

# RAMAN SPECTROSCOPY DECONVOLUTION USING STACKED AUTO-ENCODERS WITH NON-NEGATIVITY CONSTRAINTS

*Jacob S. Larsen, Flavia D. Frumosu, Jakob Thrane, Maximillian F. Vording, Tommy S. Alstrøm*

Department of Applied Mathematics and Computer Science, Technical University of Denmark

## ABSTRACT

**Index Terms**— One, two, three, four, five

## 1. INTRODUCTION

Raman scattering is a technique used to detect and identify molecules using the interaction of photons. It is a technique that is commonly used in chemistry and physics for detecting molecules by observing their vibrational and rotational modes. This is completed with the use of laser light. The photon excites the sample which causes scattering, meaning the photon has a change in energy over a short time period. This energy-change is reflected as a shift in frequency, also called a stokes shift, and by analysing the spectrum the molecules and combination of molecules can be identified.

Surface-enhanced Raman Scattering (SERS) is an enhancement of Raman Scattering, that uses surfaces such as metal or nanostructures to absorb molecules. This enables the identification of single molecules. For instance, noble metal nanostructures can concentrate light which greatly enhances the electromagnetic field near the nanostructure. These areas become so called "hot spots" that amplify weak Raman scattering signals which may not have been possible to detect with regular Raman scattering. The placement and design of these nanostructures with high SERS performance is beyond the scope of this work, however additional information can be found in [1] and references herein.

Complex mixtures in spectra.

NMF/MCR for source separation and mixture classification.

Computational heavy for increased resolution of Raman spectra.

Sparse autoencoder

## 2. PRIOR WORK

Primary reference [2]

---

Thanks to XYZ agency for funding.

## 3. METHODS

This work seeks to compare traditional and proven methods for SERS analysis, i.e. curve resolution, such as NMF to denoising autoencoders with non-negativity constraints. **Add how NMF/MCR is used in raman spectroscopy.**

### 3.1. Non-negative matrix factorization

Non-negative matrix factorization (NMF) consists of factorizing a original matrix  $V$ , with only positive elements, into two positive matrices  $W$  and  $H$ . [3]

$$V \approx W \times H \quad (1)$$

Where columns of  $W$  are considered basis vectors, and each column in  $H$  is considered an encoding with a one-to-one relationship with the columns of  $V$ . So for instance, a matrix  $V$  of size  $n \times m$  can be factorized into a matrix  $W$  of size  $n \times k$  and a matrix  $H$  of size  $k \times m$ , where  $k$  is the number of components. Approaching the intuitive point of view,  $W$  and  $H$  can be seen as components that combined approximate the original signal  $V$ .

An iterative algorithm is considered for NMF, which shares similar monotonic convergence as the EM algorithm. [4]. Moreover, the rules of update preserve non-negativity of  $W$  and  $H$ . The algorithm approaches the problem by initialization of  $W$  and  $H$  as non-negative, and then update the values in  $W$  and  $H$  until local maxima is obtained and both matrixes are considered stable. The rules of multiplicative update can be defined as:

$$H^{n+1} \leftarrow H^n \frac{(W^n)^T V}{(W^n)^T W^n H^n} \quad (2)$$

and

$$W^{n+1} \leftarrow W^n \frac{V (H^{n+1})^T}{W H^{n+1} (H^{n+1})^T} \quad (3)$$

This however leads to an optimization that is convex for  $W$  and  $H$  which can require many iterations and poor local minima. For this problem, alternating least squares (ALS) is commonly used. This consists by alternating the optimization between  $W$  and  $H$ , e.g. one iteration consists of 1) keeping

$H$  fixed while solving for  $W$ , and 2) keeping  $W$  fixed while solving for  $H$ . Further details can be found in [5] and references herein. This is, for the remainder for this work, defined as NMF-ALS.

3.2. Sparse auto-encoder with non-negativity constraint

Reconstruction of input vector through unsupervised learning. Parametric deterministic mapping [6]

$$y = f(x) = \mathbf{W}x + \mathbf{b}$$
 (4)

Where  $x$  is the input vector and  $\mathbf{W}$  is a mtrix of weights and  $\mathbf{b}$  is a bias vector.  $y$  is the latent representation that is reconstructed in input space by

$$z = g(y) = \mathbf{W}'x + \mathbf{b}'$$
 (5)

Primary reference implementation. Describe the constraint and the training architecture

4. SETUP

Dataset origin, description of wavenumbers, raman map size. Definitions for the rest of the paper.

architecture drawing with stacked training and softmax/sigmoid layer.

MNIST verification

5. RESULTS

Training and test data.

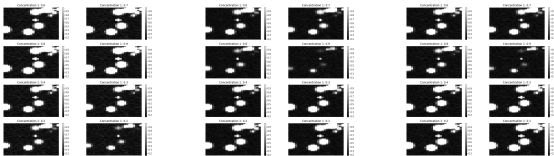
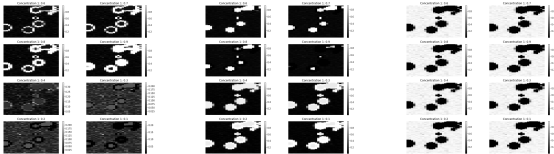
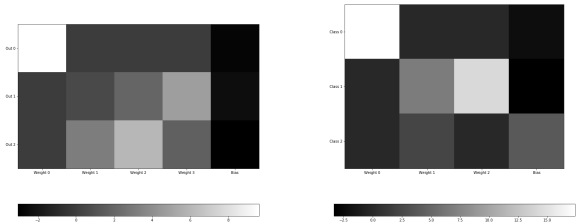
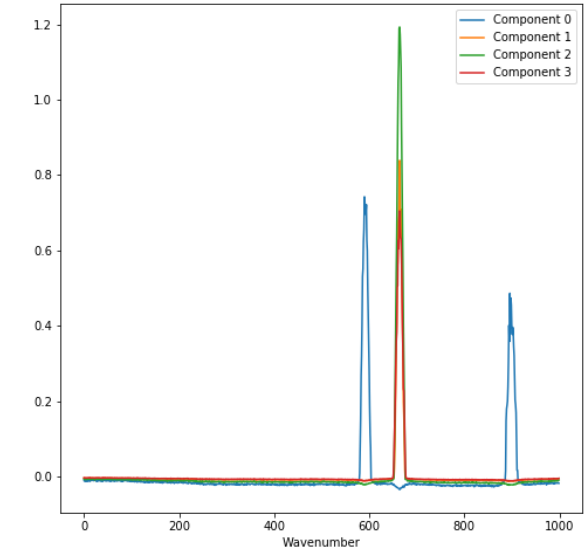
Encodings and basis vectors

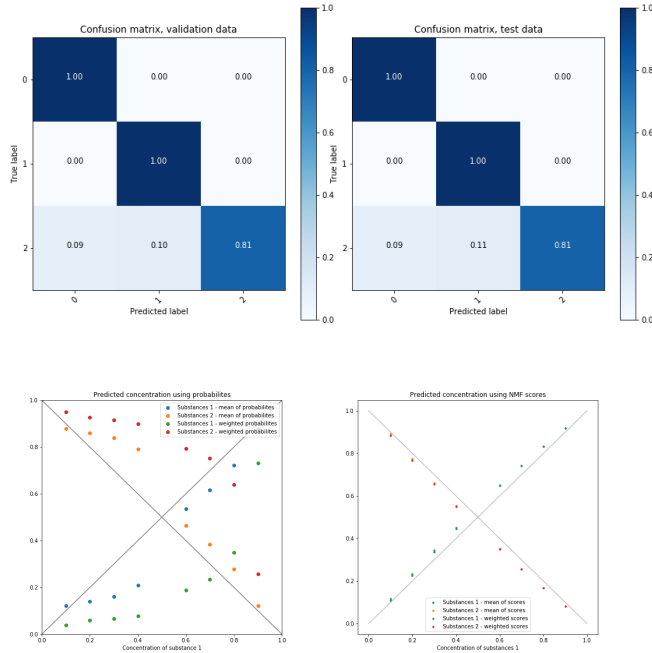
Classifier results

6. DISCUSSION

Comparison with NMF

7. CONCLUSION





[6] Pascal Vincent, Hugo Larochelle, Yoshua Bengio, and Pierre-Antoine Manzagol, “Extracting and Composing Robust Features with Denoising Autoencoders,” .

## 8. REFERENCES

List and number all bibliographical references at the end of the paper. The references can be numbered in alphabetic order or in order of appearance in the document. When referring to them in the text, type the corresponding reference number in square brackets as shown at the end of this sentence .

## 9. REFERENCES

- [1] Hong Wei and Hongxing Xu, “Hot spots in different metal nanostructures for plasmon-enhanced Raman spectroscopy,” *Nanoscale*, vol. 5, no. 22, pp. 10794, oct 2013.
- [2] Ehsan Hosseini-Asl, Jacek M. Zurada, and Olfa Nasraoui, “Deep Learning of Part-Based Representation of Data Using Sparse Autoencoders With Nonnegativity Constraints,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 12, pp. 2486–2498, dec 2016.
- [3] H. Sebastian Seung and Daniel D. Lee, “Learning the parts of objects by non-negative matrix factorization,” *Nature*, vol. 401, no. 6755, pp. 788–791, oct 1999.
- [4] A. P. Dempster, N. M. Laird, and D. B. Rubin, “Maximum Likelihood from Incomplete Data via the EM Algorithm,” 1977.
- [5] Amy N. Langville, Carl D. Meyer, Russell Albright, James Cox, and David Duling, “Algorithms, Initializations, and Convergence for the Nonnegative Matrix Factorization,” jul 2014.