https://doi.org/10.2352/ISSN.2470-1173.2017.17.COIMG-445

Deep Convolutional Neural Networks for the Classification of Snapshot Mosaic Hyperspectral Imagery

***Konstantina Fotiadou*[1,2]*, Grigorios Tsagkatakis*[1]*, Panagiotis Tsakalides*[1,2]**
[1] ***ICS- Foundation for Research and Technology - Hellas (FORTH), Crete, Greece***
[2] ***Department of Computer Science, University of Crete , Greece***

Abstract

Spectral information obtained by hyperspectral sensors enables better characterization, identification and classification of the objects in a scene of interest. Unfortunately, several factors have to be addressed in the classification of hyperspectral data, including the acquisition process, the high dimensionality of spectral samples, and the limited availability of labeled data. Consequently, it is of great importance to design hyperspectral image classification schemes able to deal with the issues of the curse of dimensionality, and simultaneously produce accurate classification results, even from a limited number of training data. To that end, we propose a novel machine learning technique that addresses the hyperspectral image classification problem by employing the state-of-the-art scheme of Convolutional Neural Networks (CNNs). The formal approach introduced in this work exploits the fact that the spatio-spectral information of an input scene can be encoded via CNNs and combined with multi-class classifiers. We apply the proposed method on novel dataset acquired by a snapshot mosaic spectral camera and demonstrate the potential of the proposed approach for accurate classification.

Introduction

Recent advances in optics and photonics are addressing the demand for designing Hyperspectral Imaging (HSI) systems with higher spatial and spectral resolution, able to revile the physical properties of the objects in a scene of interest [1]. This type of data is crucial for multiple applications, such as remote sensing, precision agriculture, food industry, medical and biological applications [2]. Although hyperspectral imaging systems demonstrate substantial advantages in structure identification, HSI acquisition and processing stages, usually introduce multiple factorial constraints. Slow acquisition time, limited spectral and spatial resolution, and the need for linear motion in the case of traditional spectral imagers, are just a few of the limitations that hyperspectral sensors admit, and must be addressed. Snapshot Spectral imaging addresses that problem by sampling the full spatio-spectral cube during each exposure, though a mapping of pixels to specific spectral bands [13], [14], [15].

High spatial and spectral resolution hyperspectral imaging systems demonstrate significant advantages concerning object recognition and material detection applications, by identifying the subtle differences in spectral signatures of various objects. This discrimination of materials based on their spectral profile, can be considered as a classification task, where groups of hyperpixels are labeled to a particular class based on their reflectance properties, exploiting training examples for modeling each class. State-of-the-art hyperspectral classification approaches are composed of two steps; first, hand-crafted feature descriptors are extracted from the training data and second the computed features are used to train classifiers, such as such Support Vector Machines (SVM) [3]. Feature extraction is a significant process in multiple computer vision tasks, such as object recognition, image segmentation and classification. Traditional approaches consider carefully designed hand-crafted features, such as the Scale Invariant Feature Transform (SIFT) [5], or the Histogram of Oriented Gradients (HoG) [6]. Despite their impressive performance, they are not able to efficient encode the underlying characteristics of higher dimensional image data, while significant human intervention is required during the design.

In hyperspectral imagery, various feature extraction techniques have been proposed, including decision boundary feature extraction (DBFE) [7] and Kumar's *et al.* scheme [8], based on a combination of highly correlated adjacent spectral bands into fewer features by means of top-down and bottom-up algorithms. Additionally, in Earth monitoring remote sensing applications, characteristic feature descriptors are the Normalized Vegetation Difference Index (NDVI) and the Land Surface Temperature(LST). Nevertheless, it is extremely difficult to discover which features are significant for each hyperspectral classification task, due to the high diversity and heterogeneity of the acquired materials. This motivates the need for efficient feature representations directly extracted from input data through deep representation learning [10], a cutting edge paradigm aiming to learn discriminative and robust representations of the input data for use in higher level tasks.

The objective of this work is to propose a novel approach for discriminating between different objects in a scene of interest, by introducing a deep feature learning based classification scheme on snapshot mosaic hyperspectral imagery. The proposed system utilizes the Convolutional Neural Networks (CNN) [9] and several multi-class classifiers, in order to extract high-level representative spatio-spectral features, substantially increasing the performance of the subsequent classification task. Unlike traditional hyperspectral classification techniques that extract complex hand-crafted features, the proposed algorithm adheres to a machine learning paradigm, able to work even with a small number of labeled training data. To the best of our knowledge, the proposed scheme is the first deep learning-based technique focused on the classification of the hyperspectral snapshot mosaic imagery.

A visual description of the proposed scheme is presented in Figure 1 where we present the block diagram for the classification of snapshot mosaic spectral imagery, using am deep CNN. The rest of the paper is organized as follows. Section 2 presents a brief review of the related work concerning deep learning ap-

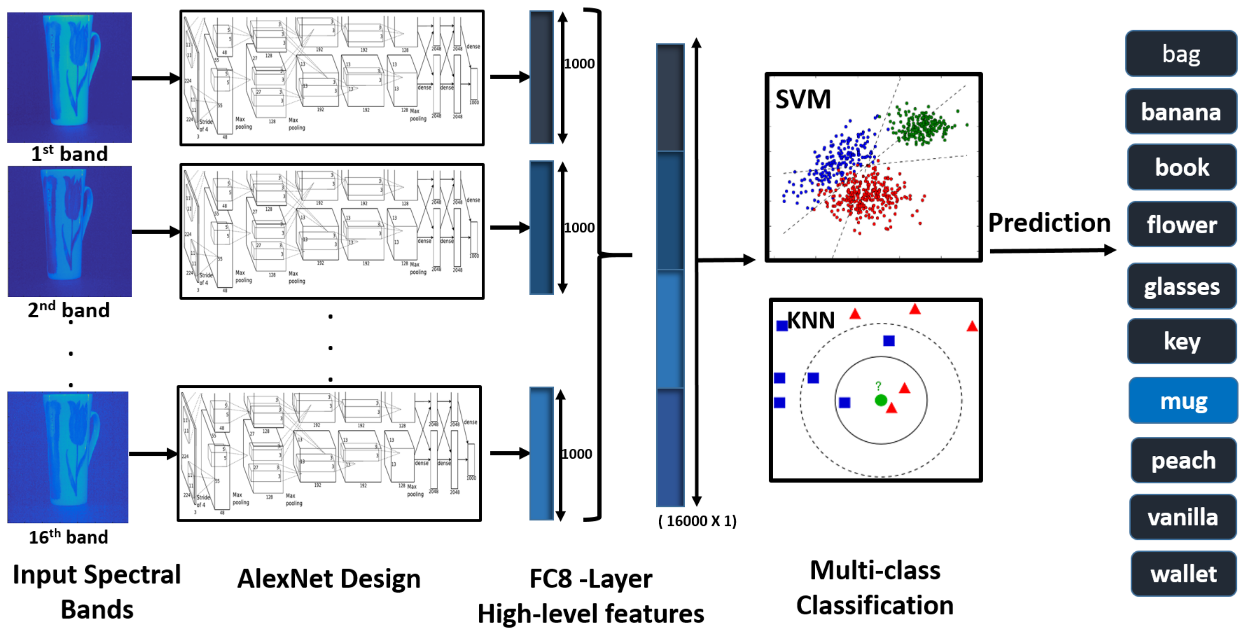

Figure 1: **Block diagram of the proposed scheme**: Our system decomposes the input hypercubes into their distinct spectral bands, and extracts AlexNet-based high level features for each spectral observation. The concatenated feature vectors are given as inputs to multi-class classifiers in order to implement the final prediction.

proaches for the classification of hyperspectral data. In Section 3 we outline the key theoretical components of the CNN. Section 4 provides an overview of the generated hyperspectral dataset along with experimental results, while the paper concludes in Section 5.

Related Work

Deep learning (DL) is a special case of representation learning which aims at learning multiple hierarchical levels of representations, leading to more abstract features that are more beneficial in classification [16]. Recently, DL has been considered for various problems in the remote sensing community, including land cover detection [17], [18], building detection [21], and scene classification [22]. Specifically, the authors in [17] considered the paradigm of Stacked Sparse Autoencoders (SSAE) as a feature extraction mechanism for multi-label classification of hyperspectral images. Another feature learning approach for the problem of multi-label land cover classification was proposed in [18], where the authors utilize single and multiple layer Sparse Autoencoders in order to learn representative features able to facilitate the classification process of MODIS multispectral data.

In addition to the Sparse Autoencoders framework, over the past few years, CNNs have been established as an effective class of models for understanding image content, producing significant performance gains in image recognition, segmentation, and detection, among others [19], [20]. However, a major limitation of CNNs is the extensively long periods of training time necessary to effectively optimize the large number of the parameters that are considered. Although, it has been shown that CNNs achieve superior performance on a number of visual recognition tasks, their hard computational requirements have limited their application in a handful of hypespectral feature learning and classification tasks. Recently, the authors in [23] utilize CNN's for large scale remote sensing image classification, and propose an efficient procedure in order to overcome the problem of inefficient training data. Additionally, in [24] the authors design a CNN able to extract spatio-spectral features for classification purposes. Another class of techniques, solve the classification problem by extracting the principal components of the hyperspectral scenes and incorporating convolutions only at the spatial domain [25], [31].

Feature Learning for Classification

The main purpose of this work is to classify an N spectral band image, utilizing both its spatial and spectral dimensions. In order to accomplish this task, we employ a sequence of filters $\mathbf{w}_{x,y}$, of size $(m \times m)$, which are convolved with the "hyper-pixels" of the spectral cube, aiming at encoding spatial invariance. To achieve scale invariance, each convolution layer is followed by a pooling layer. These learned features are considered as input to a multi-class SVM classifier, in order to implement the labelling task. In the following section, we present CNNs formulation and how they can be applied in the concept of snapshot mosaic hyperspectral classification.

Convolutional Neural Networks

While in fully-connected deep neural networks, the activation of each hidden unit is computed by multiplying the entire input by the correspondent weights for each neuron in that layer, in CNNs, the activation of each hidden unit is computed for a small input area. CNNs are composed of convolutional layers which alternate with subsampling (pooling) layers, resulting in a hierarchy of increasingly abstract features, optionally followed by fully connected layers to carry out the final labeling into categories. Typically, the final layer of the CNN produces as many outputs as the number of categories, or a single output for the case of binary labeling.

At the convolution layer, the previous layer's feature maps are first convolved with learnable kernels and then are passed through the activation function to form the output feature map. Specifically, let $n \times n$ be a square region extracted from a training input image $\mathbf{X} \in \mathbb{R}^{N \times M}$, and $\mathbf{w}$ be a filter of kernel size $(m \times m)$. The output of the convolutional layer $\mathbf{h}$, of size $(n-m+1) \times (n-m+1)$ is formulated as:

$$\mathbf{h}_{ij}^{\ell} = f\Big(\sum_{k=0}^{m-1}\sum_{l=0}^{m-1} \mathbf{w}_{ab}\mathbf{x}_{(i+k)(j+l)}^{\ell-1} + \mathbf{b}_{ij}^{\ell}\Big),$$

where $\mathbf{b}$ is the additive bias term, and $f(\cdot)$ stands for the neuron's activation unit.

The activation function $f(\cdot)$, is a formal way to model a neurons output as a function of its input. Typical choices for the activation function are the logistic sigmoid function, defined as: $f(x) = \frac{1}{1+e^{-x}}$, the hyperbolic tangent function: $f(x) = tanh(x)$, and the Rectified Linear Unit (ReLU), given by: $f(x) = max(0, x)$. The majority of state-of-the-art approaches employ the ReLU as the activation function for the CNNs. The results and analysis carried out in [32] suggest that deep CNN with ReLU activation functions, train several times faster compared to equivalent designs with other activation function. Additionally, taking into consideration the training time required for the gradient descent process, the saturating gradients of non-linearities like the tanh and logistic sigmoid, lead to slower convergence time compared to the ReLU function.

The output of the convolutional layer is directly utilized as input to a sub-sampling (*i.e* pooling) layer that produces downsampled versions of the input maps. There are several types of pooling, two common types are the max- and the average-pooling. Pooling operators partition the input image into a set of non-overlapping or overlapping patches and output the maximum or average value for each such sub-region. By pooling, the model can reduce its computational complexity for upper layers, and can provide a form of translation invariance. Formally, this procedure is formulated as:

$$\mathbf{h}_{ij}^{\ell} = f(\beta_j^{\ell}\ \mathrm{down}(\mathbf{h}_{ij}^{\ell-1} + \mathbf{b}_{ij}^{\ell})),$$

where $\mathrm{down}(\cdot)$ stands for a sub-sampling function. This function sums over each distinct $(m \times m)$ block in the input image so that the output image is m-times smaller along the spatial dimensions. Additionally, β represents the multiplicative bias of the output feature map, while $\mathbf{b}$ is the additive bias.

Classification Algorithms

Classification is the process of learning from a set of classified objects a model that can predict the class of previously unseen objects. In this work, we deal with a multi-class classification problem, since we aim to discriminate between 10 distinct image categories. As a result, the proper selection of the classification algorithm is a critical step. In the following paragraphs we provide a brief overview of three *state-of-the-art* classification techniques that we have experimented with, namely: the multi-class Support Vector Machines, the K-Nearest Neighbours and the Decision Trees.

Support Vector Machines

The state-of-the-art linear support vector machines (SVM's) is originally formulated for binary classification tasks [11], [12]. Specifically, consider a set of training data along with their corresponding labels: $(\mathbf{x}_n, \mathbf{y}_n),\ n = 1, \cdots, N,\ \mathbf{x}_n \in \mathbb{R}^D,\ t_n\ \in \{1, +1\}$. SVM's solve the following constrained optimization problem:

$$\min_{\mathbf{w},\xi_n} \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_{n=1}^{N}\xi_n, \text{ subject to}$$

$$\mathbf{w}^T\mathbf{x}_n t_n \geq 1 - \xi_n\ \forall n, \xi_n \geq 0 \forall n,$$

where the slack variable ξ_n penalize data points that violate the margin requirements. The unconstrained and differentiable variation of the aforementioned equation is:

$$\min_{\mathbf{w}} \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_{n=1}^{N} max(1 - \mathbf{w}^T\mathbf{x}_n t_n, 0)^2$$

The class prediction of the testing data $\mathbf{x}$, outcomes from the solution of the following optimization problem:

$$\underset{t}{\mathrm{argmax}}(\mathbf{w}^T\mathbf{x})t$$

The majority of classification applications utilize the softmax layer objective in order to discriminate between the different classes. Nevertheless, in our approach, the extracted features from the Convolutional Neural Network, are directly utilized for *Multi-Label Classification* among the different hyperspectral image categories. For K class problems, K-linear SVMs will be trained independently, while the data from the rest classes form the negative cases. Consider the output of the k-th SVM as:

$$\alpha_k(x) = \mathbf{w}^T\mathbf{x}$$

Then, the predicted class is estimated by solving the following optimization problem:

$$\underset{k}{\mathrm{argmax}}\,\alpha_k(\mathbf{x})$$

K-Nearest Neighbour

The K-Nearest Neighbor algorithm (KNN) [26], [27], [28] is among the simplest of all machine learning classification techniques. Specifically, KNN classifies among the different categories based on the closest training examples in the feature space. The training process for this algorithm only consists of storing feature vectors and labels of the training images. In the classification process, the unlabelled testing data is assigned to the label of its K-nearest neighbours, while the testing data are classified based on the labels of their K-nearest neighbors by majority vote. The most common distance metric function for the KNN is the Euclidean distance, defined as:

$$d(\mathbf{x}, \mathbf{y}) = ||\mathbf{x} - \mathbf{y}|| = \sum_{i=1}^{m}(x_i - \mathbf{y}_i)$$

A key advantage of the KNN algorithm is that it performs well with multi-label classification problems, since the final prediction is based on a small neighbourhood of similar classes. Nevertheless, a major drawback of the KNN algorithm is that it uses all the features equally, leading to classification errors, especially when there is a small amount of training data.

Decision Trees

Decision Trees [29] are classification models in the form of tree graphs. The typical structure of a decision tree includes the root node, that contains all training data, a set of internal nodes, *i.e* the splits, and the set of terminal nodes, the leaves. In a decision tree, each internal node splits the feature space into two or more sub-spaces according to a certain discrete function of the input data. Consider $\mathbf{x}$ as the feature vector to be predicted. Then the

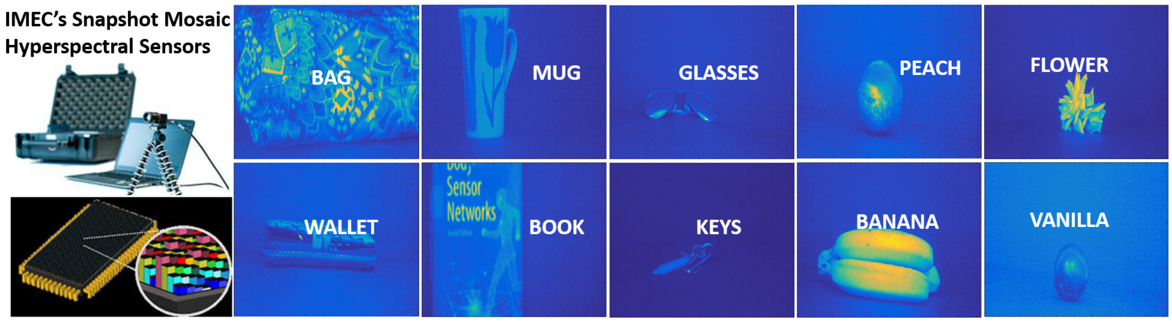

Figure 2: **Proposed Hyperspectral Dataset**: 10-Category Hyperspectral Image dataset, acquired by IMEC's Snapshot Mosaic Sensors.

value of **x**, goes through the nodes of the tree, and in each node **x** is tested whether it is higher or smaller than a certain threshold. Depending on the outcome, the process continues recursively in the right of left sub-tree, until a leaf is encounter. Each leaf contains a prediction that is returned. Typically, Decision Trees are learnt from the training data using a recursive greedy search algorithm. These algorithms are usually composed of three steps: splitting the nodes, determining which nodes are the terminal nodes and assigning the corresponding class labels to the terminal nodes [30].

Data Acquisition & Experimental Setup

In this section we explicitly describe the data acquisition process and the simulation results obtained through a thorough evaluation of the proposed hyperspectral classification scheme. To validate the merits of the proposed approach, we explored the classification of hyperspectral images acquired using a Ximea camera, equipped with the IMEC Snapshot Mosaic sensor [13], [14], [15]. These flexible sensors optically subsample the 3D spatio-spectral information on a two-dimensional CMOS detector array, where a layer of Faby-Perot spectral filters is deposited on top of the detector array. The hyperspectral data is initially acquired in the form of 2D mosaic images. In order to generate the 3D hypercubes, the spectral components are properly rearranged into separate spectral bands. In our experiments, we utilize a 4×4 snapshot mosaic hyperspectral sensor resolving 16 bands in the spectrum range of $470 - 630$ nm, with a spatial dimension of 256×512 pixels.

For genearation of the dataset, we considered 10 distinct object categories, namely: bag, banana, peach, glasses, wallet, book, flower, keys, vanilla and mug. Our hyperspectral dataset consists of 90 images. The images were acquired under different illumination conditions and from different view-points, thus producing the first snapshot mosaic spectral image dataset used for classification purposes. Fig. 2 presents an example of the proposed hyperspectral dataset.

Simulation Results

Each training hypercube encodes 16 spectral bands, where for each spectral observation, we extract high-level features using a pre-trained state-of-the-art CNN, the AlexNet [34]. AlexNet was trained on RGB images of size 227×227 from various categories. In order to comply with AlexNet's input image specifications, we downscale the spatial dimension of each spectral band and replicated each spectral bands to a three dimensional tensor. To train the classifier, the features corresponding to the FC8 fully-connected layer were extracted, mapping the input images to 1000-dimensional feature vectors. To quantify the capabilities of the proposed scheme, we experimented with different number of training images, ranging from the extremely limited case of 10, up to 50 training examples, and evaluate the performance on the remaining 40 spectral cubes, and report the results over 10 independent trials. The classification accuracy is defined as:

$$Accuracy = \frac{Number\ of\ Correct\ Predictions}{Total\ Number\ of\ Predictions}$$

Fig. 3 investigates the impact of the three comparable classification techniques, namely: the KNN, the Multi-Class SVM model using linear kernel, and the Decision Trees, on the proposed system's classification accuracy. Specifically, we illustrate the classification accuracy, with respect to the different number of randomly selected spectral observations.

Concerning the KNN classifier, we observe that the first scenario of using only 10 training images led to low classification performance, of **73**%, while the scenario where we use 50 training hypercubes achieves the best performance of **88**%. We observe that the number of training hypercubes has a great impact on the classification quality. As the number of training images, grows, the classification accuracy also grows. Additionally, when we utilize the highest possible spatio-spectral information, of all 16 acquired spectral bands, KNN classifier presents the highest classification accuracy.

In the second scenario, we investigate the classification accuracy when the SVM algorithm with linear kernel is utilized. We observe that the SVM classifier achieves the highest accuracy, of **86**% in the scenario where we use a large number of training data, *i.e.* 50 training images. In contrast, the minimum classification accuracy of **74**% is achieved when we use only 10 training hypercubes. In this classifier, we observe that among the different number of utilized spectral bands the classification accuracy does not depict serious variations.

Finally, in the last scenario we exploit the Decision Trees for the multi-class labelling of the proposed hyperspectral dataset. As we may observe, the Decision Trees achieve the lowest classification accuracy among the three comparable classification techniques. Specifically, in the first scenario, where we use only 10 input training hypercubes, the classification accuracy remains stable (25%) among the different number of utilized spectral bands. As the number of training images grows, the accuracy of the proposed classifier also grows. However, in comparison with the other two classification techniques, the Decision Trees achive the

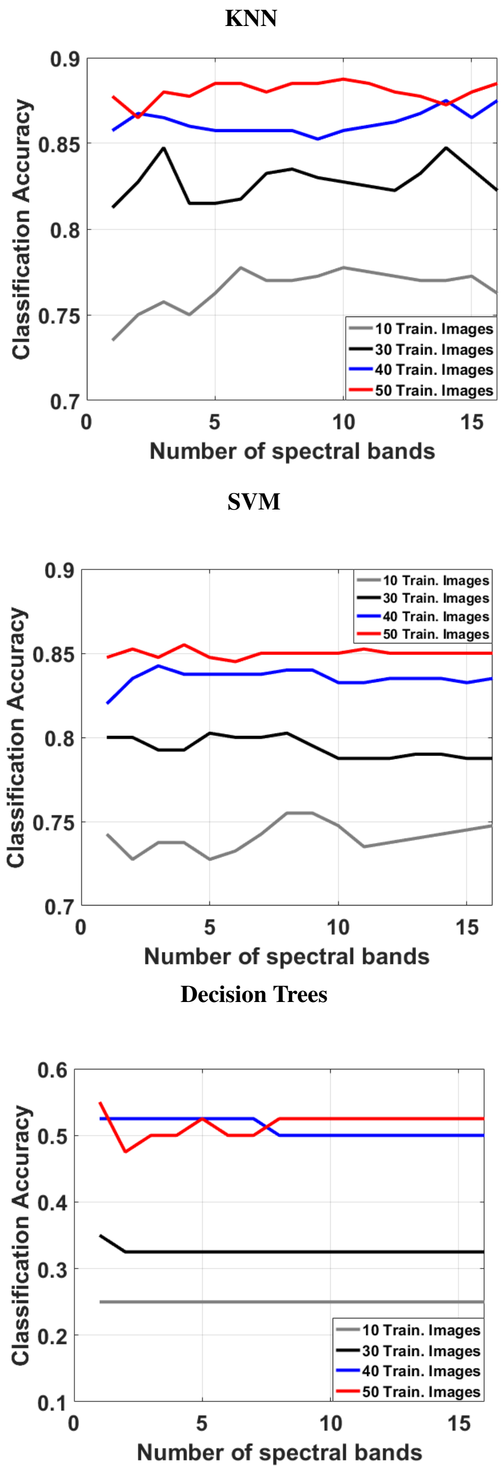

Figure 3: Classification Accuracy versus the number of spectral bands, tested on variant number of training images, for the three comparable classification techniques.

lowest performance, of **52%** in the proposed hyperspectral classification scheme. Comparing the three classication techniques, we observe the KNN classifier outperforms both the sophisticated SVM and the Decision trees techniques.

Conclusions and Future Work

Unlike traditional hyperspectral classification approaches that extract handcrafted features, in this work we propose a scheme which exploits a deep feature learning architecture for efficient feature extraction. The state-of-the-art method of CNNs is able to identify representative features, encoding both spatial and spectral variations of hyperspectral scenes, and successfully assign each hyper-cube to a predefined class. The proposed deep feature learning scheme is focused on the classification of snapshot mosaic hyperspectral imagery, while a new hyperspectral classification dataset of indoor scenes is constructed. Our flexible scheme can be easily extended in working with many more classes, and even with hyperspectral video sequences.

Acknowledgments

This work was funded by the PHySIS project, contract no. 640174, and by the DEDALE project contract no. 665044 within the H2020 Framework Program of the European Commission.

References

[1] J. M. Bioucas-Dias , A. Plaza, G. Camps-Valls, P. Scheunders, N. M. Nasrabadi, and J. Chanussot, Hyperspectral remote sensing data analysis and future challenges, Geoscience and Remote Sensing Magazine, IEEE, 1(2), pg. 6–36.(2013).

[2] N. Hagen, and M. W. Kudenov, Review of snapshot spectral imaging technologies, Optical Engineering 52(9), 090901-090901, (2013).

[3] F. Melgani and L. Bruzzone, Classification of hyperspectral remote-sensing images with support vector machines,IEEE Trans. Geosci. Remote Sensing, vol. 42, no. 8, pp. 1778-1790, (2004).

[4] G. Tsagkatakis, and P. Tsakalides, Compressed hyperspectral sensing, IS&T/SPIE Electronic Imaging, International Society for Optics and Photonics.(2015).

[5] D. G. Lowe(1999). Object recognition from local scale-invariant features. In Computer vision, 1999. The proceedings of the seventh IEEE international conference on (Vol. 2, pp. 1150-1157).

[6] N. Dalal, and B. Triggs (2005, June). Histograms of oriented gradients for human detection. In 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) (Vol. 1, pp. 886-893). IEEE.

[7] C. Lee, and D.A. Landgrebe(1993). Feature extraction based on decision boundaries. IEEE Transactions on Pattern Analysis and Machine Intelligence, 15(4), 388-400.

[8] S. Kumar, J. Ghosh, and M.M. Crawford (2001). Best-bases feature extraction algorithms for classification of hyperspectral data. IEEE Transactions on Geoscience and remote sensing, 39(7), 1368-1379.

[9] P. Y. Simard, D. Steinkraus, and J. C. Platt, . Best practices for convolutional neural networks applied to visual document analysis. In ICDAR (Vol. 3, pp. 958-962).

[10] Y. Bengio, A. Courville, and P. Vincent. Representation learning: A review and new perspectives. IEEE transactions on pattern analysis and machine intelligence 35.8 (2013): 1798-1828.

[11] J.A. Suykens, and J. Vandewalle (1999). Least squares support vector machine classifiers. Neural processing letters, 9(3), 293-300.

[12] Y.Tang (2013). Deep learning using linear support vector machines. arXiv preprint arXiv:1306.0239.

[13] B.Geelen, T.Nicolaas, L.Andy, A compact snapshot multispectral

imager with a monolithically integrated per-pixel filter mosaic, Spie Moems-Mems. Intern. Society for Optics and Photonics. (2014).

[14] B.Geelen, et al., A tiny VIS-NIR snapshot multispectral camera, SPIE OPTO. International Society for Optics and Photonics. (2015).

[15] A. Lambrechts, et al, A CMOS-compatible, integrated approach to hyper-and multispectral imaging, IEEE International Electron Devices Meeting (IEDM),(2014).

[16] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. Nature 521.7553 (2015): 436-444.

[17] G. Tsagkatakis, and P. Tsakalides. Deep Feature Learning for Hyperspectral Image Classification and Land Cover Estimation.ESA Symbosium, 2016.

[18] K. Karalas, et al: Deep learning for multi-label land cover classification, SPIE Remote Sensing. Intern. Society for Optics and Photonics. (2015).

[19] S. Lawrence, C. L. Giles, A.C. Tsoi, and A.D. Back (1997). Face recognition: A convolutional neural-network approach. IEEE transactions on neural networks, 8(1), 98-113.

[20] J. Long, E. Shelhamer, and T. Darrell. (2015). Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 3431-3440).

[21] M. Vakalopoulou, K. Karantzalos, N. Komodakis, and N. Paragios. Building detection in very high resolution multispectral data with deep learning features. In Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International, pages 18731876. IEEE, 2015.

[22] Y. Zhong, F. Fei, and L. Zhang. Large patch convolutional neural networks for the scene classification of high spatial resolution imagery. Journal of Applied Remote Sensing, 10(2):025006025006,2016.

[23] E. Maggiori, Y. Tarabalka, G. Charpiat, and P. Alliez (2017). Convolutional Neural Networks for Large-Scale Remote-Sensing Image Classification. In IEEE Transactions on Geoscience and Remote Sensing, 55(2), 645-657.

[24] Y. Chen, H. Jiang , C. Li, X. Jia, and P. Ghamisi(2016). Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. IEEE Transactions on Geoscience and Remote Sensing, 54(10), 6232-6251.

[25] J. Yue, W. Zhao, S. Mao, and H. Liu. Spectralspatial classification of hyperspectral images using deep convolutional neural networks. In Remote Sensing Letters, vol. 6, no. 6, pp. 468477, 2015.

[26] D. Bremner, E. Demaine, J. Erickson, J. Iacono, S. Langerman, P. Morin, G. Toussaint, Output-sensitive algorithms for computing nearestneighbor decision boundaries, Discrete and Computational Geometry, 2005, pp. 593604.

[27] T. Cover, P. Hart. Nearest-neighbor pattern classification, Information Theory, IEEE Transactions on, Jan. 1967, pp. 21-27.

[28] J. I. N. H. O. KIM, B. S.Kim, and S. Savarese (2012). Comparing image classification methods: K-nearest-neighbor and support-vector-machines. Ann Arbor, 1001, 48109-2122.

[29] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone. Classification and Regression Trees. Chapman & Hall, Boca Raton, 1993.

[30] P. N. Tan, M. Steinbach, and & V. Kumar (2006). Classification: basic concepts, decision trees, and model evaluation. Introduction to data mining, 1, 145-205.

[31] K. Makantasis, K. Karantzalos, A. Doulamis, and N. Doulamis. Deep supervised learning for hyperspectral data classification through convolutional neural networks. In IEEE IGARSS. IEEE, 2015, pp. 49594962.

[32] N. Srivastava, G.E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, (2014). Dropout: a simple way to prevent neural networks from overfitting. Journal of Machine Learning Research, 15(1), 1929-1958.

[33] A. Ng. "Sparse autoencoder." CS294A Lecture notes 72 (2011): 1-19.

[34] Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems.

[35] Y. LeCun, L. Bottou, G.B Orr, K. R. Muller: Efficient backprop. In Neural networks: Tricks of the trade (pp. 9-48). Springer Berlin Heidelberg. (2012)

Author Biography

Konstantina Fotiadou *is currently pursuing the PhD degree in Computer Science from the Computer Science Department of the University of Crete. She received her M.Sc. degree in Computer Science from the Computer Science Department of the University of Crete, and B.Sc. degree in Applied Mathematics from the Department of Applied Mathematics, in 2014 and 2011 respectively. Her main research interests involve machine learning techniques for computational imaging applications.*

Grigorios Tsagkatakis *received his B.S. and M.S. degrees in Electronics and Computer Engineering from Technical University of Crete, in 2005 and 2007 respectively. He was awarded his PhD in Imaging Science from the Center for Imaging Science at the Rochester Institute of Technology, USA in 2011. He is currently a postdoctoral fellow at the Institute of Computer Science - FORTH, Greece. His research interests include signal and image processing with applications in sensor networks and imaging systems.*

Panagiotis Tsakalides *received the Diploma degree from Aristotle University of Thessaloniki, Greece, and the Ph.D. degree from the University of Southern California, Los Angeles, USA, in 1990 and 1995, respectively, both in electrical engineering. He is a Professor and the Chairman with the Department of Computer Science, University of Crete, and Head of the Signal Processing Laboratory, Institute of Computer Science, Crete, Greece. He has coauthored over 150 technical publications, including 30 journal papers. He has been the Project Coordinator in seven European Commission and nine national projects. His research interests include statistical signal processing with emphasis in non-Gaussian estimation and detection theory, sparse representations, and applications in sensor networks, audio, imaging, and multimedia systems.*