



Tensor-based embedding for graph-based semi-supervised approaches

Georgoulas Ioannis
johnniegeo@mail.ntua.gr
National Technical University of Athens
Athens, Greece

Protopapadakis Eftychios
eftprot@uom.edu.gr
University of Macedonia
Thessaloniki, Greece

Makantasis Konstantinos
konstantinos.makantasis@um.edu.mt
University of Malta
Msida, Malta

Doulamis Anastasios
adoulam@cs.ntua.gr
National Technical University of Athens
Athens, Greece

ABSTRACT

This paper presents a novel approach to multiclass classification tasks, utilizing tensor-based embeddings for graph-based semi-supervised learning. The proposed method utilizes a tensor decomposition algorithm to create embeddings that capture the essential features of the data. These are used by various graph-based semi-supervised approach to construct a graph capable to propagate the information from labeled to unlabeled nodes, classifying available data. The proposed method was tested on hyperspectral datasets. The results demonstrate the potential of such combinatory tensor-based semi-supervised approaches.

CCS CONCEPTS

• **Computing methodologies** → **Machine learning.**

KEYWORDS

tensor-based embedding, graph-based, semi-supervised learning, hyperspectral data

ACM Reference Format:

Georgoulas Ioannis, Protopapadakis Eftychios, Makantasis Konstantinos, and Doulamis Anastasios. 2023. Tensor-based embedding for graph-based semi-supervised approaches. In *Proceedings of the 16th International Conference on Pervasive Technologies Related to Assistive Environments (PETRA '23)*, July 05–07, 2023, Corfu, Greece. ACM, New York, NY, USA, 6 pages. <https://doi.org/10.1145/3594806.3596550>

1 INTRODUCTION

The widespread adoption of digital technologies has resulted in the creation of large, complex, and heterogeneous data sets that are often characterized by a high degree of dimensionality, diversity, and variability, commonly referred to as high-order data, which present new opportunities and challenges for research. Semi-supervised

learning (SSL) approaches emerged naturally, as a promising solution for dealing with such data, as they allow exploiting both labeled and unlabeled data to learn predictive models [15].

In particular, graph-based SSL methods have gained significant attention due to their effectiveness in capturing the underlying structure of high-order data [14]. By representing the data as a graph, where each point corresponds to a node and the edges represent the pairwise similarity between them, these methods can propagate the label information from the labeled data to the unlabeled data in a smooth and coherent way. However, the performance of graph-based SSL methods heavily relies on the quality of the embeddings used to construct the graph, which should capture the essential features of the data while reducing its dimensionality.

Tensor-based embedding (TBE) techniques, such as tensor decomposition, can be used to extract low-dimensional representations of the data while preserving the underlying structure [8]. Tensor-based embeddings can capture complex patterns and relationships in the data, and can be particularly useful in applications where interpretability of the features is important. However, tensor-based embedding techniques can be computationally expensive, and may not be suitable for very large datasets.

In this paper we investigate the, possible, benefits of a synergistic approach between TBE and graph-based SSL. The idea lies in utilizing TBE, as an outcome of the encoder part, in a classification model, to facilitate the manifold creation, prior to the employment of a graph-based classifier. The main hypothesis is that the tensor-based embeddings will result in a better manifold projection, increasing the semi-supervised classifiers performance. The experimental scenarios involve semantic segmentation tasks in hyperspectral images (HSI).

2 RELATED WORK

Semi-supervised classification methods are particularly relevant to scenarios where labelled data is scarce. In those cases, it may be difficult to construct a reliable supervised classifier [6, 7, 16]. In practice, semi-supervised learning methods have also been applied to scenarios where no significant lack of labelled data exists: if the unlabelled data points provide additional information that is relevant for prediction, they can potentially be used to achieve improved classification performance [15].

He et al. [5] proposed a graph-based semi-supervised algorithm combined with particle cooperation and competition, which can



This work is licensed under a Creative Commons Attribution International 4.0 License.

PETRA '23, July 05–07, 2023, Corfu, Greece

© 2023 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-0069-9/23/07.

<https://doi.org/10.1145/3594806.3596550>

improve the model performance effectively by using unlabeled samples. The key idea lied in adopting a particle competition and cooperation mechanism into label propagation, which could detect and re-label misclassified samples dynamically, thus stopping the propagation of wrong labels and allowing the overall model to obtain better classification performance by using predicted labeled samples in HSI datasets.

Miao et al. [12] proposed a random multigraph and ensemble strategy to tackle hyperspectral imagery classification problems. The approach employs the construction multiple anchor graphs, for the label propagation, using the cross-entropy regularization. The entire setup tries to address all problems that come from the utilization of a single classifier, i.e. creation of one adjacent graph, which fails to effectively learn the complex structures and intrinsic properties of HSI.

Protopapadakis et al. [13] presented a two-step approach for NIR image classification based on stacked autoencoders and multiple SSL approaches. The encoder part extracted the underlying features of the and SSL approaches utilize them to create soft labels for the unlabeled instances, leveraging both the provided spectral and spatial information. Then, a deep network (classifier) is trained using custom loss functions, that encouraged the embeddings to preserve the spatial structure of the data. Experimental results demonstrate the effectiveness of the proposed method in achieving state-of-the-art performance compared to other methods, including traditional classification algorithms and deep learning approaches.

Tensor-based approaches are highly useful for handling complex data structures that have multiple modes or dimensions. These approaches can extract and model relationships between features that may not be apparent using traditional methods, leading to more accurate and meaningful insights. These methods have been successfully applied in a wide range of applications, such as computer vision, natural language processing, and bioinformatics, demonstrating their broad utility and versatility.

The work of Makantasis et al. [8] introduced a Rank-R Feedforward Neural Network for HSI analysis. In this non-linear classifier, the network weights are constrained to satisfy a rank-R Canonical Polyadic Decomposition, which promotes the learning capabilities, avoiding underfitting problems, related to limited data availability. Such models have significantly less trainable parameters and work well under noisy data inputs.

Tensor-based learning is highly beneficial in HSI analysis, as it can handle multi-dimensional data and extract meaningful features that may not be apparent using traditional methods. Additionally, semi-supervised learning offers advantages, as it can leverage the large amounts of unlabeled data to improve the accuracy of classification models, which is often a challenging task in HSI data analysis. Nevertheless, there is limited work on synergies among these research fields.

3 PROBLEM FORMULATION

In this work we address a multi-class classification problem, using as inputs hyper-spectral data. There is a collection of data points, l in total, denoted as: $D_L = ((x_i, y_i))_{i=1}^l$. The object $x_i \in \mathcal{X}$, will be used as an input, and has an associated label y_i . Further details on how the x values are produced, given an input space \mathcal{X} , can

be found in subsection 4.1. Additionally, to D_L , it is likely to have access to another collection of unlabeled data point, denoted as: $D_U = (x_i)_{i=l+1}^{l+u}$.

Let $G = (V, E)$ denote a graph, where V is a set of $|V| = n$ nodes and $E \subseteq V \times V$ is a set of $|E|$ edges between nodes, and $\mathbf{W} \in \mathbb{R}^{n \times n}$ are the associated weights on the edges. Graph-based approaches, which are by nature transductive, have as an objective to learn a predictive function $f : G, Y_L \rightarrow Y_U$, to infer the missing labels Y_U , for the unlabeled nodes [4]. Consequently, a graph-based SSL approach starts by creating the graph and then propagates the information to the nodes.

The graph creation can occur either in an unsupervised way, e.g. kernel-based or locally-linear reconstruction methods, among many other approaches. Yet, instead of utilizing the raw data values, $\{x_i\}_{i=1}^n$, we exploit the corresponding TBES, $\{\Phi(x_i)\}_{i=1}^n$ ones, generated by applying a tensor-based approach described in subsection 3.2. The label propagation has, also, many approaches; the adopted ones are explained in subsection 3.1.

3.1 Graph-based approaches

A set of three different graph-based SSL approaches are utilized, to investigate the impact of TBE in performance, namely: a) label propagation, b) label spreading and c) properly-weighted Laplace learning. In the following paragraphs a brief description is provided.

The label propagation approach is described in the work of [18]. As the name suggests, the approach operates over a graph structure to to propagate the labels of nearby labeled data points to the unlabeled points. This process can be repeated iteratively, although is not always necessarily [1], improving the accuracy of the labels assigned to the unlabeled data.

Assume a matrix $Y \in \mathbb{R}^{N \times c}$, where $N = l + u$ and c denotes the number of classes. Also, for convenience, let us adopt the following notation: $Y = [Y_L Y_U]^T$, where Y_L denotes the known and Y_U the unknown labels. Matrix Y rows can be interpreted as the probability of a point to correspond to a specific class. Also, assume a (given) graph with a weight matrix $\mathbf{W} \in \mathbb{R}^{N \times N}$, and a degree matrix $\mathbf{D} \in \mathbb{R}^{N \times N}$, defined as a diagonal matrix, whose i -th entry is given by: $D_{ii} = \sum_j W_{ij}$. Then, the labels' update is provided according to the following equation: $\mathbf{F}^{(t+1)} = \mathbf{D}^{-1} \mathbf{W} \mathbf{F}^{(t)}$, where, usually, $\mathbf{F}^{(0)} = Y$.

Label spreading has many similarities to label propagation [17], but adds a regularization to be more robust to noise. It adopts a slightly different approach over the effect of the label distributions. In particular, the updated soft labels are provided according to the following formula: $\mathbf{F}^{(t+1)} = \alpha \mathbf{S} \mathbf{F}^{(t)} + (1 - \alpha) Y$. In this case, matrix S is defined as: $S = \mathbf{D}^{-1/2} \mathbf{W} \mathbf{D}^{-1/2}$. Also, the matrix \mathbf{W} is defined as: $W_{ij} = \exp(-\|x_i - x_j\|^2) / 2\sigma^2$, if $i \neq j$, and $W_{ii} = 0$. Laplace learning, also called label propagation, may result in poor performance, attributed to the fact that the solutions develop localized spikes near the labeled points and are almost constant far from the labeled points [2]. In this work we adopted a properly weighted Laplacian approach, as described in [3].

3.2 Tensor-based embeddings

In the context of this study, tensor embeddings correspond to one-dimensional vector representations of tensor data points (three-dimensional objects) produced by the Rank-R FNN model proposed and experimentally and theoretically validated in [8–11]. The Rank-R FNN model is a two-layer neural network that receives as input D-dimensional tensor objects. The inputs are propagated to the hidden layer via a set of weights on which the Canonical-Polyadic (CP) decomposition has been applied. CP decomposition can significantly reduce the number of trainable parameters, making the Rank-R FNN model very efficient for small sample setting problems that involve high dimensional data, like data points in tensor format.

Specifically, the Rank-R FNN models the weights connecting input layer to hidden layer as:

$$\mathbf{W}^{(q)} = \llbracket \mathbf{W}_1^{(q)}, \dots, \mathbf{W}_D^{(q)} \rrbracket \in \mathbb{R}^{I_1 \times \dots \times I_D}, \quad (1)$$

or else

$$\text{vec}(\mathbf{W}^{(q)}) = (\mathbf{W}_D^{(q)} \odot \dots \odot \mathbf{W}_1^{(q)}) \mathbf{1}_R \in \mathbb{R}^{\prod_{d=1}^D I_d}, \quad (2)$$

where $\mathbf{1}_R$ stands for a vector with R ones, q stands for the number of hidden layer's neurons, and " \odot " operator to the Khatri-Rao product. The output of the q -th hidden neuron of the Rank-R FNN is given by

$$u_q = g(\langle \mathbf{W}^{(q)}, \mathbf{X} \rangle) = \text{trace}\left(g\left(\mathbf{W}_d^{(q)T} \mathbf{Z}_{\neq d}^{(q)}\right)\right), \quad (3)$$

where \mathbf{X} corresponds to a D-dimension tensor input, $\mathbf{W}_d^{(q)} \in \mathbb{R}^{I_d \times R}$, R is the rank for the CP decomposition and

$$\mathbf{Z}_{\neq d}^{(q)} = \mathbf{X}_{(d)} (\mathbf{W}_D^{(q)} \odot \dots \odot \mathbf{W}_{d+1}^{(q)} \odot \mathbf{W}_{d-1}^{(q)} \odot \dots \odot \mathbf{W}_1^{(q)}) \quad (4)$$

with $\mathbf{X}_{(d)}$ to denote the mode- d matricization of tensor \mathbf{X} .

After the transformation of the tensor input into a vector at the first hidden layer, the information is propagated to the output layer of the Rank-R FNN. The parameters of the model are estimated based on the available labelled data using an appropriate loss function (e.g., cross-entropy for classification problems of mean squared error for regression), the back-propagation algorithm and gradient-based optimization (for more information about the training procedure, we refer the interested reader to [10]).

In this study, the tensor input embeddings correspond to the output of the hidden layer of the trained Rank-R FNN. In other words, the tensor inputs are represented by q -dimensional vectors that are obtained by the output of the penultimate layer of the Rank-R FNN.

4 EXPERIMENTAL SETUP

The proposed schemes are evaluated over HSI images, in a semantic segmentation setup. A hyperspectral image is a 3-order tensor of dimensions $p_1 \times p_2 \times p_3$, where p_1 and p_2 correspond to height and width of the image, while p_3 corresponds to its spectral bands.

To conduct pixel-wise classification, i.e., to classify each pixel $I_{x,y}$ at location (x, y) according to the material it depicts, we follow the approach proposed in [8]. Specifically, it is assumed that the label of a square patch $\mathbf{X}_{x,y}$ of size $s \times s \times p_3$ centered at (x, y) has the same label with pixel $I_{x,y}$. Denoting as $\mathbf{t}_{x,y}$ the ground truth label of $I_{x,y}$, we form the dataset $\mathcal{D} = \{(\mathbf{X}_{x,y}, \mathbf{t}_{x,y})\}$ for training and evaluation purposes. In the experiments we set parameter s

equal to 5 or 7. That way we exploit spatial information of pixels, and, at the same time, satisfy the assumption that, for the majority of pixels the square patch $\mathbf{X}_{x,y}$ has same label as $I_{x,y}$. Lastly, the CP decomposition rank was set as, $R = 5$, and we worked with 75-dimensional vectors, i.e. $q = 75$.

4.1 Dataset description

Three widely known and publicly available datasets, captured by three different sensors, are used. In particular, we use i) the Salinas (corrected) dataset, which has been captured by AVIRIS sensor and consists of 224 spectral bands, minus 20 water absorption bands, and 5000 labeled pixels assigned to 16 different classes, ii) the Pavia University dataset, which has been captured by ROSIS sensor and consists of 103 spectral bands and 42,776 labeled pixels assigned to 9 different classes, and iii) the Botswana dataset, which has been captured by Hyperion sensor and consists of 145 spectral bands and 3,248 labeled pixels assigned to 14 different classes. Figure 1 presents two of the employed datasets, along with their ground truth.

4.2 Data preprocessing

In this study, we choose to train the models using a limited number of samples. This way, we focus on models' capacity to learn small sample setting classification tasks; employing a small number of training samples is a common limitation in many real-world applications such as HSI. For this reason, in our experiments, we vary the number of samples per class used for training to evaluate the proposed tensor-based model's performance when the number of training data is limited. In particular, we randomly select a specific number of samples α per class for training, while the rest are used for testing purposes.

In our case, the samples per class used for training are $\alpha = 10$ and 50. If some class includes fewer samples, we select a portion of 50% randomly for training. A stratified 10-fold cross-validation approach has been adopted for evaluating the performance of the adopted graph-based SSL approaches. This approach ensures that the distribution of classes in each fold is representative of the overall distribution of classes in the entire data set.

4.3 Experimental results

Fig. 2 demonstrates the average accuracy and F1 score for the test sets. For Pavia University and Salinas (corrected) region, the weighted Laplacian approach slightly exceeds the performance of the TNN. Yet, for the Botswana case, the performance is significantly worse. The remaining two approaches, i.e. LPC and LSC maintain a similar performance in all three investigated areas.

Fig. 3 illustrates the impact of the window size, during data preparation, on the overall classifiers' performance. The findings indicate that for the weighted Laplacian approach, an increase in window size adds on average 10% in F1 score. The reason that TNN appears better is attributed to the low performance in Botswana set, as described in fig. 2.

Fig. 4 presents comprehensive information regarding the influence of varying window sizes during the data preparation stage on the classification performance. Notably, the classification scores for the Botswana dataset exhibit significant enhancement. Conversely,

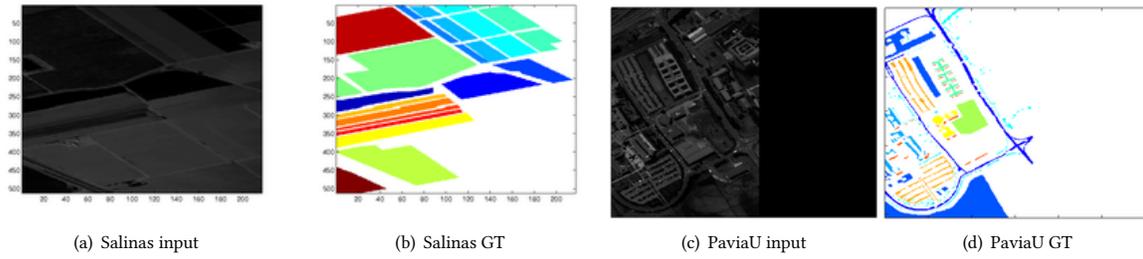


Figure 1: Some of the datasets utilized



Figure 2: Average performance scores, for the test sets (10 in total)

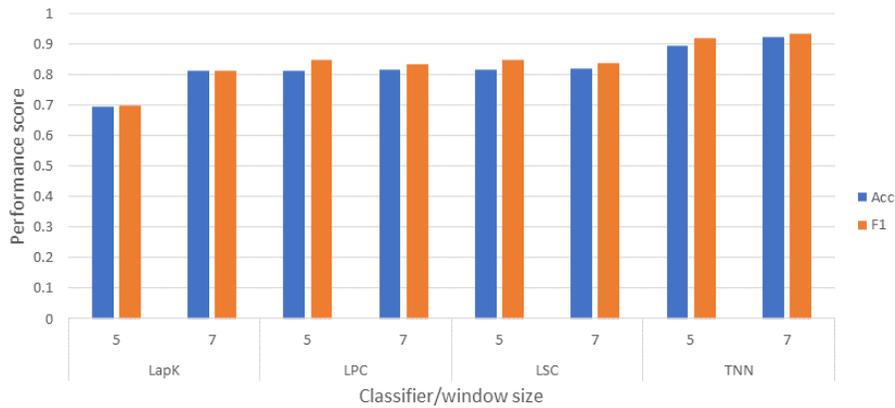


Figure 3: Performance variation depending on the size of the window.

the results for the Pavia University dataset demonstrate a marginal decline with an increase in window size. On the other hand, a slight improvement is observed for the "Salinas" dataset. It is important to note that the outcomes for both the Salinas and Pavia University datasets do not exhibit statistical significance.

Another intriguing topic involves the possible variations in classification outputs. Fig. 5 demonstrates such case for the Salinas

(corrected) dataset. Two quick notes involve: a) the similarity between LPC and LSC, where minor fluctuations are observed, b) the difference between weighted Laplacian and Tensor-based approach. The former is explained by the implementation similarity between LSC and LPC. The latter case indicates that the graph-based approach does not scatter miss-classified pixels in as many classes as the TNN or, if it did, it's less than the TNN.

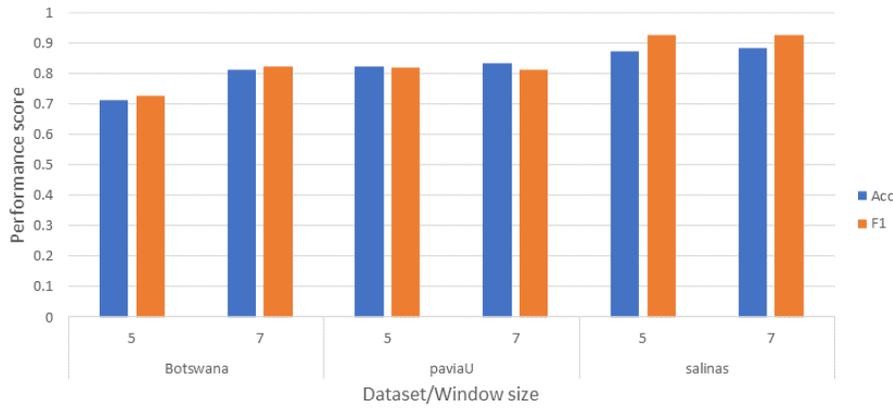


Figure 4: Performance variation depending on the size of the window.

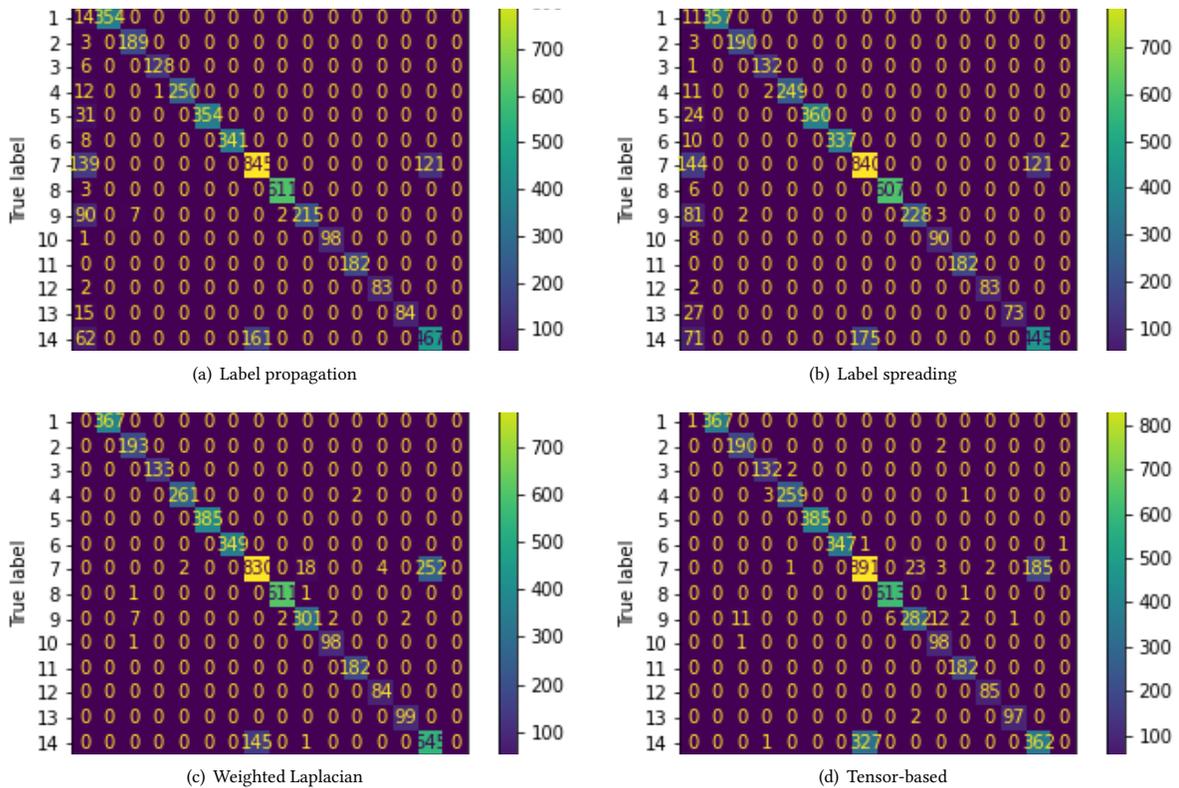


Figure 5: Variations in confusion matrices, given the adopted classification techniques

5 CONCLUSIONS

The use of tensor-based embeddings in graph-based SSL classification task could be beneficial in HSI semantic segmentation. The experiments conducted in this study demonstrate the potential of such combinatory schemes. By leveraging the spatial and spectral information present in hyperspectral data, tensor-based embeddings can effectively extract useful features for segmentation tasks.

Although the proposed method performs well, there is still room for further research and development. The impact of various parameters and hyperparameters on the performance of the method can be explored, and the generalization of the method to other datasets can be studied. In addition, the method's efficiency can be improved by adopting different computation methods for tensor-based embeddings and designing more efficient graph-based SSL approaches.

Overall, the proposed method has significant potential for various applications in remote sensing, agriculture, and environmental monitoring, among others, and could pave the way for more accurate and efficient hyperspectral image segmentation in the future.

ACKNOWLEDGMENTS

This paper is supported by the H2020 Photonics project “A Cost-Effective Photonics-based Device for Early Prediction, Monitoring and Management of Diabetic Foot Ulcers” funded under the ICT H2020 framework and the grand agreement no. 871908.

REFERENCES

- [1] Zalan Bodo and Lehel Csató. 2015. A note on label propagation for semi-supervised learning. *Acta Universitatis Sapientiae, Informatica* 7, 1 (2015), 18–30.
- [2] Jeff Calder, Brendan Cook, Matthew Thorpe, and Dejan Slepčev. 2020. Poisson learning: Graph based semi-supervised learning at very low label rates. In *International Conference on Machine Learning*. PMLR, 1306–1316.
- [3] Jeff Calder and Dejan Slepčev. 2020. Properly-weighted graph Laplacian for semi-supervised learning. *Applied mathematics & optimization* 82 (2020), 1111–1159.
- [4] Wenzheng Feng, Jie Zhang, Yuxiao Dong, Yu Han, Huanbo Luan, Qian Xu, Qiang Yang, Evgeny Kharlamov, and Jie Tang. 2020. Graph random neural networks for semi-supervised learning on graphs. *Advances in neural information processing systems* 33 (2020), 22092–22103.
- [5] Ziping He, Kewen Xia, Tiejun Li, Baokai Zu, Zhixian Yin, and Jiangnan Zhang. 2021. A constrained graph-based semi-supervised algorithm combined with particle cooperation and competition for hyperspectral image classification. *Remote Sensing* 13, 2 (2021), 193.
- [6] Maria Kaselimi, Athanasios Voulodimos, Eftychios Protopapadakis, Nikolaos Doulamis, and Anastasios Doulamis. 2020. EnerGAN: A GENERATIVE ADVERSARIAL NETWORK FOR ENERGY DISAGGREGATION. In *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 1578–1582. <https://doi.org/10.1109/ICASSP40776.2020.9054342>
- [7] Dimitrios I. Kosmopoulos, Athanasios S. Voulodimos, and Anastasios D. Doulamis. 2013. A System for Multicamera Task Recognition and Summarization for Structured Environments. *IEEE Transactions on Industrial Informatics* 9, 1 (2013), 161–171. <https://doi.org/10.1109/TII.2012.2212712>
- [8] Konstantinos Makantasis, Anastasios Doulamis, Nikolaos Doulamis, Antonis Nikitakis, and Athanasios Voulodimos. 2018. Tensor-based nonlinear classifier for high-order data analysis. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2221–2225.
- [9] Konstantinos Makantasis, Anastasios D Doulamis, Nikolaos D Doulamis, and Antonis Nikitakis. 2018. Tensor-based classification models for hyperspectral data analysis. *IEEE Transactions on Geoscience and Remote Sensing* 56, 12 (2018), 6884–6898.
- [10] Konstantinos Makantasis, Alexandros Georgogiannis, Athanasios Voulodimos, Ioannis Georgoulas, Anastasios Doulamis, and Nikolaos Doulamis. 2021. Rank-r fnn: A tensor-based learning model for high-order data classification. *IEEE Access* 9 (2021), 58609–58620.
- [11] Konstantinos Makantasis, Athanasios Voulodimos, Anastasios Doulamis, Nikolaos Doulamis, and Ioannis Georgoulas. 2019. Hyperspectral image classification with tensor-based rank-R learning models. In *2019 IEEE International Conference on Image Processing (ICIP)*. IEEE, 3148–3125.
- [12] Yanling Miao, Mulin Chen, Yuan Yuan, Jocelyn Chanussot, and Qi Wang. 2021. Hyperspectral imagery classification via random multigraphs ensemble learning. *IEEE journal of selected topics in applied earth observations and remote sensing* 15 (2021), 641–653.
- [13] Eftychios Protopapadakis, Anastasios Doulamis, Nikolaos Doulamis, and Evangelos Maltezos. 2021. Stacked autoencoders driven by semi-supervised learning for building extraction from near infrared remote sensing imagery. *Remote Sensing* 13, 3 (2021), 371.
- [14] Zixing Song, Xiangli Yang, Zenglin Xu, and Irwin King. 2022. Graph-based semi-supervised learning: A comprehensive review. *IEEE Transactions on Neural Networks and Learning Systems* (2022).
- [15] Jesper E Van Engelen and Holger H Hoos. 2020. A survey on semi-supervised learning. *Machine learning* 109, 2 (2020), 373–440.
- [16] Athanasios Voulodimos, Dimitrios Kosmopoulos, Galina Veres, Helmut Grabner, Luc Van Gool, and Theodora Varvarigou. 2011. Online classification of visual tasks for industrial workflow monitoring. *Neural Networks* 24, 8 (2011), 852–860. <https://doi.org/10.1016/j.neunet.2011.06.001> Artificial Neural Networks: Selected Papers from ICANN 2010.
- [17] Dengyong Zhou, Olivier Bousquet, Thomas Lal, Jason Weston, and Bernhard Schölkopf. 2003. Learning with local and global consistency. *Advances in neural information processing systems* 16 (2003).
- [18] Xiaojin Zhu and Zoubin Ghahramanin. 2002. Learning from labeled and unlabeled data with label propagation. (2002).