Applying Machine Learning to Analysis of Hyperspectral Images of Mars

Abraar Chaudhry,
As part of a research group including:
Michael Littman, Jack Mustard, Jesse Tarnas, Michael Mao

Fall 2018

Abstract

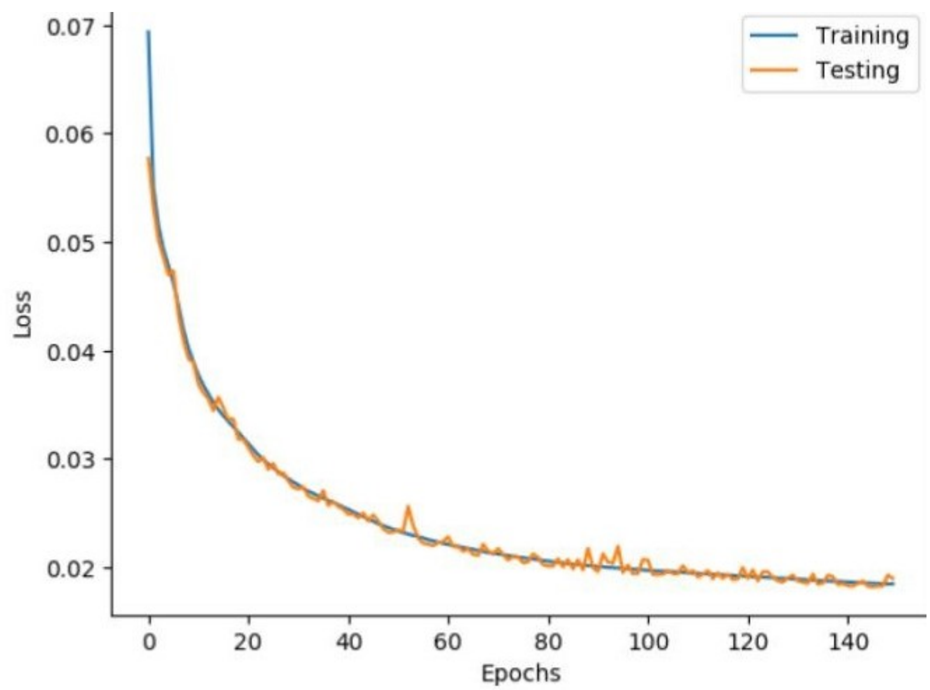
There exists a large amount of hyperspectral images of the planet Mars. Scientists have analyzed parts of these images to determine local geology. Machine learning may be used to improve and automate this analysis. We explore autoencoders in particular to denoise images and to replicate more conventional methods. Deep learning and convolutional networks are often powerful tools to analyze images; these approaches are used as well. Various normalization and preprocessing steps are also employed.

1 Details

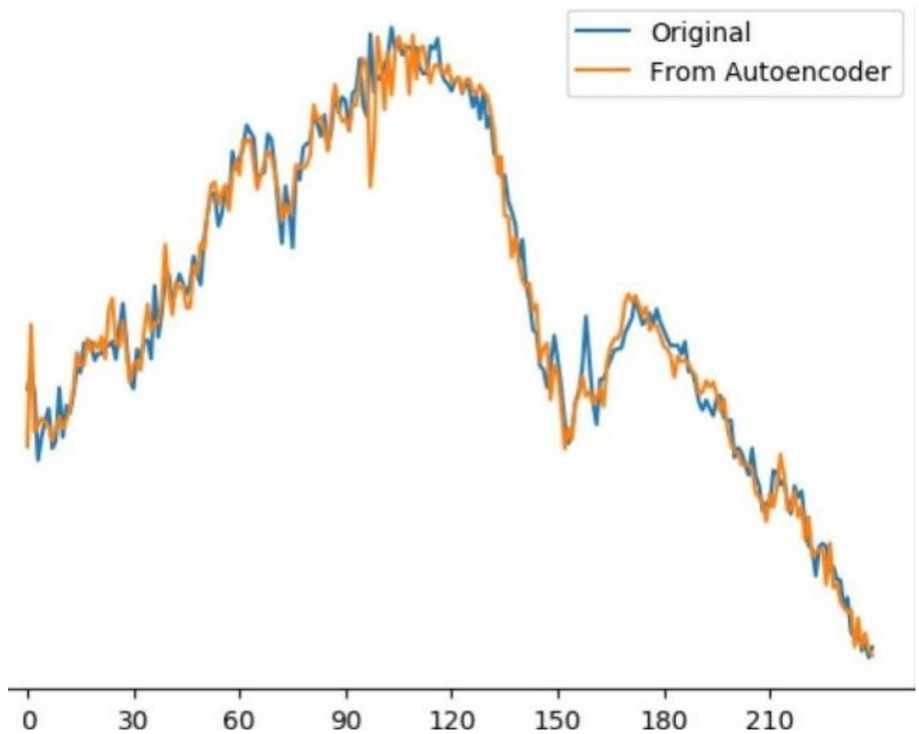
The first issue was missing data as parts of the images examined have a great dearth of usable data. This issue is especially prevalent at the edges of images and this is explained as an artifact of the imaging process. This issue is partially resolved by using images with atmospheric correction and by truncating the edges of images. Furthermore, we also truncate data to a range of spectral bands that is well studied. The first method attempted was a simple PCA. This method did not seem to work well and only the first one or two components seemed meaningful. The first component seemed to represent the albedo of the surface at a particular point and this was normalized out for subsequent models. The next model attempted was a pixel-wise deep autoencoder. Min-max scaling was used so that targets would be in a suitable range for a final sigmoid activation in the network. Mean absolute error was chosen as the loss for the network in order to not be oversensitive to noise in the data. This method had some promise however due to the relation of the spectral bands to each other, it seemed like a convolution network would yield better results. A pixel-wise convolutional network did, in fact, perform better. To clarify, the convolutions were along the spectral bands, not in a spatial dimension. This network achieved the lowest loss and results are displayed as an example below. Another model used was based on encoding a 3x3 grid of pixels and from that reconstructing the spectra of the central pixel. This method did not have quite a low loss as the pixel-wise convolutional autoencoder but it shows promise.

2 Results

The following figure shows the losses of the pixel-wise convolutional autoencoder over the course of its training.



The following figure shows a comparison of the spectrum of a pixel compared to the output of the autoencoder given the same spectrum as input.



3 Next Steps

Further steps include further tuning of the grid based model, including possibly expanding its input from 3x3 to perhaps 5x5 or 7x7. The other important next step is to see if the information encoded by the autoencoder is useful in identifying features of geological interest. If the network encoded useful information, then it should be possible to augment the network to identify important geological features automatically.