Machine Learning-Based Analysis of Molar and Enantiomeric Ratios and Reaction Yields Using Images of Solid Mixtures

Yuki Ide[1]*, Hayato Shirakura[2], Taichi Sano[2], Muthuchamy Murugavel[1], Yuya Inaba[2], Sheng Hu[1], Ichigaku Takigawa[1,3]*, and Yasuhide Inokuma[1,2]*

[1]Institute for Chemical Reaction Design and Discovery (WPI-ICReDD), Hokkaido University, Kita 21, Nishi 10, Kita-ku, Sapporo, Hokkaido 001-0021, Japan.

[2]Division of Applied Chemistry, Faculty of Engineering, Hokkaido University, Kita 13, Nishi 8, Kita-ku, Sapporo, Hokkaido 060-8628, Japan.

[3]Center for Innovative Research and Education in Data Science (CIREDS), Institute for Liberal Arts and Sciences, Kyoto University, 302 Konoe-kan, 69 Konoe-cho, Sakyo-ku, Kyoto 606-8315, Japan.

ABSTRACT: Visual observations are frequently used as a preliminary evaluation of the chemical contents of mixtures, but their accuracy largely depends on the observer's experience and intuition, which are difficult to share. Here, we report component ratio prediction using image-based machine learning (ML), which is applicable to analysis of various solid mixtures, such as mixtures of organics and inorganics, polymorphous crystals, and enantiomers. The trained model with 300 images could predict the sugar/dietary salt weight ratio from an image within 4% error. The ML prediction pipeline was shown to be broadly applicable to polymorphic glycine, D/L-tartaric acid, and four-component systems. As an application demonstration, we also used our ML system to analyze yield of a solid-state decarboxylation reaction. These results demonstrated that accumulation of researchers' experience derived from visual information can be shared as trained ML models and used as a quantitative analysis method.

INTRODUCTION

Quantitative analysis of chemical compounds is indispensable processes in chemical science. Although many precise analysis methods, such as nuclear magnetic resonance, high-resolution mass spectrometry, and optical spectroscopies, have played central roles in chemical laboratories, researchers still rely on visual observations with the naked eye for preliminary evaluation. Visual inspection is a quick and non-destructive analysis methods that requires no spectrometer. In the 19th century, Louis Pasteur succeeded in the optical resolution of a tartrate salt through careful visual observation of the crystals, which led to the discovery of molecular chirality.[1] While visual observation has acted as an impetus for historic breakthroughs, its accuracy largely depends on the observer's experience and intuition. For example, experienced researchers can sometimes estimate the approximate component ratio between the desired product and other products just by looking at a solid reaction mixture based on empirical knowledge of the color, texture, grain sizes, or transparency of the products (Figure 1a), but less experienced researchers cannot. Such differences in experience may influence the outcome of the chemical research, when thousands of similar samples are to be analyzed. The experience and intuition of researchers are, however, often difficult to explicitly formulate and thus contain considerable uncertainty, which makes visual observations subjective and unreliable.

Recently, machine learning (ML) has begun to be used in chemical science to identify and analyze the multidimensional correlations between explanatory variables (i.e., descriptors/features) and target variables, as evidenced by many successful cases, such as catalyst development,[2,3] reaction conditions,[4,5] and functional materials.[6-8] Because the prediction performance of ML systems generally depends on the information and representation they use as input, various potentially related factors or quantities associated with molecular structures, chemical reactivities, and spectroscopic measurements have been validated and used as descriptors.[9] The raw pixel values of a photograph have also been used as input variables in image-related tasks, as recently demonstrated in evaluation of nanoparticles[10,11] and crystal growth.[12,13] This success primarily relies on the recent advances in deep learning, in particular, the capability of *representation learning* to directly learn relevant feature representations from low-level sensory data such as photographs. Accordingly, visual inspections in experimental laboratories could potentially be replaced by image-based ML as a more accurate and reliable tool. However, a big problem lies in how to formulate the visual observations of human researchers and input them as data into ML in a consistently relevant way.[14] From the photograph in Figure 1a (right), researchers would instantly recognize that a glass flask is being held by a person and anticipate that the chemicals to be analyzed are inside the flask. Such visual perception is grounded not simply in the visual stimulus itself, but also in the rich contextual information from all the past experiences and knowledge. Developing a similar machine recognition system would be, if not impossible, time-consuming and require an astronomical number of diverse training images that can cover all potential contextual variations that researchers encounter in daily activities.

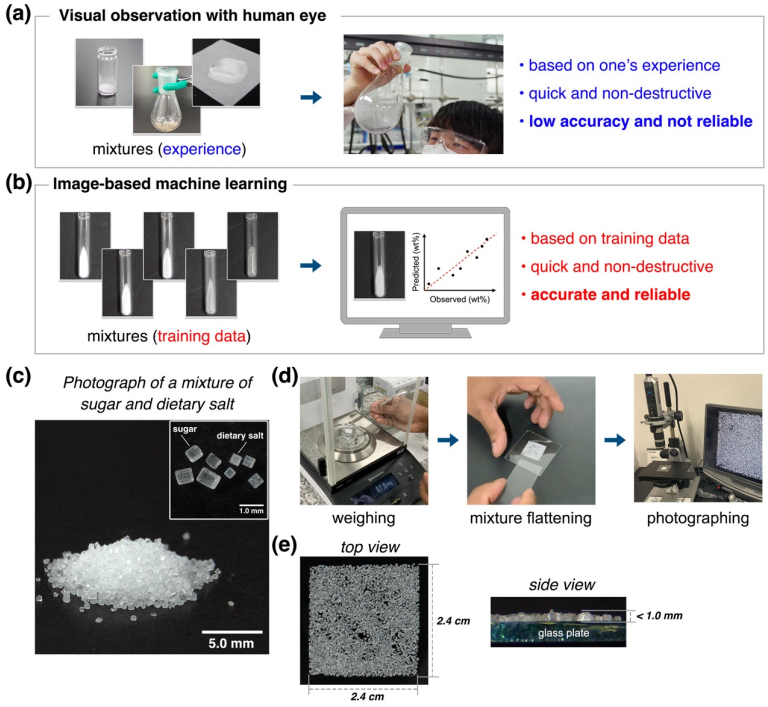


Figure 1. (a) Prediction of the mixing ratio of a reaction mixture by observation with the human eye. (b) Prediction of the mixing ratio of components by image-based ML. (c) Photograph of the mixture of sugar and dietary salt used in this experiment. The inset shows a magnified photograph of the sugar and dietary salt mixture. (d) Preparation procedure of the samples for the datasets. (e) Settings for photographing the samples.

To establish practical ML-based visual observation analysis, it is thus of great importance to frame the problem design in a reproducible, less uncertain, and robust form (Figure 1b) despite the inevitable variability and diversity of visual observations. Here, we report mixing-ratio prediction using image-based ML and its applicability to analysis of weight (molar) ratios, enantiomeric ratios, and reaction yields. The initial pilot case for the design, development, and optimization of the ML models was to predict the weight ratio between sugar (sucrose) and dietary salt (NaCl) from a given image of their mixture. The model achieved a prediction with a mean absolute error (MAE) of 3.9% for a physically independent test dataset. Suitable trained models for ML were obtained using 300 images as training data for each case. For reproducibility, we open-sourced the codebase, examined all randomness-included calculations multiple times, and carefully performed prediction assessment to avoid any potential data leakage, which also provides a web-based prediction system to enhance the usability for third-party assessment by non-experts. Using optimized conditions for ML model making, we were able to predict the molar (or weight) ratio of polymorphic glycine, D/L-tartaric acid, and four-component organic/inorganic mixture systems. The ML-based prediction system was also applied to analyze the reaction yield of solid state decarboxylation of *p*-aminosalicylic acid. To further expand the ease of applicability, we also show that this ML model trained with images taken by a research-level microscope can be transferred to the images recorded by a consumer-level mobile camera after fine-tuning calibration with a few images.

RESULTS AND DISCUSSIONS

An initial investigation of the image-based ML system targeted a pilot case study of predicting the weight ratios of given mixtures of sugar and dietary salt (Figure 1c). The sugar and dietary salt samples used in this experiment had averaged grain sizes of 0.46 and 0.41 mm, and averaged grain weights of 0.17 and 0.14 mg, respectively. Because the grain-size distributions showed considerable overlap (Figure S1 in the in the Supporting Information), it is difficult for the human eye to clearly distinguish these solids based on the size, color, and transparency of the image. An initial training dataset of a total of 100 mixtures of sugar and dietary salt with different weight ratios was prepared by weighing (Figure 1d): sugar/dietary salt (mg/mg) = 0:10, 1:9, 2:8, …, 9:1, and 10:0 with 9–10 samples for each ratio. Each well-blended sample with a total weight of 200 mg was spread on a glass plate inside a square area of 2.4 cm × 2.4 cm with a sample thickness of 1.0 mm so that all of the grains could be photographed without grain overlap (Figure 1e). To avoid the influence of the background, the sample glass plate was placed on a black stage, and a photograph was taken with a microscope camera at resolution of 2048 × 1536 pixels. The collected images were trimmed to a square shape of 1536 × 1536 pixels. To evaluate the prediction performance of the ML system, another set of 100 images was prepared and used as the test dataset to evaluate the error between the predictions and the actual weight ratios. To avoid any unintended data leakage, this test dataset was prepared as a physically independent dataset from the training dataset: for each image, a new sample was taken, weighed, mixed, and spread on a glass plate. Hence, the test dataset contained the natural batch effect from potential confounders. The prediction accuracy was evaluated by multiple standard metrics: the MAE, root-mean-squared error (RMSE), and coefficient of determination (R^2). A random subset of 10% of the training dataset was used as a validation dataset to choose the best model, and the remaining 90% of the training dataset was used for training, unless specified otherwise.

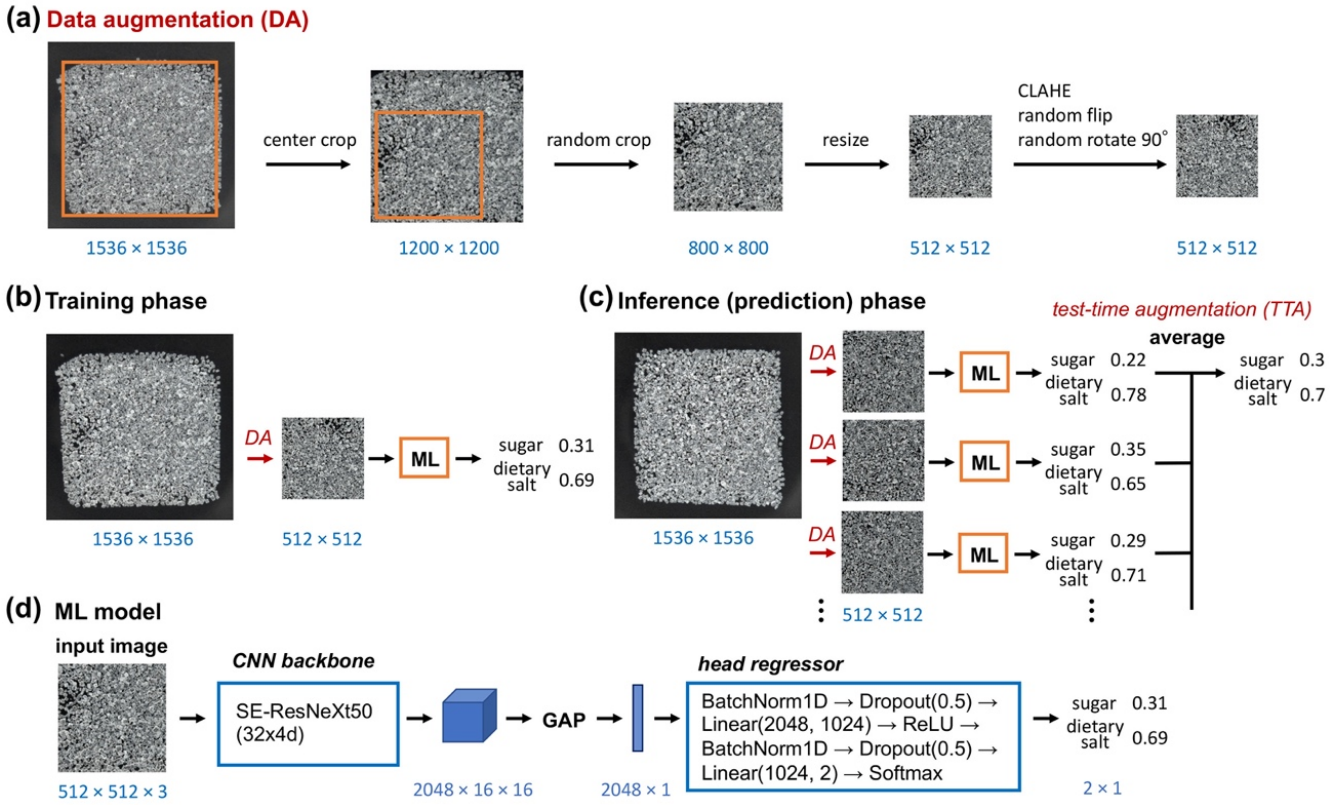


Figure 2. (a) Data augmentation (DA) process. Center cropping of 1200 × 1200 pixels, random cropping of 800 × 800 pixels, resizing to 512 × 512 pixels, CLAHE, random horizontal flipping with 50% probability, random vertical flipping with 50% probability, and random 90° rotation are sequentially applied every time an image is given to ML. (b) Training phase. The ML model is trained over subimages from DA. (c) Inference phase. DA is applied multiple times, and each augmented image is given to ML to obtain a predicted ratio. These multiple predicted ratios are averaged to give a single number as the final prediction. (d) ML model. Three-channel color 512 × 512 pixels are input into a convolutional neural network (CNN), and global average pooling (GAP), a head regressor by fully-connected layers in sequence, and finally softmax are applied to output the predicted ratios.

Considering the variability of visual observations, the ML model was designed as ensemble regression over randomly cropped subimages from an input image (Figure 2a–c), which is called *test-time augmentation* (TTA) in inference. The base model was thus trained over 512 × 512 pixel subimages associated with the same ground truth of parent images (Figure 2b). Each subimage was obtained from the original 1536 × 1536 input image by center cropping to 1200 × 1200 pixel size to remove the edges of the sample area, random cropping at different locations to 800 × 800 pixel size, random flipping and 90° rotation with a probability of 50%, contrast normalization by contrast-limited adaptive histogram equalization (CLAHE), and resizing to 512 × 512 pixels (Figure 2a). These sizes were fixed to obtain subareas that were sufficiently large to estimate the average weight ratio. To predict the weight ratio of a given 1536 × 1536 image, the same data augmentation procedure was applied multiple times to obtain several 512 × 512 subimages from a single input image, and the predicted weight ratios of each subimage were averaged to obtain the final predicted value of the weight ratio of the original input image (Figure 2c). The ML model had a standard architecture with a convolutional neural network (SE-ResNeXt-50)[15] backbone, global average pooling, and a head regressor by fully-connected layers with softmax outputs, and it was trained with soft cross-entropy loss (Fig. 2d, also see Methods). For the sugar/salt two-component example, the ML model output two values as predictions of the weight ratios of sugar and dietary salt in wt%, satisfying $0 \leq$ sugar (wt%) ≤ 1, $0 \leq$ dietary salt (wt%) ≤ 1, and sugar (wt%) + dietary salt (wt%) = 1.

The observed–predicted plots of the initial pilot case of weight-ratio prediction by the ML models are shown in Figure 3. When the ML model was directly trained with the input images (with only non-random preprocessing of center cropping to 1200 × 1200 pixels, resizing to 512 × 512 pixels, and CLAHE), the prediction performance was MAE = 10 %, RMSE = 11%, and $R^2 = 0.85$ (Figure 3a), presumably because of the small number of only 100 training images, and it also exhibited unstable prediction variability. Adding data augmentation by random flipping and 90° rotation improved the prediction accuracy to MAE = 7.1%, RMSE = 9.1%, and $R^2 = 0.90$ without TTA (Figure 3b). When the model was trained with data augmentation, a differently augmented image was generated for each epoch, and thus 12,000 images were given to ML during training of 120 epochs. We re-trained the ML model by further adding random cropping to the data augmentation, and we examined the TTA prediction performance of the obtained model with varying number of augmented images from 1 to 50 (Figure S2). As a result, 30-image TTA with MAE = 5.9%, RMSE = 7.7%, and $R^2 = 0.93$ was selected as the setup for further investigation, considering the balance between the efficiency and the significance of the accuracy improvement (Figure 3c).

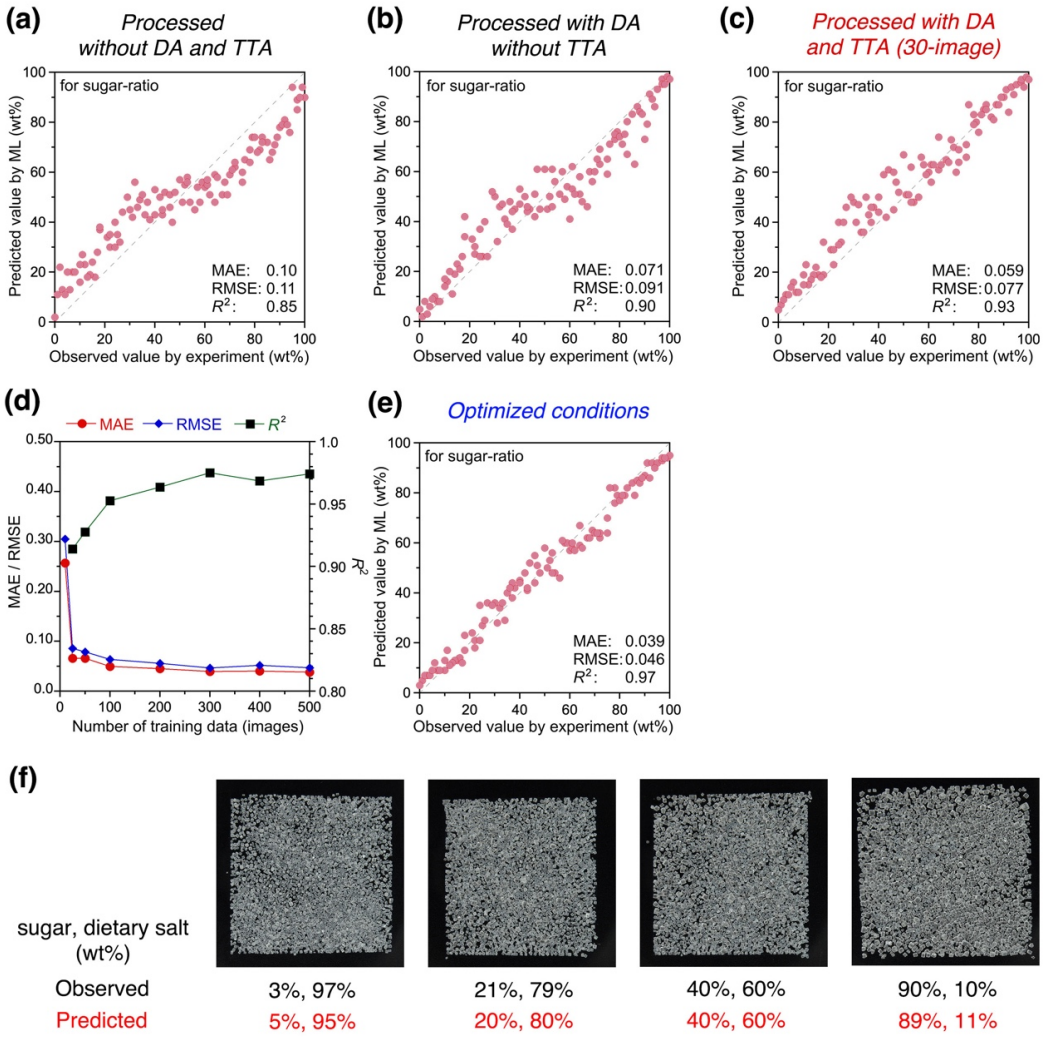


Figure 3. (a–c) Observed–predicted sugar-ratio plots for 100 test images using different data processing in ML: (a) without data augmentation (DA) and test-time augmentation (TTA), (b) with DA (except for random cropping) and without TTA, and (c) with DA and 30-image TTA. (d) Relationship between the training image data and the prediction accuracy. (e) Prediction of the mixing ratio under the optimized conditions. (f) Selected examples of the prediction of the mixing ratio under the optimized conditions.

We then optimized the photography conditions and number of training images. When the accuracy of ML-based analysis was evaluated using images with different sample densities (10, 20, and 35 mg/cm^2), the densest condition (35 mg/cm^2) showed the best MAE and RMSE values (Figure S3 and S4). The interval of the ground truth values in the training data also affected the accuracy. When 100 images of training data were used with 1wt% increments, the MAE and RMSE were lower than those with 10 and 20wt% increments (Figure S5 and S6). Although ML can complement data that have not actually been input during the training processes, preparing various mixing ratios for training was found to be effective for model construction.

Using the optimized conditions (30-image TTA, 35 mg/cm^2, and 1wt% increment), we investigated the relationship between the number of training images and the prediction accuracy to evaluate the cost–performance ratio. A suitable prediction model could not be constructed using only 10 images even with data augmentation, but using 25 images gave moderate-level prediction. The MAE and RMSE gradually decreased with increasing number of original images for training until 300 images, after which the MAE and RMSE became stable (Figure 3d and S7). Using 300 images of the sugar/dietary salt mixtures for the training dataset (Figure 3f), the accuracy of ML-driven prediction improved to MAE = 3.9%, RMSE = 4.6%, and R^2 = 0.97 (Figure 3e). The inference time for computing the predicted values for 100 images was 34.79 seconds with 30-image TTA on an NVIDIA A100 GPU (0.348 seconds per image), 7.52 seconds with one-image TTA (0.075 seconds per image) on a NVIDIA A100 GPU, and 20.25 seconds with one-image TTA on an AMD EPYC 7252 3.1GHz/8-core CPU (0.202 seconds per image).

We further compared the prediction accuracy and stability with human visual inspection by assigning the same prediction task to five examinees, including one expert engaged in actual image collection and four non-experts not related at all to this project (Figure S8). The five examinees were asked to estimate the weight ratios of 10 images of sugar/dietary salt mixtures randomly chosen from the test dataset, referring to the same 300 images as the training dataset for ML (see the Supporting Information). The results indicated that the prediction accuracy of the ML system was much more accurate and stable than human visual inspections, which greatly varied (MAE = 7.5%–25 %, RMSE = 9.3%–34 %, Figure S8).

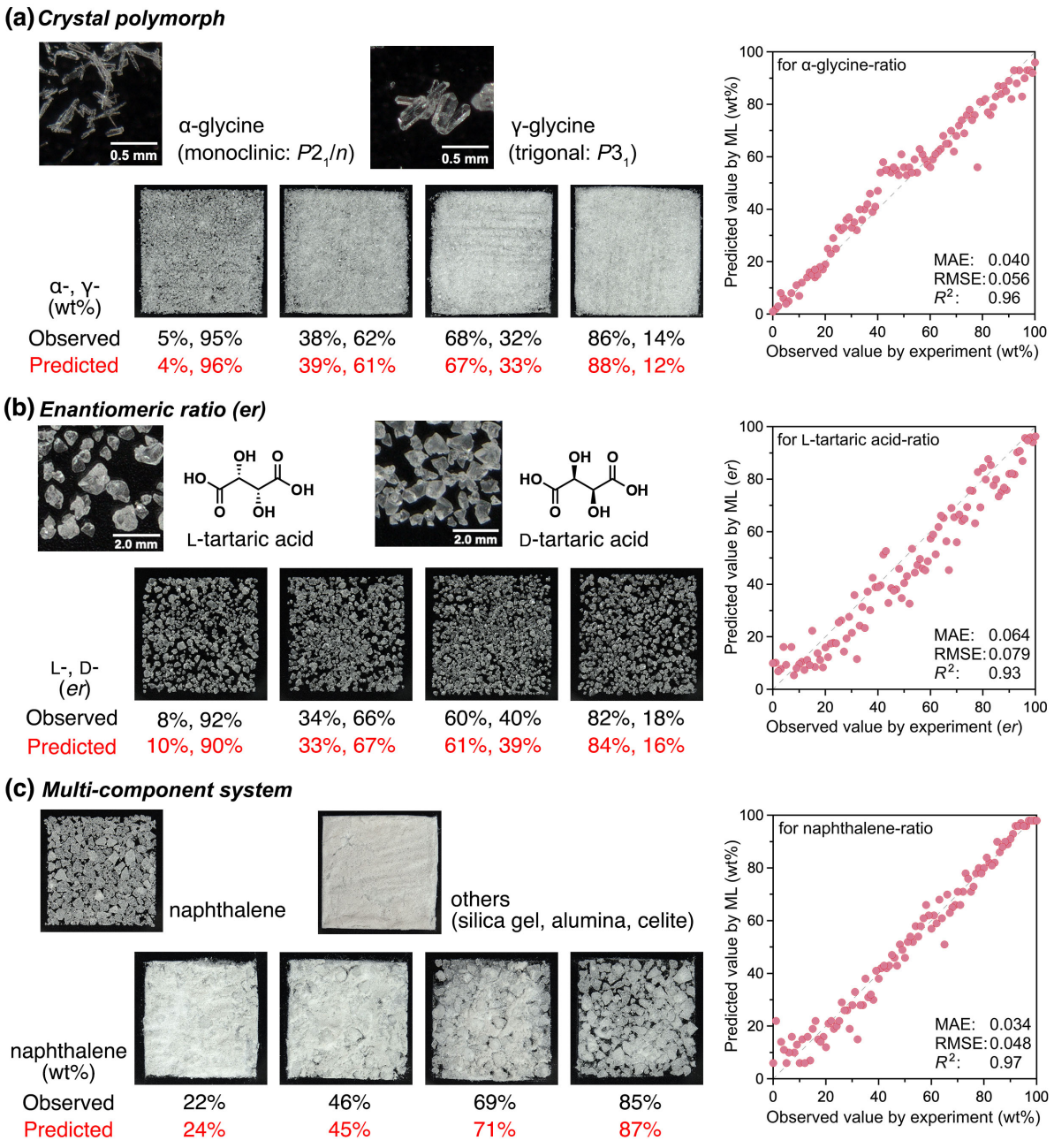


Figure 4. (a) Analysis of the mixing ratios of α- and γ-glycine. Typical crystal morphologies of the α- and γ-forms, observed–predicted plot, and representative results of ML-based prediction. (b) Prediction of the enantiomeric ratio (*er*) of mixtures of L- and D-tartaric acid. Crystals of each enantiomer, observed–predicted plot, and representative results of ML-based prediction. (c) Analysis of the naphthalene content of solid mixtures containing naphthalene, silica gel, alumina, and celite.

Next, we investigated the applicability of ML-based mixture analysis to different chemical solids. Crystal polymorphism is a phenomenon where different types of crystals are formed from an identical compound. Because the solubility of the crystals generally differs between two polymorphs, distinguishing polymorphs is of great importance in the pharmaceutical community.[16] When glycine is crystallized from neutral water, rod-like crystals of α-glycine in the monoclinic $P2_1/n$ space group exclusively form. Conversely, crystallization from hydrochloric acid gives a different polymorph[17]: γ-glycine as prismatic crystals in the trigonal $P3_1$ space group (see the Supporting Information). Using authentic solids of α- and γ-glycine, 300 images of mixtures with different weight ratios were taken as training dataset for ML. The ML pipeline for the initial pilot case was applied to this dataset, and its accuracy was evaluated using another physically independent set of 100 images as the test dataset. To our delight, the observed–predicted plot (Figure 4a) showed a good correlation with MAE, RMSE, and R^2 of 4.0%, 5.6%, and 0.96, respectively. It should be noted that methods for quantitative analysis of polymorphic mixtures are quite limited because they are composed of the same molecules. A reported method using infrared spectroscopy for evaluation of α- and γ-glycine has an uncertainty of 2%.[18] ML-based prediction showed that the ratio of polymorphs can be predicted at comparable accuracy to infrared spectroscopy using only a photograph of the mixture.

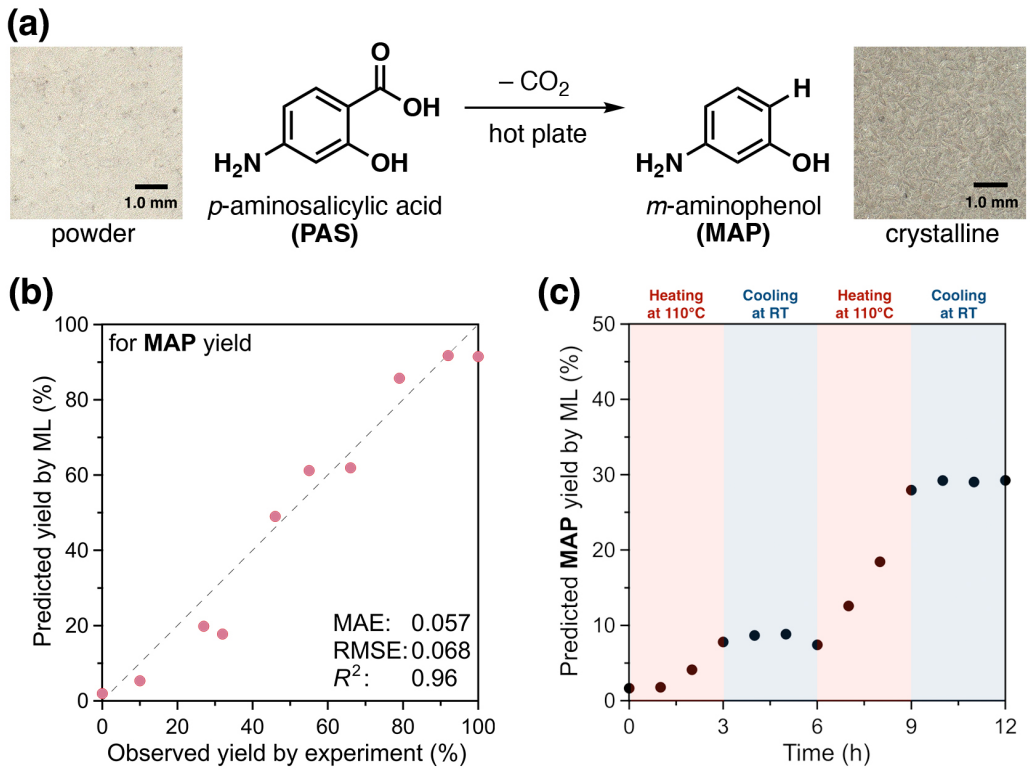


Figure 5. (a) Decarboxylation reaction of **PAS** to **MAP**. Insets show the changes in appearance. (b) Prediction of the **MAP** yield (%) from pictures of reaction mixtures. (c) Reaction monitoring using the ML-based yield prediction system under discontinuous heating conditions. Images of reaction mixtures were taken every 1 hour (5 images), and the averaged yields from 5 images at each time-point were used as the predicted **MAP** yield values in the graph.

Enantiomers, two stereoisomers in a mirror-image relationship, are one of the most difficult combinations to distinguish in molecular science. The purity of an enantiomer is described by the enantiomeric excess[19] or enantiomeric ratio (*er*),[20] an important index that is frequently used to evaluate the outputs in asymmetric synthesis. To analyze the *er* of a mixture of enantiomer crystals by image-based ML, we prepared 300 different mixtures of an enantiomer pair: L- and D-tartaric acids. Unlike sugar and dietary salt, the crystalline solid of L-tartaric acid shows almost the same color and size distribution as its enantiomer, D-tartaric acid. Our established ML pipeline excluding random flipping from data augmentation was applied to the 300 training images of the mixtures. The observed–predicted plot (Figure 4b) for another independent test dataset of 100 images showed that from the images of crystalline mixtures, image-based ML can predict the *er* with MAE, RMSE, and R^2 of 6.4%, 7.9%, and 0.93, respectively. Although the appearances of enantiomeric crystals are too confusing to quantitatively distinguish with the human eye, image-based ML can rapidly and quantitatively determine the *er*.

We further examined ML-based analysis of mixtures containing more than two solid components for future applications of ML-based analysis to estimation of the reaction yields using photographs of complex mixtures. As a model case, we prepared 300 mixture samples of naphthalene, silica gel, alumina, and celite with various weight ratios, and we used the 300 images as the training dataset. The weight ratio of naphthalene was assigned as the ground truth of each image, and the ML model was trained to predict the weight ratio of naphthalene. The observed–predicted plot (Figure 4c) of 100 independent test images showed that image-based ML can also be applied to multi-component mixture analysis, with MAE, RMSE, and R^2 of 3.4%, 4.8%, and 0.97 for prediction of the weight ratio of naphthalene, respectively.

Our ML-based system enables non-destructive analysis of solid-state chemical reaction. *p*-Aminosalicylic acid (**PAS**) is an antibiotic used in the medical treatment of tuberculosis, and is known to undergo thermal decarboxylation in the solid state to form *m*-aminophenol (**MAP**). This solid-state reaction results in continuous changes in physical appearance as the reaction progresses; when fine powder **PAS** was heated above 110 °C, crystalline of **MAP** gradually appears (Figure 5a). Considering practical cost performance, we collected 75 different images of reaction mixtures of **PAS** at different progress rates as a training dataset, and built an ML model for reaction yield prediction using the NMR-based reaction yield of **MAP** as ground truth for each image. Validation of the ML model was carried out using further images of reaction mixtures; when using an ML model built from 25 out of 75 training images, the prediction of reaction yields was failed. Another ML model based on 50 training images gave a better prediction score (Figure S15). With a ML model generated from 75 images, the best MAE, RMSE, and R^2 were determined to be 5.7%, 6.8%, and 0.96, respectively (Figure 5b).

Encouraged by the validation results, we conducted image-based reaction yield analysis of the **PAS** decarboxylation reaction. Solid **PAS** (*ca.* 200 mg) was placed on a glass plate, and heated with a hot plate using various temperature profiles (Figure S16). When the reaction was carried out at a constant

temperature, 110 °C, for 6 h, an almost monotonic increase in yield was observed as a function of reaction time. These changes were in good agreement with previously reported kinetic parameters of the **PAS** decarboxylation reaction.[21] Next, we examined a slightly more complicated temperature profile for ML-based reaction analysis; **PAS** was first heated at 110 ºC for 3 h, then the heating was stopped and left at room temperature for 3 h, and then reheated at 110 ºC for a further 3 h. We took 5 different images of reaction mixtures at each 1 h, and analyzed using the ML model (Figure 5c). The predicted yield appeared to increase during heating, while remaining approximately the same at room temperature. Reaction analysis using ^{1}H NMR spectroscopy also confirmed similar changes in reaction rate, demonstrating the reliability of ML-based reaction analysis. Although NMR spectroscopy is more accurate than our ML-based analysis, it requires sample preparation at each analysis step, *i.e.* taking a small portion of the sample to dissolve in a deuterated solvent. In this sense, ML-based analysis is more straightforward and non-destructive, because image data can be acquired during the reaction without contacting the sample. Thus, our ML-based system would be useful for quality management of pharmaceutical compounds that may degrade during storage.

Although image-based ML is applicable to analyzing images of various solid mixtures, dependence on the type of camera can be a common limitation. In fact, weight-ratio prediction was unsuccessful for the test images taken by a mobile camera even though the visual difference seemed to be negligible to the human eye. For example, in the case of sugar/dietary salt mixtures, the MAE, RMSE, and R^2 were 19%, 22%, and 0.58, respectively (Figure S18a), for six test images taken by a mobile camera when the ML model was trained with the 300 microscope images. This can be attributed to the subtle differences in the image resolution and brightness between the mobile camera and microscope camera. When the ML model pre-trained with 300 microscope images was fine-tuned by additional training with 24 mobile-camera images for calibration, the MAE, RMSE, and R^2 were improved to 6.0%, 6.4%, and 0.97, respectively (Figure S18b). The applicability of using different cameras for image-based ML might allow very large numbers of training datasets constructed in different laboratories to be merged toward the same targets for constructing more precise models.

To confirm the reproducibility and further expand the ease of applicability, we also developed a stand-alone prediction system that can be used by any external non-experts for third-party evaluation. Once a suitable ML model for a given task is obtained by training, it can be shared with others and separately used for making predictions of any unseen images of the task. Demonstration of the sugar/salt example with instructions for sample preparation and photographing is provided as a freely accessible Google Colaboratory notebook (see the Supporting Information). We have already tested the usability of this prediction system, and we confirmed that it worked well for a third-party (Figure S19). Although it should be carefully operated for input test images to match the training conditions (size or morphology), this suggests that image-based ML models can be used as a storage of visual information from experiments to be shared among researchers.

CONCLUSION

We have developed mixing-ratio prediction for various solid mixtures using image-based ML. Although the visual observations frequently used as preliminary evaluation of chemical contents are non-quantitative and largely depends on the experiences and intuition of the observers, ML enables accurate prediction of the weight ratio from photographs. The data-augmentation process allows construction of a reasonable model for mixture analysis using 300 training images. The trained models allowed prediction of the ratios of sugar and dietary salt, polymorphic crystals, and enantiomers with better accuracy than prediction by researchers. Image-based ML is also applicable to more complex mixtures containing four components, thus enabled non-destructive reaction analysis of solid-state decarboxylation of **PAS**. The limitation of the camera dependency was resolved by fine-tuning the models with a few calibration data. Although training image datasets should be collected when different size, color, or morphologies of solids are used, our ML system would be convenient when thousands of output batches produced through a common method are to be analyzed. Such situations are conceivable in combinatorial chemistry, or in laboratorial screening processes using robotic technologies[22] or automation synthesis.[23] In addition, ML-based prediction could act as a checker for quality assessment of combined medicines during long-term storage, reaction management in plant synthesis, and environmental soil surveys. Although it is generally difficult to share the experience of researchers, trained ML models will assume the role of a data bank, and even visually impaired people can receive the benefits from visual observations.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

General information, image-based ML data, and characterization of the recrystallized glycine samples. (PDF)

Data S1 for instruction, answer sheet, test and training images for evaluation of image-based ML system vs. human eye prediction accuracy. (PDF)

Accession Codes

All codes and data to reproduce the results in the paper are available at https://github.com/itakigawa/mixing_ratio_prediction and https://doi.org/10.6084/m9.figshare.20521224.v2, respectively.

AUTHOR INFORMATION

Corresponding Author

* Yuki Ide – Email: ide.yuki@icredd.hokudai.ac.jp,
* Ichigaku Takigawa – Email: takigawa@icredd.hokudai.ac.jp
* Yasuhide Inokuma – Email: inokuma@eng.hokudai.ac.jp

Notes

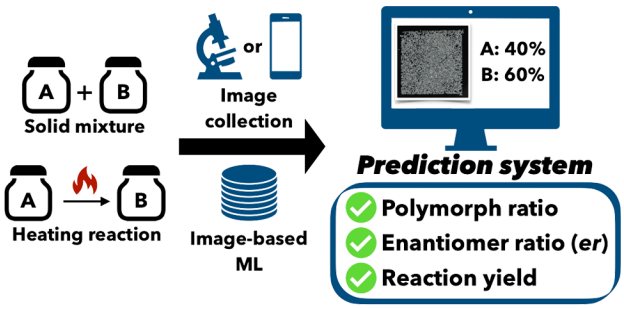
The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was partly supported by a JSPS Grant-in-Aid for Scientific Research(B) (grant no. 22H02058), a JSPS Grant-in-Aid for Challenging Research (Exploratory) (grant no. 20K21214), and the JST FOREST Program (grant no. JPMJFR211H) to Y. Inokuma. The Institute for Chemical Reaction Design and Discovery (ICReDD) was established by the World Premier International Research Initiative (WPI), MEXT, Japan. This work was supported by a JSPS Grant-in-Aid for Early-Career Scientists (grant no. 21K14597) to Y. Ide. This work was partly supported by a JSPS Grant-in-Aid for Transformative Research Areas (grant no. 20H05962), for Scientific Research (S) (grant no. 20H05657) to I. Takigawa. Y. Inaba appreciates a Grant-in-Aid for JSPS Research Fellows (grant no. 21J20973). This work was also supported by a JST SPRING Grant (grant no. JPMJSP2119) to H. S.

REFERENCES

(1) Pasteur, L. Recherches sur les relations qui peuvent exister entre la forme crystalline, la composition chimique et le sens de la polarisation rotatoire. *Ann. Chim. Phys.* **1848**, *24*, 442–459.
(2) Ahneman, D. T.; Estrada, J. G.; Lin, S.; Dreher, S. D.; Doyle, A. G. Predicting reaction performance in C–N cross-coupling using machine learning. *Science* **2018**, *360*, 186–190.
(3) Rinehart, N. I.; Zahrt, A. F.; Henle, J. J.; Denmark, S. E. Dreams false starts, dead ends, and redemption: a chronicle of the evolution of a chemoinformatic workflow for the optimization of enantioselective catalysts. *Acc. Chem. Res.* **2021**, *54*, 2041–2054.
(4) Meuwly, M. Machine learning for chemical reactions. *Chem. Rev.* **2021**, *121*, 10218–10239.
(5) Granda, J. M.; Donina, L.; Dragone, V.; Long, D.-L.; Cronin, L. Controlling an organic synthesis robot with machine learning to search for new reactivity. *Nature* **2018**, *559*, 377–381.
(6) Moosavi, S. M.; Chidambaram, A.; Talirz, L.; Haranczyk, M.; Stylianou, K. C.; Smit, B. Capturing chemical intuition in synthesis of metal-organic frameworks. *Nat. Commun.* **2019**, *10*, 539.
(7) Liu, Z.; Zhu, D.; Raju, L.; Cai, W. Tackling Photonic Inverse Design with machine learning. *Adv. Sci.* **2021**, *8*, 2002923.
(8) Ren, F.; Ward, L.; Williams, T., Laws, K. J.; Wolverton, C.; Hattrick-Simpers, J.; Mehta, A. Accelerated discovery of metallic glasses through iteration of machine learning and high-throughput experiments. *Sci. Adv.* **2018**, *4*, eaaq1566.
(9) Keith, J. A.; Vassilev-Galindo, V.; Cheng, B.; Chmiela, S.; Gastegger, M.; Müller, K.-R.; Tkatchenko, A. Combining machine learning and computational chemistry for predictive insights into chemical systems. *Chem. Rev.* **2021**, *121*, 9816–9872.
(10) Wang, X.; Li, J.; Ha, H. D.; Dahl, J. C.; Ondry, J. C.; Moreno-Hernandez, I.; Head-Gordon, T.; Alivisatos, A. P. AutoDetect-mNP: an unsupervised machine learning algorithm for automated analysis of transmission electron microscope images of metal nanoparticles. *JACS Au* **2021**, *1*, 316–327.
(11) Katsuno, H.; Kimura, Y.; Yamazaki, T.; Takigawa, I. Fast improvement of TEM images with low-dose electrons by deep learning. *Microsc. Microanal.* **2022**, *28*, 138–144.
(12) Katsuno, H.; Kimura, Y.; Yamazaki, T.; Takigawa, I. Early detection of nucleation events from solution in LC-TEM by machine learning. *Front. Chem.* **2022**, *10*, 818230.
(13) Jerez, D.; Stuart, E.; Schmitt, K.; Guerrerro-Given, D.; Christie, J. M.; Hahn, W. E.; Kamasawa, N.; Smirnov, M. S. A deep learning approach to identifying immunogold particles in electron microscopy images. *Sci. Rep.* **2021**, *11*, 7771.
(14) Zhou, Z.; Firestone, C. Humans can decipher adversarial images. *Nat. Commun.* **2019**, *10*, 1334.
(15) Hu, J.; Shen, L.; Albanie, S.; Sun, G.; Wu, E. Squeeze-and-excitation networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *42*, 2011–2023.
(16) Censi, R.; Martino, P. D. Polymorph impact on the bioavailability and stability of poorly soluble drugs. *Molecules* **2015**, *20*, 18759–18776.
(17) Han, G.; Thirunahari, S.; Chow, P. S.; Tan, R. B. H. Resolving the longstanding riddle of pH-dependent outcome of glycine polymorphic nucleation. *CrystEngComm,* **2013**, *15*, 1218–1224.
(18) Koleva, B. B.; Kolev, T. M.; Tsalev, D. L.; Spiteller, M. Determination of phenacetin and salophen analgetics in solid binary mixtures with caffeine by infrared linear dichroic and Raman spectroscopy. *J. Pharm. Biomed. Anal.* **2008**, *46*, 267–273.
(19) Moss, G. P. Basic terminology of stereochemistry (IUPAC Recommendations 1996). *Pure Appl. Chem.* **1996**, *68*, 2193–2222.
(20) Gawley, R. E. Do the terms "% ee" and "% de" make sense as expressions of stereoisomer composition or stereoselectivity? *J. Org. Chem.* **2006***, 71*, 2411–2416.
(21) Wesolowski, M. The decarboxylation and thermal stability of *p*-amino-salicylic acid and its salts. *Thermochim. Acta* **1997***, 21*, 243–253.
(22) Steiner, S.; Wolf, J.; Glatzel, S.; Andreou, A.; Grada, J. M.; Keenan, G.; Hinkley, T.; Aragon-Camarasa, G.; Kitson, P. J.; Angelone, D.; Cronin, L. Organic synthesis in a modular robotic system driven by a chemical programming language. *Science* **2019**, *363*, eaav2211.
(23) Burger, B.; Maffettone, P. M.; Gusev, V. V.; Aitchison, C. M.; Bai, Y.; Wang, X.; Li, X.; Alston, B. M.; Li, B.; Clowes, R.; Rankin, N.; Harris, B.; Sebastian, R.; Cooper, A. I. A mobile robotic chemist. *Nature* **2020**, *583*, 237–241.


A
+
B
Solid mixture
A
B
Heating reaction
or
Image collection
Image-based ML
A: 40%
B: 60%
Prediction system
Polymorph ratio
Enantiomer ratio (er)
Reaction yield