://doi.org/10.1038/s41467-017-00839-3.

18. Odabaşı, Ç., and Yildirim, R. (2019). Performance analysis of perovskite solar cells in 2013-2018 using machine-learning tools. Nano Energy *56*, 770–791.

19. Howard, J.M., Tennyson, E., Neves, B., and Leite, M. (2019). Machine learning for perovskites' reap-rest-recovery cycle. Joule *3*, 325–33713.

20. Im, J., Lee, S., Ko, T.-W., Kim, H.W., Hyon, Y., and Chang, H. (2019). Identifying Pb-free perovskites for solar cells by machine learning. NPJ Comput. Mater. *5*, 1–8.

21. Shi, D., Adinolfi, V., Comin, R., Yuan, M., Alarousu, E., Buin, A., Chen, Y., Hoogland, S., Rothenberger, A., Katsiev, K., et al. (2015). Low trap-state density and long carrier diffusion in organolead trihalide perovskite single crystals. Science *347*, 519–522.

22. Chen, Z., Turedi, B., Alsalloum, A.Y., Yang, C., Zheng, X., Gereige, I., AlSaggaf, A., Mohammed, O.F., and Bakr, O.M. (2019). Single-crystal MAPbI3 perovskite solar cells exceeding 21% power conversion efficiency. ACS Energy Lett. *4*, 1258–1259.

23. Saidaminov, M.I. (2015). Planar-integrated single-crystalline perovskite photodetectors. Nat. Commun. *6*, 8724.

24. FORMULATRIX. (2016). Drop setter for protein crystallization - NT8® by FORMULATRIX®. https://formulatrix.com/protein-crystallization-systems/nt8-drop-setter/.

25. Gong, X., Voznyy, O., Jain, A., Liu, W., Sabatini, R., Piontkowski, Z., Walters, G., Bappi, G., Nokhrin, S., Bushuyev, O., et al. (2018). Electron-phonon interaction in efficient perovskite blue emitters. Nat. Mater. *17*, 550–556.

26. Hong, K. (2018). Low-dimensional halide perovskites: review and issues. J. Mater. Chem. C *6*, 2189–2209.

27. Proppe, A.H., Wei, M., Chen, B., Quintero-Bermudez, R., Kelley, S.O., and Sargent, E.H. (2019). Photochemically cross-linked quantum well ligands for 2D/3D perovskite photovoltaics with improved photovoltage and stability. J. Am. Chem. Soc. *141*, 14180–14189.

28. Bruno, A.E., Charbonneau, P., Newman, J., Snell, E.H., So, D.R., Vanhoucke, V., Watkins, C.J., Williams, S., and Wilson, J. (2018). Classification of crystallization outcomes using deep convolutional neural networks. PLoS One *13*, https://doi.org/10.1371/journal.pone.0198883.

29. Ward, L., Liu, R., Krishna, A., Hegde, V.I., Agrawal, A., Choudhary, A., and Wolverton, C. (2017). Including crystal structure attributes in machine learning models of formation energies via Voronoi tessellations. Phys. Rev. B *96*, https://doi.org/10.1103/PhysRevB.96.024104.

30. Li, Z., Najeeb, M.A., Alves, L., Sherman, A., Parrilla, P.C., Pendleton, I.M., Zeller, M., Schrier, J., Norquist, A.J., and Chan, E. (2019). Robot-accelerated perovskite investigation and discovery (RAPID): 1. Inverse temperature crystallization. ChemRxiv. https://doi.org/10.26434/chemrxiv.10013090.v1.

31. Olson, R.S., Bartley, N., Urbanowicz, R.J., and Moore, J.H. (2016). Evaluation of a tree-based pipeline optimization tool for automating data science. In GECCO '16: Proceedings of the Genetic and Evolutionary Computation Conference 2016, T. Friedrich, ed. (GECCO), pp. 485–492.

32. opencv-python. Wrapper package for OpenCV python bindings. https://pypi.org/project/opencv-python/.

33. Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., et al. (2016). TensorFlow: A system for large-scale machine learning.12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16), 265–283.

34. Chollet, F. (2015). Keras. http://keras.io/.

35. Simonyan, K., and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. ArXiv, 1409.1556v6. https://arxiv.org/abs/1409.1556.

36. Avsec, Ž.. kopt: Keras-hyperopt (kopt); Hyper-parameter tuning for Keras using hyperopt. https://github.com/avsecz/keras-hyperopt.

37. Bergstra, J., Yamins, D., and Cox, D.D. (2013). Making a science of model search: hyperparameter optimization in hundreds of dimensions for vision architectures. In ICML'13: Proceedings of the 30th International Conference on International Conference on Machine Learning, S. Dasgupta and D. McAllester, eds. (PMLR), pp. 115–123.