Contrib. Astron. Obs. Skalnaté Pleso **53/3**, 51 – 57, (2023)
https://doi.org/10.31577/caosp.2023.53.3.51

Analysis of laser ablation spectral data using dimensionality reduction techniques: PCA, t-SNE and UMAP

M.S. Rabasovic, D.M. Pavlovic and D. Sevic

Institute of Physics, University of Belgrade, Pregrevica 118, 11080 Zemun, Serbia, (E-mail: sevic@ipb.ac.rs)

Received: September 15, 2023; Accepted: October 9, 2023

Abstract. Laser ablation is among the various methods employed for expediting the prototyping of printed circuit boards. Laser-induced breakdown spectroscopy (LIBS) serves as a convenient technique for overseeing the targeted elimination of thin layers with lasers. Consequently, this approach facilitates the rapid prototyping of printed circuit boards. In this paper the obtained LIBS data are analyzed by using data dimension reduction techniques: principal component analysis (PCA), t-SNE and UMAP to obtain an indication that copper layer is fully removed. For machine learning approach to data analysis we use Solo+Mia software package (Version 9.1, Eigenvector Research Inc, USA).

Key words: Machine learning – Dimensionality reduction – Laser induced breakdown spectroscopy

1. Introduction

Machine learning methods, including both unsupervised techniques like PCA (Principal Component Analysis) and supervised techniques like LDA (Linear Discriminant Analysis), are increasingly being employed to analyze LIBS data. These combinations of well-known machine learning algorithms with LIBS enable the swift and accurate classification of diverse samples Bellou et al. (2020); Diaz et al. (2020); Pořízka et al. (2018); Yang et al. (2020); Zhang et al. (2022).

Laser ablation offers a swift method for the rapid prototyping of printed circuit boards. A viable approach to this technique was introduced in a prior work Rabasović et al. (2016). In our application, we have harnessed laser-induced breakdown spectroscopy (LIBS) as a convenient method to both perform ablation and monitor the precise removal of thin layers using lasers. A similar technique has been detailed in a recent publication, as evidenced by Shiby & Vasa (2022). In the study by Rabasović et al. (2016), LIBS data were scrutinized through the application of correlation coefficients. Presently, with the increasing accessibility of high-speed computers capable of machine learning, the trajectory of analysis algorithms has shifted from basic numerical calculations towards more advanced artificial intelligence methods. Our initial efforts for machine learning analysis of LIBS printed circuit board data, using principal component analysis are presented in Sevic et al. (2020). Interesting applications of machine learning algorithms for analysis of LIBS data are presented in Boucher et al. (2015); Moros et al. (2013); Serranoa et al. (2014); Rabasovic et al. (2022). State of the art approaches to the problem are reviewed in Pořízka et al. (2018); Vrábel et al. (2020); Zhang et al. (2022).

In this paper we study the spectral data by using data dimension reduction techniques based on machine learning (ML). For ML approach to data analysis we use Solo+Mia software package (Version 9.1, Eigenvector Research Inc, USA) Wise et al. (2006).

2. Methods

Our experimental arrangement for acquiring the training spectra required for data dimensionality reduction is comprehensively elucidated in Rabasović et al. (2016). In brief, spectral images capturing the optical emissions from laser-induced plasma on a printed circuit board are obtained using a streak camera. These images are temporally integrated to generate spectra at different time points. Importantly, due to the time-resolved nature of streak images, we had the ability to carefully choose time intervals for spectrum integration, ensuring that the highly intense optical emissions from the initial plasma phase were excluded.

Data dimension reduction constitutes a fundamental technique within the realm of machine learning. This method involves using a dataset with a particular structure to "train" a machine, enabling it to discern specific features within the input data. These discernible characteristics are accentuated through the process of data dimension reduction. Subsequently, the machine becomes proficient at recognizing and identifying these characteristics in newly presented data that shares a similar structure and nature. A prevalent approach for automating this analysis is by teaching the machine a low-dimensional representation of the data. Within this low-dimensional representation, each object from the original high-dimensional dataset is portrayed as a point in a reduced-dimensional space. This representation is designed in such a way that proximate points correspond to similar objects, while distant points correspond to dissimilar ones. The low-dimensional embedding lends itself to straightforward visualization.

In this study we compare principal component analysis (PCA), t-distributed stochastic neighbor embedding (t-SNE), and uniform manifold approximation and projection (UMAP) methods for analysis of LIBS spectral data to obtain indication that copper layer is fully removed on given small area of printed circuit board.

The PCA was proposed long ago, see Hotelling (1933); Karhunen (1947); Lévy (1948). Efficient implementation of t-SNE is proposed in Van der Maaten

& Hinton (2008). UMAP is proposed in McInnes et al. (2018, 2020); Sainburg et al. (2020).

Similarly to most non-linear dimension reduction techniques, t-SNE and UMAP lack the strong interpretability of PCA. In particular the dimensions of the t-SNE and UMAP embedding space have no specific meaning, unlike PCA where the dimensions are the directions of greatest variance in the source data.

For obtaining the ML models we have used the set of 40 printed circuit boards spectra. For testing purposes we used spectra from beggining and from the end of ablation process.

3. Results and discussion

Plasma breakdown optical spectra of printed circuit board at the start, when only copper is ablated; and when the substrate is fully exposed, are shown in Fig. 1. Their differences could be seen by a naked eye.

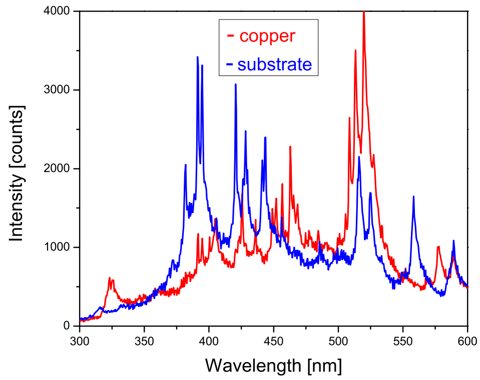

Figure 1. Plasma breakdown optical spectra of printed circuit board at the start, when only copper is ablated; and when the substrate is fully exposed.

The first two principal components of LIBS data are shown in Fig. 2. Embeddings plots for two t-SNE and UMAP components are shown in Fig. 3 and Fig. 4. Data dimension reducing techniques t-SNE and UMAP are designed for visualization of data. Using numerical results of embedding components any clustering technique will have no problem to put the test specrum into corresponding cluster.

Visual comparisons of PCA, t-SNE and UMAP plots shows the best performance of UMAP method. When reduced to two dimensions, dimension reduced spectral data are the most closely grouped by UMAP method. It should be mentioned that results presented in Fig. 4 were obtained when UMAP algorithm was initialized with parameters suggested by SOLO software. On the other hand, we had to try varying parameters for initialization of t-SNE algorithm to obtain acceptable results shown in Fig. 3.

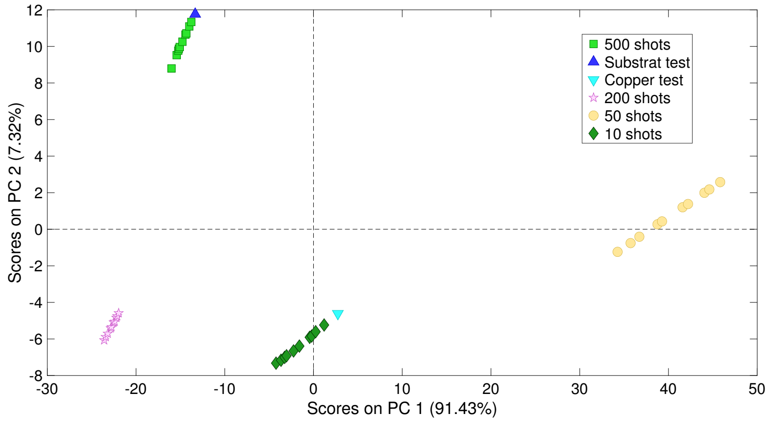

Figure 2. Scores plot of first two principal components.

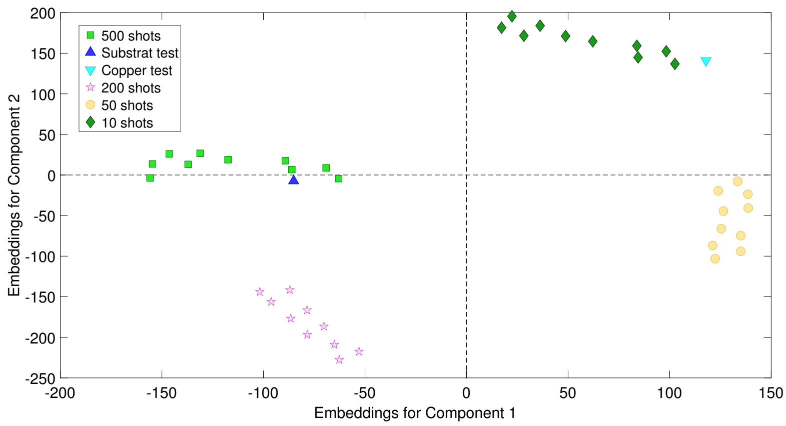

Figure 3. Embeddings plot for two t-SNE components.

4. Conclusions and Discussion

The rapid prototyping of printed circuit boards can be effectively accomplished through the application of laser ablation and LIBS. Our analysis of LIBS data related to printed circuit boards has involved the utilization of data dimension reduction techniques.

In our previous investigations, we employed correlation coefficients and PCA to determine the point at which laser ablation penetrates the composite substrate of the printed circuit board. In this particular study, we conducted a comparative analysis of PCA, t-SNE, and UMAP methods for reducing the di-

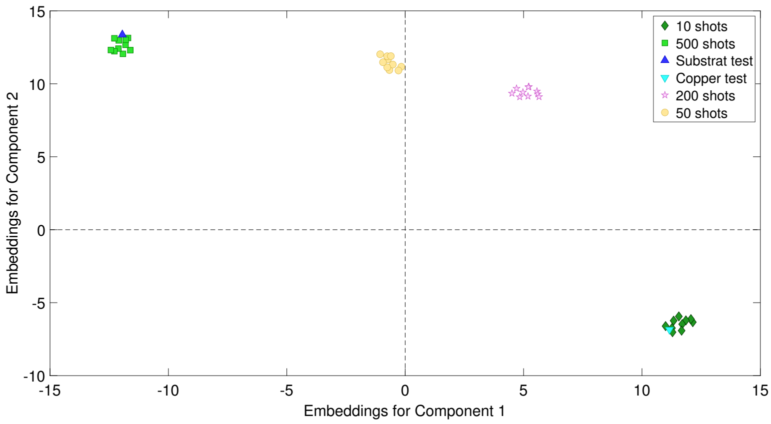

Figure 4. Embeddings plot for two UMAP components.

mensionality of spectral data. Our findings have demonstrated that the UMAP method stands out as the most promising candidate for pinpointing the precise moment when the copper layer is completely ablated.

Acknowledgements. This work was financially supported by funding provided by the Institute of Physics Belgrade, through the grant by the Ministry of Science and Technological Development of the Republic of Serbia.

References

Bellou, E., Gyftokostas, N., Stefas, D., Gazeli, O., & Couris, S., Laser-induced breakdown spectroscopy assisted by machine learning for olive oils classification: The effect of the experimental parameters. 2020, *Spectrochimica Acta*, **163**, 105746, DOI: 10.1016/j.sab.2019.105746

Boucher, T. F., Ozanne, M. V., Carmosino, M. L., et al., A study of machine learning regression methods for major elemental analysis of rocks using laser-induced breakdown spectroscopy. 2015, *Spectrochimica Acta*, **107**, 1, DOI: 10.1016/j.sab.2015.02.003

Diaz, D., Molina, A., & Hahn, D. W., Laser-Induced Breakdown Spectroscopy and Principal Component Analysis for the Classification of Spectra from Gold-Bearing Ores. 2020, *Applied Spectroscopy*, **74**, 42, DOI: 10.1177/0003702819881444

Hotelling, H., Analysis of a Complex of Statistical Variables into Principal Components. 1933, *Journal of Educational Psychology*, **24**, 417, DOI: http://dx.doi.org/10.1037/h0071325

Karhunen, K. 1947, *Ueber lineare Methoden in der Wahrscheinlichkeitsrechnung*

Lévy, P. 1948, *Processus stochastiques et mouvement brownien: suivi d'une note de M. Loève*

McInnes, L., Healy, J., & Melville, J., UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. 2018, *arXiv e-prints*, arXiv:1802.03426, DOI: 10.48550/arXiv.1802.03426

McInnes, L., Healy, J., & Melville, J., UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. 2020, *arXiv e-prints*, arXiv:1802.03426v3, DOI: 10.48550/arXiv.1802.03426

Moros, J., Lorenzo, J. A., Novotný, K., & Laserna, J. J., Fundamentals of stand-off Raman scattering spectroscopy for explosive fingerprinting. 2013, *Journal of Raman Spectroscopy*, **44**, 121, DOI: 10.1002/jrs.4138

Pořízka, P., Klus, J., Képeš, E., et al., On the utilization of principal component analysis in laser-induced breakdown spectroscopy data analysis, a review. 2018, *Spectrochimica Acta*, **148**, 65, DOI: 10.1016/j.sab.2018.05.030

Rabasovic, M. S., Marinkovic, B. P., & Sevic, D., Analysis of Printed Circuit Board LIBS Data Using Deep Learning. 2022, *Publications de l'Observatoire Astronomique de Beograd*, **102**, 219

Rabasović, M. S., Šević, D., Lukač, N., et al., Evaluation of laser-induced thin-layer removal by using shadowgraphy and laser-induced breakdown spectroscopy. 2016, *Applied Physics A: Materials Science & Processing*, **122**, 186, DOI: 10.1007/s00339-016-9697-3

Sainburg, T., McInnes, L., & Gentner, T. Q., Parametric UMAP embeddings for representation and semi-supervised learning. 2020, *arXiv e-prints*, arXiv:2009.12981, DOI: 10.48550/arXiv.2009.12981

Serranoa, J., Moros, J., Sánchez, C., Macías, J., & Laserna, J. J., Advanced recognition of explosives in traces on polymer surfacesusing LIBS and supervised learning classifiers. 2014, *Analytica Chimica Acta*, **806**, 107, DOI: 10.1016/j.aca.2013.11.035

Sevic, D., Rabasovic, M. S., P., G., Rabasovic, M. D., & Marinkovic, B. P., Principal Component Analysis of Printed Circuit Board LIBS Data. 2020, *Publications de l'Observatoire Astronomique de Beograd*, **99**, 117

Shiby, S. & Vasa, N. J., Nanosecond laser-assisted micro-scribing of a copper film on a dielectric material with laser-induced breakdown spectroscopy based monitoring. 2022, *Optics Laser Technology*, **147**, 107685, DOI: 10.1016/j.optlastec.2021.107685

Van der Maaten, J. P. L. & Hinton, G., Visualizing data using t-SNE. 2008, *Journal of Machine Learning Research*, **9**, 2579

Vrábel, J., Képeš, E., Duponchel, L., et al., Classification of challenging Laser-Induced Breakdown Spectroscopy soil sample data - EMSLIBS contest. 2020, *Spectrochimica Acta*, **169**, 105872, DOI: 10.1016/j.sab.2020.105872

Wise, B. M., Gallagher, N. B., Bro, R., et al. 2006, *Chemometrics tutorial for PLS Toolbox and Solo, ISBN: 0-9761184-1-6, Eigenvector Research, Inc. USA.*

Yang, Y., Hao, X., Zhang, L., & Ren, L., Application of Scikit and Keras Libraries for the Classification of Iron Ore Data Acquired by Laser-Induced Breakdown Spectroscopy (LIBS). 2020, *Sensors*, **20**, 1393, DOI: 10.3390/s20051393

Zhang, D., Zhang, H., Zhao, Y., et al., A brief review of new data analysis methods of laser-induced breakdown spectroscopy: machine learning. 2022, *Applied Spectroscopy Reviews*, **57**, 89, DOI: 10.1080/05704928.2020.1843175