2522
doi:10.1017/S1431927621008990

Microsc. Microanal. 27 (Suppl 1), 2021

Learning to Estimate the Composition of a Mixture with Synthetic Data

Cuong Ly[1], Cody Nizinski[2], Clement Vachet[3], Luther McDonald[2] and Tolga Tasdizen[4]

[1]University of Utah - Scientific Computing and Imaging Institute, Salt Lake City, Utah, United States, [2]University of Utah, United States, [3]Scientific Computing and Imaging Institute, United States, [4]University of Utah - Scientific Computing and Imaging Institute, United States

Identifying the precise composition of a mixed material is important in various applications. For instance, in nuclear forensics analysis, knowing the process history of unknown or illicitly trafficked nuclear materials when they are discovered is desirable to prevent future losses or theft of material from the processing facilities. Motivated by this open problem, we describe a novel machine learning approach to determine the composition of a mixture from SEM images. In machine learning, the training data distribution should reflect the distribution of the data the model is expected to make predictions for, which can pose a hurdle. However, a key advantage of our proposed framework is that it requires reference images of pure material samples only. Removing the need for reference samples of various mixed material compositions reduces the time and monetary cost associated with reference sample preparation and imaging. Moreover, our proposed framework can determine the composition of a mixture composed of chemically similar materials, whereas other elemental analysis tools such as powder X-ray diffraction (p-XRD) have trouble doing so. For example, p-XRD is unable to discern mixtures composed of triuranium octoxide (U3O8) synthesized from different synthetic routes such as uranyl peroxide (UO4) and ammonium diuranate (ADU) [1]. In contrast, our proposed framework can easily determine the composition of uranium oxides mixture synthesized from different synthetic routes, as we illustrate in the experiments.

Fig. 1 shows an overview of our proposed approach, which is divided into three steps as highlighted with the dashed rectangles. For the first step (Fig. 1a), we develop a model to generate a synthetic dataset of mixed materials using real images of pure materials only. Our image synthesis model is based on the texture synthesis work proposed by Gatesy et al. [2], which generates a new image with the same texture as a given reference image by minimizing the differences between second-order statistics of features (Gram matrix) of the generated and reference images [2]. We define each pure material as a texture. A new synthetic mixture image is initialized as an image of white noise. We then minimize the differences between the Gram matrix of this image and the weighted sum of Gram matrices of the pure reference images. In the second step (Fig. 1b), we train a convolutional neural network (CNN) to estimate mixture compositions from SEM images using synthetic images generated from the first phase. After the training process, the mixture composition estimation model is deployed to estimate the composition of a mixture. This step is shown in Fig. 1c.

We utilized mixtures of U3O8 made from ADU, uranyl hydroxide (UH), and sodium diuranate (SDU) to validate the proposed method. We refer to these U3O8 materials simply by their precipitation route, e.g., a U3O8 sample made from ADU is simply referred to as ADU for the rest of this paper. First, we used the image synthesis model to obtain the necessary dataset for training the mixture composition estimation model. In this experiment, the image synthesis model takes SEM images of ADU, UH, and SDU as input and produces images of the mixtures of prescribed percentages. Next, we used the synthetic images to train the mixture composition estimation model, which is tasked with estimating the percentage of ADU, UH, and SDU presence in an SEM image. Then, the mixture composition estimation model was tested on a set of real mixture images to validate the performance of the proposed framework. The set of real mixture images used for testing consists of six compositions: 100%ADU,100%UH, 100%SDU, 50%ADU-

50%UH, 50%UH-50%SDU, 33%ADU-33%UH-33%SDU. We also compared against a best-case CNN mixture composition estimation model that uses real samples of mixed materials for training. In other words, instead of synthesizing images, another set of real mixture images was used to train the mixture composition estimation model. The number of compositions of training samples is the same for the two experiments. Fig. 2 shows the mixture composition estimation model results. As seen in that figure, the estimated composition of the mixtures produced by the model is in good agreement with the ground truth compositions. More importantly, the performance of the mixture composition estimation model when trained with real images and when trained with synthetic images is comparable. This result demonstrates that our image synthesis model generates realistic images that can be used to replace real images for purposes of training a mixture composition estimation model.

In this work, we demonstrated that a novel framework can be used to determine the composition of a mixture. The main advantage of our proposed framework is that it does not require reference mixed samples. This advantage eliminates the costly reference mixture sample acquisition process, which can be prohibitive when the mixture is composed of large number of materials. Furthermore, our proposed framework can be applied to other types of mixtures beyond nuclear materials [3].

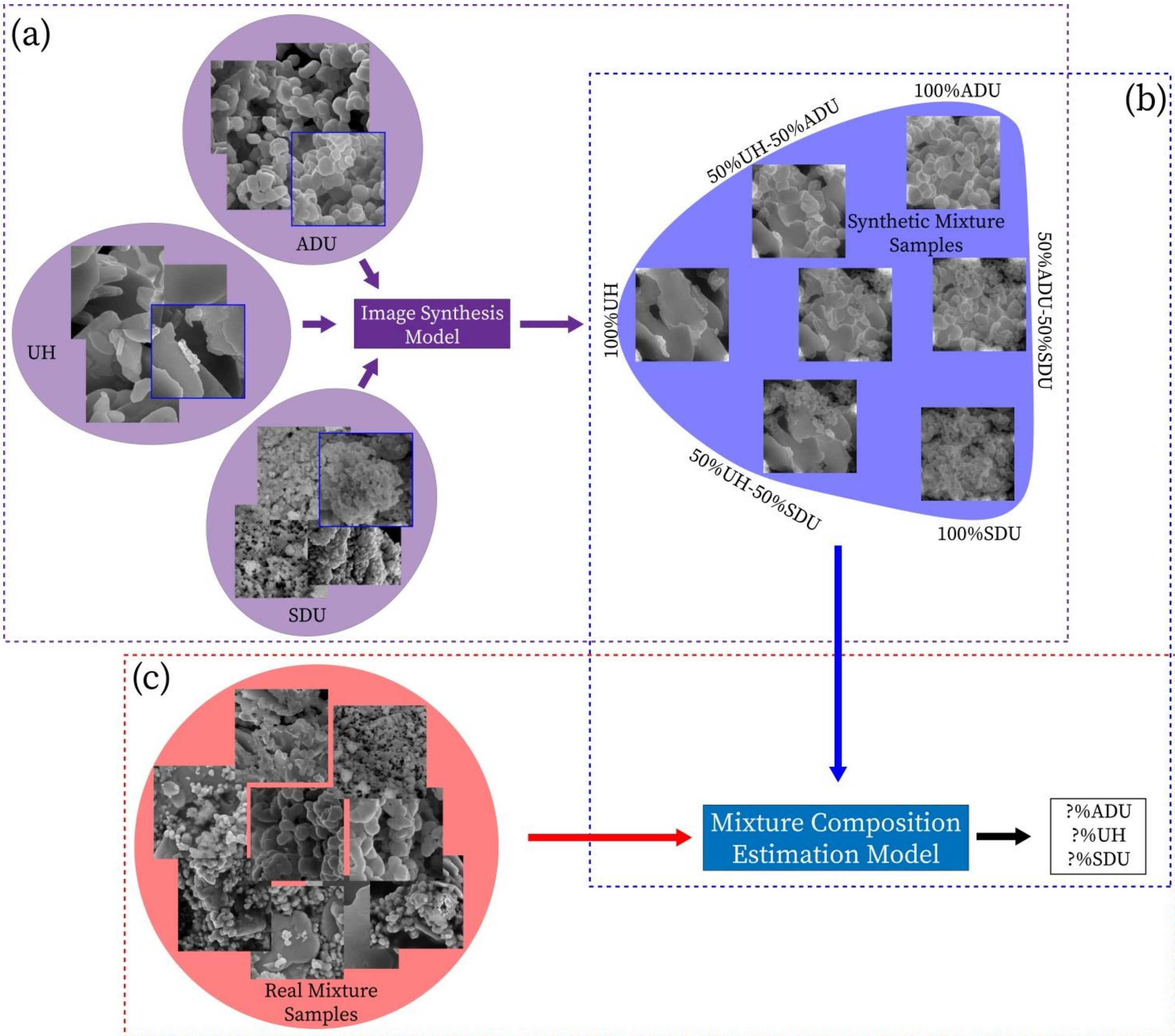

Figure 1. Fig. 1. An overview of our proposed composition of mixture estimation framework. (a) In the first phase of this framework, an image synthesis model takes SEM images of pure materials (e.g., ADU, UH, and SDU) composed in the desired mixtures as input and generates mixture images. (b) The synthetic images, obtained from (a), are then utilized to train the mixture composition estimation model, which provides an estimated percentage of each pure material presence in a mixture. (c) After the training process, the mixture composition estimation model can be deployed to estimate the composition of a mixture.

Dataset	100% ADU 0% UH 0% SDU	0% ADU 100% UH 0% SDU	0% ADU 0% UH 100% SDU	50% ADU 50% UH 0% SDU	0% ADU 50% UH 50% SDU	33% ADU 33% UH 33% SDU	RMSE	R^2
Real Images	0.991 0.00500 0.00400	0.00300 0.992 0.00500	0.0210 0.104 0.875	0.330 0.545 0.125	0.0720 0.582 0.346	0.260 0.500 0.240	0.0920	0.931
Synthetic Images	0.963 0.0260 0.0110	0.0280 0.943 0.0290	0.0600 0.0960 0.844	0.120 0.488 0.392	0.110 0.500 0.390	0.092 0.496 0.412	0.153	0.803

Figure 2. Fig. 2. Results of the mixture composition estimation model when trained with real images and with synthetic images. The table shows the mean of five repetitions of model training. The mean estimated compositions of the mixtures in the test set are shown in columns 2-7. The coefficient of determination (R^2) and root-mean-square error (RMSE) between the estimated and ground truth compositions are also shown.

References

[1] S. Heffernan, N. Ly, B.J. Mower, C. Vachet, I.J. Schwerdt, T. Tasdizen, L.W. McDonald IV, Identifying surface morphological characteristics to differentiate between mixtures of U3O8synthesized from ammonium diuranate and uranyl peroxide, Radiochimica Acta. 0(0) (2019). doi:10.1515/ract-2019-3140.

[2] L. A. Gatys, A. S. Ecker, and M. Bethge, Texture synthesis using convolutional neural networks, Neural Information Processing Systems 28: Annual Conference on Neural Information Processing Systems. (2015)pages 262–270.

[3] This work is supported by the Department of Homeland Security, Domestic Nuclear Detection Office, under grant Number 2015-DN-077-ARI092.