Machine Learning on STEM-EDS Data for Quantifying Overlapping Deep-Mantle Rock Assemblages

Hui Chen[1], Farhang Nabiei[2], James Badro[3], Duncan TL Alexander[1] and Cécile Hébert[1]

[1]École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Vaud, Switzerland, [2]University of Cambridge, Cambridge, England, United Kingdom, [3]Institute de Physique du Globe de Paris, Paris, Ile-de-France, France

Geophysical interpretations of the Earth's interior and its dynamics are significantly influenced by chemical compositions of the constituting minerals [1]. Scanning transmission electron microscopy (STEM) is a versatile technique for studying minerals and rocks. It allows several chemical analyses, such as energy-dispersive X-ray spectroscopy (EDS) and electron energy-loss spectrometry (EELS) to be done (near-)simultaneously with high spatial resolution. Combining it with the electron tomography (ET) technique, the 3D phase relationship and chemical composition of each overlapping phase in a mineral assemblage can be retrieved. However, a 2D EDS/EELS scan is more commonly used, because of the instability of certain minerals under the extensive beam exposure that is necessary for recording the tilt series data set for ET. At the same time, machine learning has begun to make an impressively positive impact in various fields [2]. Here we look to address the challenge of quantifying overlapped phases in 2D STEM-EDS spectrum imaging (SI) data by exploiting several machine learning algorithms.

In this paper, the starting material is a synthetic pyrolite glass doped with Nd, Sm, Hf, Lu, and U (0.3 wt.% for each). Four samples were made by compressing the pyrolite across a range of pressures from 46 GPa to 88 GPa, using a diamond anvil cell (DAC). The samples were molten by double-sided laser heating and then slowly cooled down below the solidus temperature before quenching. Thin sections for STEM analysis were made by the focused ion beam (FIB) lift-out technique from the samples recovered after decompression.

Figure 1(a) is a HAADF image of the 71 GPa sample. Three phases can be identified in the STEM-EDS SI data set: ferropericlase (Fp), bridgmanite (Brg), and Ca-rich perovskite (Ca-Pv). All three phases are partly overlapping, with Brg being the matrix phase. An algorithm called non-negative matrix factorization (NMF) [3] with non-negativity constraint decomposed the acquired STEM-EDS data into 3 components; the resultant loading spectra and score images are shown in Figure 1(b)-(e). NMF#0, NMF#1, NMF#2 feature Brg, Fp and Ca-Pv, respectively, as indicated by their loading spectra. Though the loading spectra of these components do not completely coincide with the true spectra of the physical phases, owing to a lack of necessary physical constraints in the data decomposition, they reveal a clearer spatial phase distribution between Fp, Brg and Ca-Pv than traditional elemental maps.

By adding the score images of NMF#1 and NMF#2, it is possible to identify and segment the SI spatial region that corresponds to pure Brg. A trainable Weka using fastRandomForest classifier [4] combined with a global smoothness tuning [5] aid in this process, thereby generating the pure Brg map shown in Figure 1(f). Figure 1(g) compares the Nd La and Sm La EDS signals integrated over this segmented pure Brg area compared to that from a confined region of Brg, as indicated by the green rectangle in Figure 1(a). The full Brg area signal integration clearly gives a much-improved signal to noise for these trace elements, which can further enable a more confident investigation of rare earth elements partitioning in minerals and melt. This protocol also applies to the quantification of Fp and Ca-Pv grains, but with some modification because of the overlapping between these two kinds of grains. Figure 2(a) and 2(b) are binary

images for labeling Fp and Ca-Pv grains, made from NMF#1 and NMF#2. The intersection of Fp, Ca-Pv, and Brg will be the same as the intersection of Figure 2(a) and 2(b), which should be avoided during the quantification of Fp or Ca-Pv. By eliminating the phase intersections, segmented Fp-Brg mixtures and Ca-Pv-Brg mixtures are demonstrated in Figure 2(c) and 2(d). Finally, by subtracting the Brg contribution from these mixtures, the pure spectra of Fp and Ca-Pv grains are shown in Figure 2(e). Therefore, we have demonstrated that the NMF-aided phase segmentation is an effective method for segmenting spectra from overlapping mineral assemblages for improved quantification, when applied appropriately. It can further be extrapolated to the data processing of other spectroscopic techniques, such as EELS.
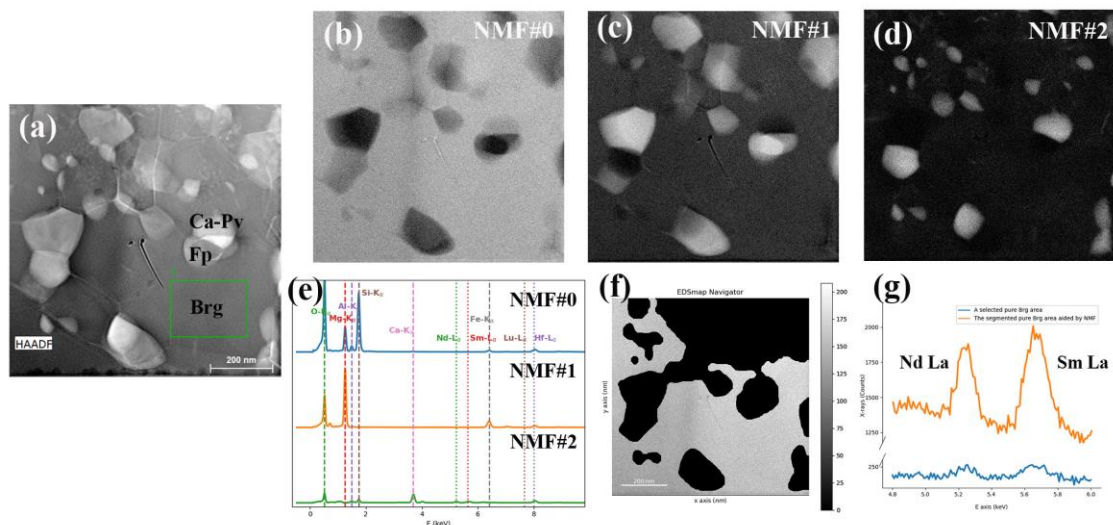

Figure 1. (a) HAADF image of the 71 GPa sample; (b)-(d) representative score images and (e) loading spectra from its NMF decomposition; (f) the segmented pure Brg map; (g) spectral comparison between the segmented pure Brg area and a confined Brg area.
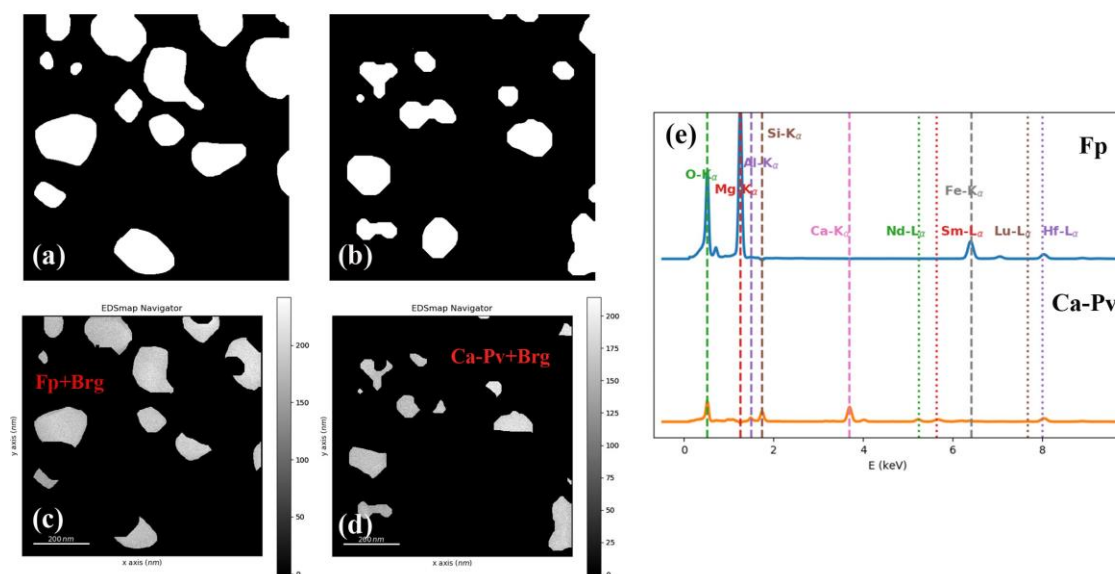

Figure 2. (a) and (b) binary masks for labeling Fp and Ca-Pv grains; (c) Fp-Brg mixtures; (d) Ca-Pv-Brg mixtures; (e) pure Fp spectrum and pure Ca-Pv spectrum.

References
[1] T Irifune et al. Science, **327** (2010), p. 193–195.
[2] J Schmidt et al., npj Computational Materials 5 (2019), no. 83.
[3] F de la Pena et al., Microscopy and Microanalysis, 23 (2017), p. 214–215.
[4] https://imagej.net/Trainable_Weka_Segmentation
[5] https://imagej.net/Graph_Cut