Neural Cell Segmentation in Large-Scale 3D Color Fluorescence Microscopy Images for Developemental Neuroscience

F. Nourbakhsh, L. Abdeladim, S. Clavreul, K. Loulier, E. Beaurepaire, J. Livet, Anatole Chessel

NEURAL CELL SEGMENTATION IN LARGE-SCALE 3D COLOR FLUORESCENCE MICROSCOPY IMAGES FOR DEVELOPEMENTAL NEUROSCIENCE

F. Nourbakhsh [†], L. Abdeladim [⋆], S. Clavreul[†], K. Loulier[†], E. Beaurepaire[⋆], J. Livet[†] and A. Chessel[⋆]

[†] Sorbonne Université
Institut de la Vision, INSERM, CNRS, Paris

[⋆] Laboratory for optics and biosciences
Ecole polytechnique, INSERM, CNRS, Paris

ABSTRACT

The cells composing brain tissue, neurons, and glia, form extraordinarily complex networks that support cognitive functions. Understanding the organization and development of these networks requires quantitative data resolved at the single cell level. To this aim, we apply novel large-scale 3D multicolor microscopy methodologies in combination with "Brainbow", a transgenic approach enabling to label neural cells with diverse combinations of spectrally distinct fluorescent proteins. In this paper, we present a pipeline based on Convolutional Neural Network (CNN) to detect and segment individual astrocytes, the main type of glial cells of the brain, and map the domains occupied by their fine processes. This bioimage analysis approach successfully handles the challenging variety of astrocyte shape, color, size and their overlap with background elements. Our method shows significant improvement compared with classical techniques, opening the way to varied biological inquiries.

Index Terms— deep learning, segmentation,

1. INTRODUCTION

A major goal in neuroscience is to understand how the complex structure of brain circuits is organized and emerges through development. To this aim, it is essential to access the layout of the cells that participate in these circuits, neurons and glial cells, and to trace back their lineage (i.e. understand how they are generated during development by the division of neural progenitor cells). Optical imaging combined with transgenic cell labeling is one of the main approaches to fulfil this aim. In particular, "Brainbow" multicolor labeling makes it possible to mark neural cells with color labels resulting from combinations of 3-4 spectrally distinct fluorescent proteins (e.g. cyan, yellow and red) [1]. These color labels can be used to 1) image individual neural cells in mature tissue and spatially reconstruct their processes [1, 2]; 2) mark individual neural progenitor cells during development and identify the ensemble of neural cells (clone) that they generate [3]. Applying these approaches over entire brain structures such as the cerebral cortex requires imaging cubic millimeters of intact neural tissue in 3 dimensions at near-micrometric resolution, which can be achieved with recent two-photon multicolor microscopy approaches [4]. The resulting image datasets weight tens to hundreds of Gb and are too large to be conveniently analyzed by human observers. Automated tridimensional segmentation of labeled cellular objects is therefore essential, but poses unique challenges: different types of neural cells are labeled by Brainbow labels; cells of the same type come with a great variability of shapes; and their processes are often juxtaposed or intermingled. Moreover, the 3D trichromatic fluorescence microscopy images are anisotropic (with different planar and axial resolution) and noisy at the single pixel level (precluding simple colorimetric approaches).

In the field of bioimage informatics [5], the detection and segmentation of complex biological objects have received a lot of attention using classical [6] and machine learning based approaches [7]. Su et al. [8] proposed a semi-supervised segmentation clustering method to detect cells in microscopy images based on a dictionary of different patterns and clustering of neighboring pixels. Mualla et al. [9] presented a detection algorithm for unstained cells relying on scale invariant feature transform (SIFT). Massoudi et al. [10] proposed a pixel level segmentation method based on interactive graph cut clustering method. But deep convolutional neural networks (CNNs) are now arguably the state of the art, using GPU and large training sets to train deep neural networks on supervised problems. CNNs show good performance on segmentation applications [11, 12, 13] and have been successfully used in microscopy image analysis. For instance, Ciresan et al. [14] presented a sliding-window system to predict the class label of each pixel on 3D grayscale images. Ronneberger et al. [15] used data augmentation to compensate for lack of training data and improved segmentation by replacing pooling with upsampling operators. Xie et al. [16] implemented an automated cell counter in microscopy images based on fully convolutional regression networks. Chen et al. [17] proposed cascade networks based on [16] to get a probability map of candidates of mitosis and then detect them in details. Xue and Ray [18] presented a CNN-based cell detection method that uses sparsely labeled pixel locations followed by CNN regresses and L1-norm optimization to recover sparse

Author Email: farshad.nourbakhsh@inserm.fr

cell locations. Dong et al. [19] utilized CNNs as a feature selection for Support Vector Machine classifier (SVM) to detect cells in larval zebrafish brain microscopy images. Ciressan et al. [20] used deep max-pooling CNNs to detect mitosis in breast histology images. Similarly Chen et al. [21] applied CNNs to produce probability maps of immune cell locations on different color channels and then detect the center of each cell by non-max suppression. However, the diversity of datasets, acquisition methods and associated biological questions has made the development of generic segmentation algorithms challenging. Moreover, the above approaches are mostly limited to 2D monochromatic (grayscale) images and not adapted to 3D multicolor volumes.

In this paper, we propose a CNN-based segmentation technique for automated cell detection and segmentation in 3D multichannel microscopy image datasets acquired by two-photon microscopy [4] from brain samples labeled with Brainbow color markers [3]. Our CNN trained based on annotated data and complemented by 3D connected component elements distinguishes labeled astrocytes (the main glial cell type of the brain) from other cell types, and enables to position these cells in 3D, segment their cell body and main branches and map the domains occupied by their fine processes. Comparison of our method with techniques commonly used in biology and evaluation of its final segmentation accuracy demonstrate its performance.

2. METHODOLOGY

2.1. Brain Tissue Labeling and Image Acquisition

Astrocytes and neurons of the mouse cerebral cortex were labeled with the MAGIC Markers strategy [3]. Briefly, a Brainbow transgene was electroporated in utero in mouse embryos at 15 days of development. This transgene, capable of integrating in the genome of transfected cells under the action of a transposase, randomly expresses cyan, yellow and red fluorescent proteins upon rearrangement by a co-expressed Cre recombinase. This procedure sparsely distributes varied combinations of the three fluorescent proteins in the cells originating from electroporated cortical progenitors: upper layer excitatory neurons and later-born glial cells, mostly astrocytes [3]. Electroporated animals were allowed to develop until 2 months of age prior imaging. After sacrifice and fixation of brain tissue, 3-dimensional 3-channel images of labeled brains were acquired with wavelength-mixing two-photon fluorescence microscopy [4]. The dataset used in this study corresponds to a $1.2 \times 2 \times 2\ mm^3$ volume of cerebral cortex continuously sampled with $0.4 \times 0.4 \times 1.5$ μm^3 voxel size. Image processing was done using Fiji [22] and Matlab. Illumination inhomogeneities in the field of view were corrected post acquisition, and mosaic stitching was performed using the Grid/Collection Stitching plugin of Fiji [23]. Finally, linear unmixing was applied to correct the limited bleedthrough between channels. Full details concerning the microscopy methodology will be published elsewhere.

2.2. Convolutional Neural Networks (CNNs)

The vast majority of modern CNNs are applied to object classification and use alternating convolution and max-pooling layers followed by a small number of fully connected layers. Pooling layers, in particular, combines the outputs of a neuron of a given layer into a single neuron in the next layer. One of the common used pooling layers approach, max-pooling, selects the maximum value of a cluster of neurons for the next layer. Researchers have made considerable efforts over the last years to improve the performance of this basic pipeline. In particular, Springenberg et al. [24] pointed out that as long as the main contribution of pooling in CNNs is defined as spatial dimensionality reduction, the max-pooling step may be replaced by a strided (larger than one) convolution and fully connected layers may be replaced by simple 1-by-1 convolutions. The resulting overall architecture is thus reduced to convolutional layers with rectified linear non-linearities and an averaging followed by softmax layer for classification.

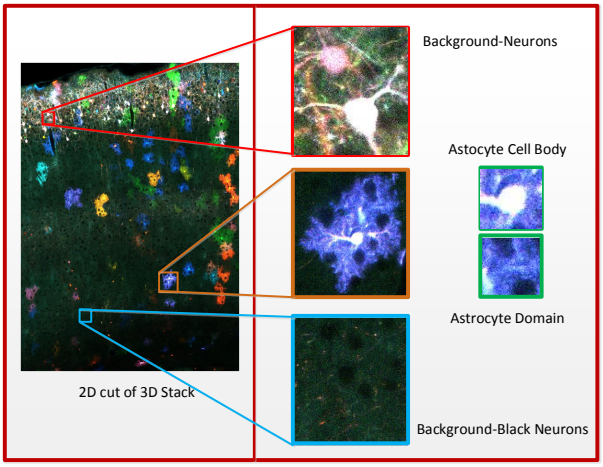

Fig. 1. Example of an image plane from one 3D stack, with the 3 categories of segmented elements.

2.3. Overall Architecture

We adapted the CNN model from [24] to our experiment. This CNN contains sixteen layers with weights; five of the layers are convolutional, three of them are fully connected and the remaining two are strided convolutions designed to reduce layer dimensionality. The output of the last fully-connected layer is replaced by simple 1-by-1 convolutions and fed to a 3-way softmax which produces a distribution over the three classes of labels. The network maximizes the multinomial logistic regression objective. The characteristics of the network

are detailed in Table 1 with the output size of each layer, the type of activation function and the number of parameters. The CNNs method was implemented in Python 3.5 and Matlab. For the deep learning implementation, we used keras and tensorflow. Training computations were conducted on a PC with two processor of 2.10 GHz Intel Core, 32 GB RAM memory and an NVIDIA GeForce GTX 1080 Ti.

	Layer (type)	Output Size	Param #	Description
1	Conv2d-1	(32, 32, 96)	2688	ReLU,strides 1
2	Dropout-1	(32, 32, 96)	0	Dout 0.2, strides 1
3	Conv2d-2	(32, 32, 96)	83040	ReLU, strides 1
4	Conv2d-3	(16, 16, 96)	83040	ReLU, strides 2
5	Dropout-2	(16, 16, 96)	0	Dout 0.5, strides 1
6	Conv2d-4	(16, 16, 192)	166080	ReLU, strides 1
7	Conv2d-5	(16, 16, 192)	331968	ReLU, strides 1
8	Conv2d-6	(8, 8, 192)	331968	ReLU, strides 2
9	Dropout-3	(8,8,192)	0	Dout 0.5, strides 1
10	Conv2d-7	(8, 8, 192)	331968	strides 1
11	activation-1	(8,8,192)	0	ReLU
12	Conv2d-8	(8, 8, 192)	37056	strides 1
13	activation-2	(8,8,192)	0	ReLU
14	Conv2d-9	(8,8,192)	579	strides 1
15	GAV Pooling	(3)	0	
16	activation-3	(3)	0	Softmax

Table 1. Architecture of our CNN network. The total number of parameters is 1,368,387

3. EXPERIMENTAL RESULTS

3.1. Dataset

Analysis of large-volume, high-resolution tridimensional multichannel microscopy images of brain tissue requires efficient procedures for automated segmentation of cellular objects. To explore the potential of CNNs for this purpose, we used a three color dataset encompassing a column of an adult mouse cerebral cortex labeled and imaged as described above. The dataset was partitioned in two sets of image stacks of 4782×2940 pixels, comprising 390 layers. These images contained large numbers of labeled astrocytes, here the object of the experiment, distributed throughout the entire cortex, but also excitatory neurons, marked in upper cortical layers in the present experimental conditions. These two cell types could be identified based on the following characteristics: for astrocytes, star-like morphology of the cell body and main branches, surrounded by dense fine processes that defined a tridimensional domain of brain tissue [25]; for neurons, round cell body and presence of long processes. Despite these characteristic features, individual astrocytes and neurons displayed strong variation in shape, size and color. Moreover, labeled astrocytes were often juxtaposed with one another, and each of them engulfed several labeled or unlabeled neurons. Unlabeled neurons appeared in reversed contrast as black holes within the domain of labeled astrocytes, while the signal from labeled neurons could overlap with that of the astrocytes.

3.2. CNN Training

Our main purpose was to segment 1) the cell body of each astrocyte from the rest of elements in the image and 2) its associated domain. We selected three categories for the segmentation task: astrocyte cell body, astrocyte domain and background (comprising labeled neurons and other unlabeled elements). Within the two image stacks, the cell bodies of all labeled astrocytes (respectively 546 and 534) were manually pointed in 3D using Fiji, as well as unlabeled neural cells engulfed within astrocyte domains. This constituted the ground truth. Data were stored as $32 \times 32 \times 3$ RGB patches. Figure 1 shows one image plane from the dataset and elements from the different classes. As can be seen, astrocyte cell bodies represented a smaller fraction of the dataset compared to astrocyte domains, themselves sparsely distributed within background. From the first image stack (comprising 546 astrocytes), we manually extracted 14,280 RGB patches for background, 8,601 for astrocyte domains, and 4,020 for astrocyte cell bodies. To deal with unbalanced class size, we used data augmentation to bring each class size to 54,000 by random translation of a few pixels. 45,000 RGB patches were randomly selected in each of the 3 categories to train the CNN network and 9,000 for testing. Figure 2 provides an outline of the method. Equivalent AlexNet with maxpooling and fully connected layers was tested on our training and testing data. Our CNN architecture had a better performance, although the difference was minor. In addition, we noticed that the performance increased by adding more conventional steps. We achieved 99.53% accuracy on testing data. We compared the performance of our CNN with the following classical classifiers: Nearest Neighbor Classifier (NN), Sparse Representation Classification (SRC) [26], Robust Sparse Coding (RSC) [27] and CNN with four convolutional layers (CNN4C). Table 2 shows the results of the comparison. Because the training dataset was too large for some of the classifiers, it was reduced by applying K-means clustering and random fashion [28, 29, 30]. Average accuracy is reported based on five iterations. The result indicates that classifiers with random selection were not as robust as those with K-means atoms selection. The NN classifier provided higher accuracy with a lower number of atoms compared to SRC and RSC, but converged faster as the number of atoms increased. SRC and RSC provided slightly higher accuracy than NN with 900 atoms but were too slow to be applied on the entire dataset. Overall, our CNN method provided the highest accuracy compared with NN and CNN4C on the full size dataset. Results of the NN were quantitatively very close but were qualitatively much less satisfying than ours in term of continuity and regularity of the segmented volumes, thanks to the regularizing effect of the translation applied for data augmentation.

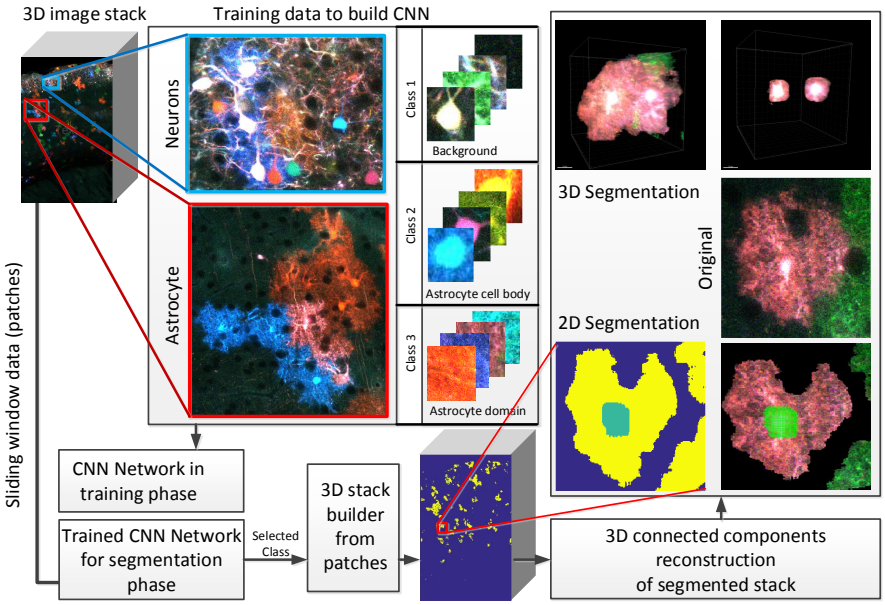

Fig. 2. Schematic presentation of the CNN based method with example of final segmentation

3.3. Detection and Segmentation

The trained CNN model was then applied on the second 3D image stacks for segmentation and detection, using a sliding window approach with $32 \times 32 \times 3$ pixels patches, layer by layer. 3D connected component was then used on the result to individualize astrocytes based on the 3D continuity and size of their cell body, and infer the spatial position of their center. This could then be compared to the ground truth position. Figure 2 shows the different steps of the methodology and an example of segmentation result. Evaluation of the final result with respect to the ground truth on the entire image dataset showed that our procedure correctly detected 95% of astrocyte cell bodies. A possible comparison gaining traction in bioimaging is U-net [15], which directly performs a segmentation on full size images. Our initial tests yielded much lower accuracy, due to false positives created by background neurons, and additional work would be needed to adapt the training set. Moreover the standard U-net method solves a two-class problem while we need at least three classes, with astrocyte domains in addition to cell bodies. Overall this proves that our CNN approach performs satisfyingly and gives results accurate enough for further biological interpretation.

4. CONCLUSION

Automated quantification of 3D multicolor microscopy is challenging but necessary to unlock the potential of this technique in advancing our understanding of living systems. Here we show an application of convolutional network for the detection and segmentation of one neural cell type, astrocytes, in Brainbow-labeled mouse brain. Further work on the algorithm pipeline shall focus on extending it to other cell types as well as speeding up the application of the trained network on large datasets using a multi-resolution strategy, and ensuring its robustness across experiments.

	Partial Training Data		
# Atoms	**3**	**30**	**900**
SRC + Random	34.13±2.22	53.50± 1.81	78.93±1.11
SRC+KMeans	39.20±0.00	61.30±0.32	85.53±0.09
RSC + Random	36.8 ± 1.87	57.21±1.72	79.10 ±1.10
RSC + KMeans	40.38±0.00	62.40±0.23	86.15±0.15
NN + Random	18.30 ±3.30	65.12± 2.40	82.33 ± 2.20
NN + KMeans	61.53± 0.00	81.42± 0.22	84.53± 0.16
# Atoms	**All Data**		
NN	92.25%		
CNN4C	96.61%		
Our-CNN	99.53%		

Table 2. Accuracy comparison of classification methods

5. ACKNOWLEDGEMENTS

Work supported by Fondation pour la Recherche Médicale, ERC-CoG 649117, Fondation ARC pour la Recherche contre le Cancer and Agence Nationale de la Recherche under contracts ANR-11-EQPX-0029, ANR-10-INBS-04 and ANR-10-LABX-65.

6. REFERENCES

[1] J. Livet, TA. Weissman, H. Kang, RW. Draft, J. Lu, RA. Bennis, JR. Sanes, and JW. Lichtman, "Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system," *Nature*, 2007.

[2] S. Hammer, A. Monavarfeshani, T. Su. J. Lemon, and MA. Fox, "Multiple retinal axons converge onto relay cells in the adult mouse thalamus," *Cell Reports*, 2015.

[3] K. Loulier, R. Barry, P. Mahou, Y. Le Franc, W. Supatto, KS. Matho, S. Ieng, S. Fouquet, E. Dupin, R. Benosman, A. Chédotal, E. Beaurepaire, X. Morin, and J. Livet, "Multiplex cell and lineage tracking with combinatorial labels," *Neuron*, 2014.

[4] P. Mahou, M. Zimmerley, K. Loulier, KS. Matho, G. Labroille, X. Morin, W. Supatto, J. Livet, D. Débarre, and E. Beaurepaire, "Multicolor two-photon tissue imaging by wavelength mixing," *Nature Methods*, 2012.

[5] E. Meijering, AE. Carpenter, H. Peng, FA. Hamprecht, and JC. Olivo-Marin, "Imagining the future of bioimage analysis," *Nature Biotechnology*, vol. 34, no. 12, pp. 1250–1255, 2016.

[6] R. C. Gonzalez and R. E. Woods, *Digital Image Processing (3rd Edition)*, Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2006.

[7] Y. Guo, Y. Liu, T. Georgiou, and M. S. Lew, "A review of semantic segmentation using deep neural networks," *International Journal of Multimedia Information Retrieval*, 2017.

[8] H. Su, Z. Yin, S. Huh, and T. Kanade, "Cell segmentation in phase contrast microscopy images via semi-supervised classification over optics-related features," *Medical image analysis*, vol. 17 7, pp. 746–65, 2013.

[9] F. Mualla, S. Schöll, B. Sommerfeldt, A. K. Maier, S. Steidl, R. Buchholz, and J. Hornegger, "Unsupervised unstained cell detection by SIFT keypoint clustering and self-labeling algorithm," in *Medical Image Computing and Computer-Assisted Intervention*, 2014, pp. 377–384.

[10] A. Massoudi, A. Sowmya, K. Mele, and D. Semenovich, "Employing temporal information for cell segmentation using max-flow/min-cut in phase-contrast video microscopy," in *International Conference of the IEEE Engineering in Medicine and Biology Society*, Boston, 2011, pp. 5985–5988.

[11] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," *CoRR*, vol. abs/1411.4038, 2014.

[12] YT. Hu, JB. Huang, and A. G. Schwing, "Maskrnn: Instance level video object segmentation," *CoRR*, vol. abs/1803.11187, 2018.

[13] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network," *CoRR*, vol. abs/1612.01105, 2016.

[14] D. C. Ciresan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, "Deep neural networks segment neuronal membranes in electron microscopy images," in *NIPS*, 2012, pp. 2852–2860.

[15] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *CoRR*, vol. abs/1505.04597, 2015.

[16] W. Xie, J. A. Noble, and A. Zisserman, "Microscopy cell counting and detection with fully convolutional regression networks," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 0, no. 0, pp. 1–10, 2016.

[17] H. Chen, Q. Dou, X. Wang, J. Qin, and P. Heng, "Mitosis detection in breast cancer histology images via deep cascaded networks," 2016.

[18] Y. Xue and N. Ray, "Cell detection with deep convolutional neural network and compressed sensing," *CoRR*, vol. abs/1708.03307, 2017.

[19] B. Dong, L. Shao, M. D. Costa, O. Bandmann, and A. F. Frangi, "Deep learning for automatic cell detection in wide-field microscopy zebrafish images," in *ISBI*. 2015, pp. 772–776, IEEE.

[20] D. C. Ciresan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, "Mitosis detection in breast cancer histology images with deep neural networks," in *MICCAI (2)*. 2013, vol. 8150 of *Lecture Notes in Computer Science*, pp. 411–418, Springer.

[21] T. Chen and C. Chefd'Hotel, "Deep learning based automatic immune cell detection for immunohistochemistry images," in *MLMI*. 2014, vol. 8679 of *Lecture Notes in Computer Science*, pp. 17–24, Springer.

[22] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, JY. Tinevez, DJ. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona, "Fiji: an open-source platform for biological-image analysis," *Nature Methods*, 2012.

[23] S. Preibisch, S. Saalfeld, and P. Tomancak, "Globally optimal stitching of tiled 3d microscopic image acquisitions," *Bioinformatics*, 2009.

[24] J. T. Springenberg, A. Dosovitskiy, T. Brox, and M. A. Riedmiller, "Striving for simplicity: The all convolutional net," *CoRR*, vol. abs/1412.6806, 2014.

[25] ME. Martone EA. Bushong and MH. Ellisman, "Maturation of astrocyte morphology and the establishment of astrocyte domains during postnatal hippocampal development," *International Journal of Developmental Neuroscience*, 2004.

[26] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.

[27] M. Yang, L. Zhang, J. Yang, and D. Zhang, "Robust sparse coding for face recognition," in *Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition*, Washington, DC, USA, 2011, CVPR '11, pp. 625–632, IEEE Computer Society.

[28] F. Nourbakhsh, E. Granger, and G. Fumera, "An extended sparse classification framework for domain adaptation in video surveillance," in *ACCV Workshops (3)*. 2016, vol. 10118 of *Lecture Notes in Computer Science*, pp. 360–376, Springer.

[29] F. Nourbakhsh, *Algorithms for Graph Compression: Theory and Experiments*, Ph.D. thesis, Dipartamento di Scienze Ambientali, Informatica e Statistica, Universitá Ca'Foscari, Venice, Italy, 2015.

[30] F. Nourbakhsh, S. Rota Bulò, and M. Pelillo, "A matrix factorization approach to graph compression with partial information," *Int. J. Machine Learning & Cybernetics*, vol. 6, no. 4, pp. 523–536, 2015.