Mapping Mineralogical Distributions on Mars with Unsupervised Machine Learning

Max Hipperson¹, Ingo Waldmann¹, Peter Grindrod², and Nikolaos Nikolaou¹

¹Department of Physics and Astronomy, University College London, London, UK
²Department of Earth Sciences, Natural History Museum, London, UK

Abstract

Knowledge of the constituents of the Martian surface and their distributions over the planet informs us about Mars’ geomorphological formation and evolutionary history. In this research, an unsupervised end-to-end unmixing model using autoencoders is proposed, that can produce physically plausible abundance maps of the surface mineralogy from hyperspectral imaging data.

1 Introduction

Mars’ formationary and evolutionary history can be inferred by studying the mineralogical constituents of the Martian surface and their distributions. Knowledge of the specific minerals present and their abundances constrain the possible histories of the planet’s surface (Liu et al., 2016).

Hyperspectral imaging of Mars from the Compact Reconnaissance Imaging Spectrometer for Mars (CRISM; Murchie et al., 2007) provides unprecedented insight into the distribution of surface mineralogy. Usually, minerals are identified from spectral absorption features and comparison to spectral libraries (see Viviano-Beck et al., 2014). This can be incredibly time consuming, and relies upon detailed a priori knowledge of the surface chemistry and geology. In order to obtain constituent mineralogical spectra (endmembers) and their fractional abundances one must perform spectral unmixing.

Machine learning methods have been used extensively for Earth based hyperspectral imaging analysis, but only a few works to our knowledge have attempted to produce end-to-end unmixing models: Guo et al. (2015); Palsson et al. (2018, 2017); Zhang et al. (2018). Data obtained from hyperspectral imaging of Mars is essentially of the same format as that obtained from Earth, so similar kinds of analysis can be performed.
2 Method

Signal decomposition techniques, namely principal component analysis (PCA) and independent component analysis (ICA) are used in order to investigate the presence - and the approximate number - of distinct, clearly detectable signals (component spectra) within CRISM imaging data.

Clustering algorithms are then used to obtain an alternative estimate of the number of distinct identifiable mineralogical components present within the spectra. The intuition is that the optimal number of clusters identified should match the number of components determined through the use of signal decomposition techniques.

Autoencoders (a type of neural network consisting of an encoder and decoder function that are trained to reproduce their input to their output) are then trained on the spectra from the image. They learn an approximation of the LMM whereby the weights of the decoder correspond to the endmember spectra and the activations of the encoder are the fractional abundances of the endmembers for a given spectrum (Palsson et al., 2018, 2017). The number of endmembers in the learned approximation to the LMM are then set as the estimates of the number of component signals from PCA and clustering. Figure 1 shows a schematic of this.

Learned endmembers are extracted from the networks and abundance maps are produced by projecting the activation of each neuron in the decoder to the pixel location of the input spectrum for all of the spectra in the image.

These techniques all get around the requirement of detailed a priori knowledge by making no assumptions about the state of the surface. To our knowledge, this is the first work to address the use of unsupervised neural networks for an end-to-end spectral unmixing of Martian hyperspectral images.

![Figure 1: Schematic view of an autoencoder used for spectral unmixing.](image-url)
3 Results and Conclusion

The signal decomposition techniques provide strong evidence for the presence of detectable mineralogical signals within the data, and the number of signal components estimated from PCA are consistent with the estimates from clustering.

PCA, ICA, and the abundance maps produced by the trained autoencoders all appear to convincingly replicate features of the surface topography. Figure 2 shows an example of this, comparing the quicklook image for phyllosilicates (a false colour RGB image constructed from wavelengths using spectral parameters; Viviano-Beck et al., 2014) to one of the produced abundance maps. This provides strong evidence that these methods have learnt mineralogically significant spectra, and that the autoencoders have learnt a physically plausible approximation to the LMM. This is despite the fact that at no spatial information is provided at any point to any of these methods – their results are purely based on spectral information.

Hence, when combined with methods to estimate the number of endmembers present, one can construct an unsupervised end-to-end linear unmixing model using autoencoders, that learns to identify physically plausible mineralogical endmembers. Once trained, this pipeline will allow the large number of CRISM data sets to be characterised and abundances mapped over the surface of the planet. This method is also extendable to any kind of planetary hyperspectral imaging, allowing simple mineralogical mapping on an unprecedented scale.

Figure 2: Quicklook image for phyllosilicates (a false colour RGB image made using spectral parameters; Viviano-Beck et al., 2014) alongside one of the produced abundance maps.
References


