Predicting Solid Mixing Ratios Using Image-based Machine Learning and Application to Chemical Reaction Analysis

([1]*Grad. Sch. Eng., Hokkaido Univ.*, [2]*WPI-ICReDD, Hokkaido Univ.,* [3]*CIREDS, Kyoto Univ.*)
○Hayato Shirakura,[1] Taichi Sano,[1] Yuki Ide,[2] Sheng Hu,[2] Ichigaku Takigawa,[2,3] Yasuhide Inokuma[1,2]

Keywords: Machine Learning; Enantiomer Ratio; Mixing Ratio Prediction; Reaction Yield Prediction; Organic Crystal

Machine learning is recently used as a powerful tool for discovering relationships between complex chemical parameters, and enabling predictions in chemical reactions and molecular design. Our laboratory recently reported machine learning systems that instantly predicts mixing ratios of sugar and dietary salt from their images[1]. Image-based prediction has potentials to analyze objects that are difficult to identify with human eyes. In this research, we explored the practical applicability of this machine learning system for chemical research.

Crystal polymorphs are defined as different crystal structures of the same compound, and instrumental analysis methods are limited to the solid state. It was assumed that glycine with different crystalline polymorphs of α- and γ-from can be visually distinguished. Machine learning model built from 300 images of 200 mg α- and γ-glycine mixtures and their mixing ratios (training data) can predict mixing ratios from images with an error of 4.0% (**Figure 1a**). Prediction of the enantiomeric ratio of D-/L-tartaric acid was also possible, although the difference between the crystals was more difficult to distinguish with naked eyes. Furthermore, yield prediction was conducted as an application to chemical reaction. Reaction yield prediction of *m*-aminophenol obtained by solid-state decarboxylation for *p*-aminosalicylic acid was achieved with an error of 5.7% by using sample images and ^{1}H NMR yields (**Figure 1b**). As an example of more practical systems in chemical research, we attempted to predict the mixing ratios using sample solutions. The enantiomeric ratio was predicted from evaporated solid-state images of solutions containing (*R*)- or (*S*)-1,1'-binaphthyl with an error of 6.8% (**Figure 1c**).

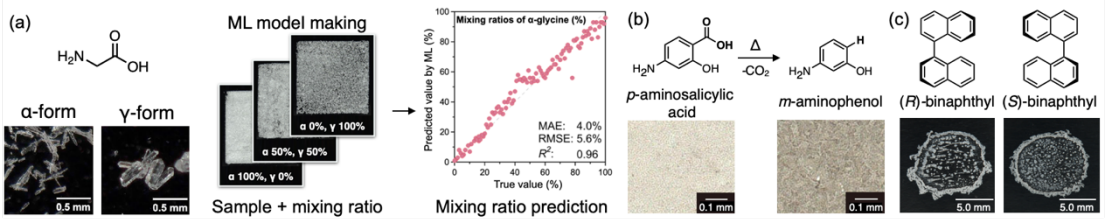

Figure 1. (a) Crystal polymorph mixing ratio prediction results from α- and γ-glycine crystal images. (b) Yield prediction using solid-state reaction images. (c) Enantiomeric ratio prediction using droplet drying images.

1) Y. Ide, H. Shirakura, T. Sano, M. Murugavel, Y. Inaba, S. Hu, I. Takigawa, Y. Inokuma, *Ind. Eng. Chem. Res.* **2023**, *62*, 13790.