6

Conclusion

The goal of this thesis has been to propose and validate novel machine and deep learning methods and, in particular, to cope with some of the limitations which usually restrict their range of applicability. This is achieved by integrating statistical modeling in order to inject as much *a priori* information as possible. Indeed, in the last years, several new remote sensing technologies have been designed, developed, and deployed. At the same time, future sensors and missions are already on sight, being planned every day. In contrast with the constant evolution of the quality and typology of the sources of information, the availability of automatic methods able to jointly exploit all the complementary data sources is almost steady.

The ideal solution would be exploiting the huge power that deep learning has shown in many aspects of signal processing. However, the field of remote sensing is cursed by the huge amount of available data together with the scarcity of ground truth data which can be used for training the models. For that reason, the vast majority of supervised method struggles to find its way through the remote sensing field. It is in fact very hard to get a vast and reliable training set by crowdsourcing, as ImageNet was created for object recognition, both due to the generally higher difficulty for non expert to interpret the multisource data they see and to the fact that many datasets are still classified and not publicly available. For that reason, applications based on deep learning methods risk to remain limited to the benchmark datasets and online competitions, fighting for few additional score points without ever seeing the light in a real case scenario. The aim of this thesis is trying to overcome such limitations by integrating statistical modeling and being able to exploit the power of machine and deep learning in the whole procedure which lead from the raw data to the final product.

In this framework, first, in the context of multisensor image registration, an approach capable to exploit the domain adaptation and image-to-image translation capabilities of conditional generative adversarial networks has been proposed for registration purposes. The idea was to to transform a multisensor registration problem into a simpler, more canonical and faster-to-solve single-sensor registration problem. This was done by translating the optical image into the SAR domain, by creating a fake SAR image. The major cost was the necessity of having some coregistered optical and SAR patches to train a cGAN. But, once the model was trained, the generated fake SAR image was in most cases so accurate to be almost impossible to be distinguished from the real SAR image by a human. At this point, the great advantage to reduce the computation cost of the problem was the use of a very simple correlation-type metric to match the real and fake SAR images. Remarkably, such correlation-type metric are usually ineffective in the application to multisensor image registration, but has been made effective within the proposed approach thanks to the integration of the registration method with the powerful image translation capabilities of the cGAN architecture. The experimental results suggested the capability of the method to obtain sub-pixel error and/or visually accurate results, exhibiting a rather small impact of seasonality issues and outperforming a previous area-based approach that used an information-theoretic metric. Moreover, the method proved to be also resilient to the application to different resolutions and to data with different distributions without retraining or even fine tuning.

Second, an approximation of a fully connected conditional random field designed to work with convolutional neural networks has been proposed. When beyond the context of online competitions and benchmark datasets, no detailed pixel-level ground truths are available for training the models. The resulting weakly trained CNN models tend to provide results with much less accuracy and usually cannot even match the accuracies of other, more consolidated, state-of-the-art methods. For that reason, the idea was to integrate as much *a priori* information as possible in the resulting maps. The ideal solution would be a fully connected conditional random field, which is however intractable from the computational viewpoint. Then, the proposed approach integrates intermediate information acquired by the inner layer of the CNN in the form of activations (which are usually discarded after the classification) so to define long range relationships among the pixels. This is done by applying a clustering process to a tensor which includes both original input data and increasingly semantic information coming from the network. Thanks to the additional structure made of clusters, it was possible to deploy an approximation of a fully connected CRF, capable of enforcing inference thought pixels across the image, regardless of their spatial distance. The proposed method has been validated on two well-known subdecimetric semantic segmentation benchmark datasets, in conjunction with two different CNN architectures. The proposed method has proven able to partially fill the gap between the densely and weakly trained models, by retrieving objects that were even completely missed by the CNN, and to outperform even the most recent state-of-the-art methods.

Finally, a large scale multi-sensor data fusion method has been proposed. The method is based on consensus theory, Markov random fields and integrates a Cascade approach and consists of specific formulations and implementations which allow the application of such methods to very large scale scenarios within reasonable computational times. The idea was to reformulating the iterated conditional modes minimization algorithm so that is could be performed by mean of a convolution operation. Thanks to the the interest acquired by such specific operation due to convolutional neural networks, highly optimized strategies for performing convolutions are available. Several stages of validation, starting from the most specific case up to a quantitative validation over an entire subcontinental area, proved the proposed method to be able to greatly improve the results as compared to what could be achieved using single-sensor Sentinel-1 and Sentinel-2 time series. The proposed method is capable of discriminating the correct information from the input data sources though logarithmic opinion pool and to enforce contextual information through MRF. Then, for what regards historical classification, the proposed method is capable of integrating a cascade approach and enforcing temporal coherence into the classification maps and to recover errors caused by scarcity of input images in the past. For what regards the computational complexity, a central aspect when dealing with such large scale applications, the proposed methods proved to be remarkably fast. In fact, thanks to the proposed formulation, the time required to run MRF on a full 10 meters resolution Sentinel-2 tile, which has an extension of 10980×10980 pixels, is of about 8 minutes on a standard desktop machine. In the historical case, the corresponding 30 meters resolution tile has an extent of 3660×3660 pixels and the whole processing including MRF and Cascade requires and average of 30 to 40 seconds only.

The previous paragraphs have been meant to briefly recall the conclusions related to the specific methodologies that have been proposed in this thesis. Nevertheless, it is also worth analysing the possible impact of the proposed methods from a broader perspective. A whole end-to-end deep learning framework would allow to open new possibilities to more and more accurate studies related to several EO crucial aspects, first of all, climate changes. This is in fact receiving primary attention by the remote sensing community and, in particular, by the space agencies. The collection of radar and optical data by Sentinel-1 and Sentinel-2, which is made available by the European Space Agency, together with the Landsat archives by NASA and USGS, and the recent international contests organized by communities like the IEEE Geoscience and Remote Sensing Society, are a clear example of such an interest. Given the huge success that deep learning methodologies are having in many other fields, they are catching more and more attention by the remote sensing community. Indeed, they allow integrating complementary data sources in diverse applications and, due to the heterogeneous nature of their input data, are flexible enough to meet the requirements of the ever growing sets of diverse data that are currently available and that will be available in the future. However, several obstacles usually arise, in particular in relation with the need of training data for such models. This thesis tries to tackle down some of these issues, opening the applicability of deep learning solutions to real case remote sensing applications, both considering the current perspectives and the ones which can be projected in the next future.

Bibliography

[1] P. Krähenbühl and V. Koltun, "Efficient inference in fully connected CRFs with Gaussian edge potentials," in *Proc. Neural Information Processing Systems*, 2011.

[2] J. Le Moigne, N. S. Netanyahu, and R. D. Eastman, Eds., *Image Registration for Remote Sensing*, Cambridge University Press, 2011.

[3] John Alan Richards and JA Richards, *Remote sensing digital image analysis*, vol. 3, Springer, 1999.

[4] Siamak Khorram, Cynthia F Van der Wiele, Frank H Koch, Stacy AC Nelson, and Matthew D Potts, *Principles of applied remote sensing*, Springer, 2016.

[5] Anil K Jain, *Fundamentals of digital image processing*, Prentice-Hall, Inc., 1989.

[6] Luiz Velho, Alejandro C Frery, and Jonas Gomes, *Image processing for computer graphics and vision*, Springer Science & Business Media, 2009.

[7] Charles Elachi and Jakob J Van Zyl, *Introduction to the physics and techniques of remote sensing*, John Wiley & Sons, 2021.

[8] John Lee, *Introduction to topological manifolds*, vol. 202, Springer Science & Business Media, 2010.

[9] Gabriele Moser and Josiane Zerubia, *Mathematical Models for Remote Sensing Image Processing*, Signals and Communications Technology. Springer, Jan. 2018.

[10] R.A. Schowengerdt, *Remote Sensing, Models, and Methods for Image Processing*, Academic Press, 1997.

[11] Leung Tsang, Jin Au Kong, and Kung-Hau Ding, *Scattering of electromagnetic waves: theories and applications*, vol. 27, John Wiley & Sons, 2004.

[12] David A Landgrebe, *Signal theory methods in multispectral remote sensing*, vol. 24, John Wiley & Sons, 2003.

[13] Claudia Kuenzer and Stefan Dech, *Thermal infrared remote sensing: sensors, methods, applications*, vol. 17, Springer Science & Business Media, 2013.

[14] Dimitris G Manolakis, Ronald B Lockwood, and Thomas W Cooley, *Hyperspectral imaging remote sensing: physics, sensors, and algorithms*, Cambridge University Press, 2016.

[15] Luciano Alparone, Bruno Aiazzi, Stefano Baronti, and Andrea Garzelli, *Remote sensing image fusion*, Crc Press, 2015.

[16] Takashi Fujii and Tetsuo Fukuchi, *Laser remote sensing*, CRC press, 2005.

[17] Fawwaz Tayssir Ulaby, David G Long, William J Blackwell, Charles Elachi, Adrian K Fung, Chris Ruf, Kamal Sarabandi, Howard A Zebker, and Jakob Van Zyl, *Microwave radar and radiometric remote sensing*, vol. 4, University of Michigan Press Ann Arbor, 2014.

[18] RA Schowengerdt, "Remote sensing: Models and methods for image processing. ny (usa)," 2006.

[19] Frank Rosenblatt, *The perceptron, a perceiving and recognizing automaton Project Para*, Cornell Aeronautical Laboratory, 1957.

[20] M Minsky, "Perceptrons, expanded edition. an introduction to computational geometry by marvin minsky and seymour a. papert," 1969.

[21] Kunihiko Fukushima, "Neocognitron: A hierarchical neural network capable of visual pattern recognition," *Neural networks*, vol. 1, no. 2, pp. 119–130, 1988.

[22] Christopher M Bishop, "Pattern recognition," *Machine learning*, vol. 128, no. 9, 2006.

[23] Ian Goodfellow, Yoshua Bengio, and Aaron Courville, *Deep learning*, MIT Press, 2017.

[24] Xavier Glorot, Antoine Bordes, and Yoshua Bengio, "Deep sparse rectifier neural networks," in *Proceedings of the fourteenth international conference on artificial intelligence and statistics*. JMLR Workshop and Conference Proceedings, 2011, pp. 315–323.

[25] Geoffrey E Hinton, Nitish Srivastava, Alex Krizhevsky, Ilya Sutskever, and Ruslan R Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors," *arXiv preprint arXiv:1207.0580*, 2012.

[26] Mike Schuster and Kuldip K Paliwal, "Bidirectional recurrent neural networks," *IEEE transactions on Signal Processing*, vol. 45, no. 11, pp. 2673–2681, 1997.

[27] Sepp Hochreiter and Jürgen Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[28] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio, "Generative adversarial nets," *Advances in neural information processing systems*, vol. 27, 2014.

[29] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proc. CVPR*, 2017.

[30] Diederik P Kingma and Max Welling, "Auto-encoding variational bayes," *arXiv preprint arXiv:1312.6114*, 2013.

[31] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov 1998.

[32] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*, USA, 2012, NIPS'12, pp. 1097–1105, Curran Associates Inc.

[33] B. Hariharan, P. Arbeláez, R. Girshick, and J. Malik, "Hypercolumns for object segmentation and fine-grained localization," in *Proc. Computer Vision and Pattern Recognition*. 2015, IEEE/CVF.

[34] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[35] L. A. Gatys, A. S. Ecker, and M. Bethge, "Image style transfer using convolutional neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2414–2423.

[36] Philippe Maillard, "Comparing texture analysis methods through classification," *Photogrammetric Engineering & Remote Sensing*, vol. 69, no. 4, pp. 357–367, 2003.

[37] Pedram Ghamisi, Mauro Dalla Mura, and Jon Atli Benediktsson, "A survey on spectral–spatial classification techniques based on attribute profiles," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 5, pp. 2335–2353, 2014.

[38] Mathieu Fauvel, Yuliya Tarabalka, Jon Atli Benediktsson, Jocelyn Chanussot, and James C Tilton, "Advances in spectral-spatial classification of hyperspectral images," *Proceedings of the IEEE*, vol. 101, no. 3, pp. 652–675, 2012.

[39] D. Koller, N. Friedman, and F. Bach, *Probabilistic graphical models: principles and techniques*, MIT press, 2009.

[40] Z. Kato and J. Zerubia, "Markov random fields in image segmentation," *Foundations and Trends in Machine Learning*, vol. 5, no. 1-2, pp. 1–155, 2012.

[41] Chaohui Wang, Nikos Komodakis, and Nikos Paragios, "Markov random field modeling, inference & learning in computer vision & image understanding: A survey," *Computer Vision and Image Understanding*, vol. 117, no. 11, pp. 1610–1627, 2013.

[42] Haluk Derin and Patrick A Kelly, "Discrete-index markov-type random processes," *Proceedings of the IEEE*, vol. 77, no. 10, pp. 1485–1510, 1989.

[43] A.H.S. Solberg, T. Taxt, and A.K. Jain, "A markov random field model for classification of multisource satellite imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 34, no. 1, pp. 100–113, 1996.

[44] Qiong Jackson and David A Landgrebe, "Adaptive bayesian contextual classification based on markov random fields," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 11, pp. 2454–2463, 2002.

[45] Lisa Gottesfeld Brown and Lisa Gottesfeld, "A survey of image registration techniques," *ACM Computing Surveys*, vol. 24, no. 4, pp. 325–376, Dec. 1992.

[46] N. Merkle, S. Auer, R. Müller, and P. Reinartz, "Exploring the Potential of Conditional Adversarial Networks for Optical and SAR Image Matching," *IEEE J. Sel. Top. Appl. Earth Observ. Rem. Sens.*, vol. 11, pp. 1811–1820, June 2018.

[47] M. J. D. Powell, "A Direct Search Optimization Method That Models the Objective and Constraint Functions by Linear Interpolation," in *Advances in Optimization and Numerical Analysis*, S. Gomez and J.-P. Hennart, Eds., pp. 51–67. Springer Netherlands, Dordrecht, 1994.

[48] Luca Maggiolo, David Solarna, Gabriele Moser, and Sebastiano Bruno Serpico, "Automatic area-based registration of optical and sar images through generative adversarial networks and a correlation-type metric," in *IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium*. Sept. 2020, IEEE.

[49] Béatrice Pinel-Puysségur, Luca Maggiolo, Michel Roux, Nicolas Gasnier, David Solarna, Gabriele Moser, Sebastiano B Serpico, and Florence Tupin, "Experimental comparison of registration methods for multisensor sar-optical data," in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*. IEEE, 2021, pp. 3022–3025.

[50] Jacqueline Le Moigne, "Parallel registration of multisensor remotely sensed imagery using wavelet coefficients," in *Wavelet Applications*. International Society for Optics and Photonics, 1994, vol. 2242, pp. 432–443.

[51] Hui Li, BS Manjunath, and Sanjit K Mitra, "A contour-based approach to multisensor image registration," *IEEE transactions on image processing*, vol. 4, no. 3, pp. 320–334, 1995.

[52] Hui Li, BS Manjunath, and Sanjit K Mitra, "Optical-to-sar image registration using the active contour model," in *Proceedings of 27th*

Asilomar Conference on Signals, Systems and Computers. IEEE, 1993, pp. 568–572.

[53] H-M Chen, Manoj K Arora, and Pramod K Varshney, "Mutual information-based image registration for remote sensing data," *International Journal of Remote Sensing*, vol. 24, no. 18, pp. 3701–3706, 2003.

[54] David G Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.

[55] Bin Fan, Chunlei Huo, Chunhong Pan, and Qingqun Kong, "Registration of optical and sar satellite images by exploring the spatial relationship of the improved sift," *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 4, pp. 657–661, 2012.

[56] Flora Dellinger, Julie Delon, Yann Gousseau, Julien Michel, and Florence Tupin, "Sar-sift: a sift-like algorithm for sar images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 1, pp. 453–466, 2014.

[57] Wenping Ma, Zelian Wen, Yue Wu, Licheng Jiao, Maoguo Gong, Yafei Zheng, and Liang Liu, "Remote sensing image registration with modified sift and enhanced feature matching," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 1, pp. 3–7, 2016.

[58] Yuming Xiang, Feng Wang, and Hongjian You, "Os-sift: A robust sift-like algorithm for high-resolution optical-to-sar image registration in suburban areas," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 6, pp. 3078–3090, 2018.

[59] Jonghye Woo, Maureen Stone, and Jerry L Prince, "Multimodal registration via mutual information incorporating geometric and spatial context," *IEEE Transactions on Image Processing*, vol. 24, no. 2, pp. 757–769, 2014.

[60] Yuanxin Ye, Jie Shan, Lorenzo Bruzzone, and Li Shen, "Robust registration of multimodal remote sensing images based on structural similarity," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 5, pp. 2941–2958, 2017.

[61] Younggi Byun, Jaewan Choi, and Youkyung Han, "An area-based image fusion scheme for the integration of sar and optical satellite imagery," *IEEE journal of selected topics in applied earth observations and remote sensing*, vol. 6, no. 5, pp. 2212–2220, 2013.

[62] Yue Wu, Wenping Ma, Qiguang Miao, and Shanfeng Wang, "Multimodal continuous ant colony optimization for multisensor remote sensing image registration with local search," *Swarm and Evolutionary Computation*, vol. 47, pp. 89–95, 2019.

[63] Lloyd Haydn Hughes, Diego Marcos, Sylvain Lobry, Devis Tuia, and Michael Schmitt, "A deep learning framework for matching of sar and optical imagery," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 169, pp. 166–179, 2020.

[64] Mary Katherine Bennett, *Affine and projective geometry*, John Wiley & Sons, 2011.

[65] Robert B. Ash, *Information Theory*, Courier Corporation, Jan. 1990, Google-Books-ID: yZ1JZA6Wo6YC.

[66] L. Zagorchev and A. Goshtasby, "A comparative study of transformation functions for nonrigid image registration," *IEEE transactions on image processing*, vol. 15, no. 3, 2006.

[67] B. Zitova and J. Flusser, "Image registration methods: a survey," *Image Vision Comput.*, vol. 21, no. 11, pp. 977–1000, Oct. 2003.

[68] David G. Lowe, David G. Lowe, and David G. Lowe, "Object Recognition from Local Scale-Invariant Features," in *Proceedings of the International Conference on Computer Vision-Volume 2 - Volume 2*, Washington, DC, USA, 1999, ICCV '99, pp. 1150–, IEEE Computer Society.

[69] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool, "SURF: Speeded Up Robust Features," in *Computer Vision – ECCV 2006*, Aleš Leonardis, Horst Bischof, and Axel Pinz, Eds., Berlin, Heidelberg, 2006, Lecture Notes in Computer Science, pp. 404–417, Springer.

[70] M. Donoser and H. Bischof, "Efficient Maximally Stable Extremal Region (MSER) Tracking," in *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, June 2006, vol. 1, pp. 553–560, ISSN: 1063-6919.

[71] Chris Harris and Mike Stephens, “A combined corner and edge detector,” in *In Proc. of Fourth Alvey Vision Conference*, 1988, pp. 147–151.

[72] Richard O. Duda and Peter E. Hart, “Use of the Hough transformation to detect lines and curves in pictures,” *Communications of the ACM*, vol. 15, no. 1, pp. 11–15, Jan. 1972.

[73] Xavier Descombes, Robert Minlos, and Elena Zhizhina, “Object Extraction Using a Stochastic Birth-and-Death Dynamics in Continuum,” *Journal of Mathematical Imaging and Vision*, vol. 33, no. 3, pp. 347–359, Mar. 2009.

[74] Chunlei Huo, Keming Chen, Zhixin Zhou, and Hanqing Lu, “Hybrid approach for remote sensing image registration,” in *MIPPR 2007: Remote Sensing and GIS Data Processing and Applications; and Innovative Multispectral Technology and Applications*. International Society for Optics and Photonics, 2007, vol. 6790, p. 679006.

[75] David Solarna, Alberto Gotelli, Jacqueline Le Moigne, Gabriele Moser, and Sebastiano B Serpico, “Crater detection and registration of planetary images through marked point processes, multiscale decomposition, and region-based analysis,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 9, pp. 6039–6058, 2020.

[76] David Solarna, Gabriele Moser, Jacqueline Le Moigne, and Sebastiano B Serpico, “Planetary crater detection and registration using marked point processes, multiple birth and death algorithms, and region-based analysis,” in *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. IEEE, 2017, pp. 2337–2340.

[77] Xiaolei Huang, Yiyong Sun, Dimitris Metaxas, Frank Sauer, and Chenyang Xu, “Hybrid image registration based on configural matching of scale-invariant salient region features,” in *2004 Conference on Computer Vision and Pattern Recognition Workshop*. IEEE, 2004, pp. 167–167.

[78] Maoguo Gong, Shengmeng Zhao, Licheng Jiao, Dayong Tian, and Shuang Wang, “A novel coarse-to-fine scheme for automatic image registration based on sift and mutual information,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 7, pp. 4328–4338, 2013.

[79] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Transactions on Medical Imaging*, vol. 16, no. 2, pp. 187–198, 1997.

[80] P. Thevenaz and M. Unser, "Optimization of mutual information for multiresolution image registration," *IEEE Transactions on Image Processing*, vol. 9, no. 12, pp. 2083–2099, 2000.

[81] Dirk Loeckx, Pieter Slagmolen, Frederik Maes, Dirk Vandermeulen, and Paul Suetens, "Nonrigid image registration using conditional mutual information," *IEEE Transactions on Medical Imaging*, vol. 29, no. 1, pp. 19–29, 2010.

[82] Emanuel Parzen, "On Estimation of a Probability Density Function and Mode," *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 1065–1076, 1962.

[83] Weihao Gao, Sewoong Oh, and Pramod Viswanath, "Demystifying fixed k-nearest neighbor information estimators," *IEEE Transactions on Information Theory*, vol. 64, no. 8, pp. 5629–5661, 2018.

[84] Dimitre H. Hristov and B. Gino Fallone, "A grey-level image alignment algorithm for registration of portal images and digitally reconstructed radiographs.," *Medical physics*, vol. 23 1, pp. 75–84, 1996.

[85] J.N. Sarvaiya, Suprava Patnaik, and Salman Bombaywala, "Image registration by template matching using normalized cross-correlation," in *2009 International Conference on Advances in Computing, Control, and Telecommunication Technologies*, 2009, pp. 819–822.

[86] Melanie (Computer scientist) Mitchell and Melanie, *An introduction to genetic algorithms*, MIT Press, 1996.

[87] Peter J. M. van Laarhoven and Emile H. L. Aarts, *Simulated annealing*, pp. 7–15, Springer Netherlands, Dordrecht, 1987.

[88] Michael Gadermayr, Lotte Heckmann, Kexin Li, Friederike Bähr, Madlaine Müller, Daniel Truhn, Dorit Merhof, and Burkhard Gess, "Image-to-image translation for simplified mri muscle segmentation," *Frontiers in Radiology*, vol. 1, pp. 3, 2021.

[89] Richard Zhang, Phillip Isola, and Alexei A Efros, "Colorful image colorization," in *European conference on computer vision*. Springer, 2016, pp. 649–666.

[90] Mario Beaulieu, Samuel Foucher, Dan Haberman, and Colin Stewart, "Deep image-to-image transfer applied to resolution enhancement of sentinel-2 images," in *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2018, pp. 2611–2614.

[91] Y. Jing, Y. Yang, Z. Feng, J. Ye, Y. Yu, and M. Song, "Neural style transfer: A review," *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 11, pp. 3365–3385, 2020.

[92] Leon A. Gatys, Alexander S. Ecker, and Matthias Bethge, "A neural algorithm of artistic style," *CoRR*, vol. abs/1508.06576, 2015.

[93] Alireza Makhzani, Jonathon Shlens, Navdeep Jaitly, Ian Goodfellow, and Brendan Frey, "Adversarial Autoencoders," Nov. 2015.

[94] Rafael C Gonzalez, Richard E Woods, and Barry R Masters, "Digital image processing," 2009.

[95] David Solarna, Luca Maggiolo, Gabriele Moser, and Sebastiano Bruno Serpico, "A tiling-based strategy for large-scale multisensor image registration.," in *Submitted to: IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium*. 2022, IEEE.

[96] M. J. D. Powell, *A Direct Search Optimization Method That Models the Objective and Constraint Functions by Linear Interpolation*, pp. 51–67, Springer Netherlands, Dordrecht, 1994.

[97] W. Zhao, L. Denis, C. Deledalle, H. Maitre, J.-M. Nicolas, and F. Tupin, "Ratio-based multi-temporal SAR images denoising," *IEEE Trans. Geosci. Rem. Sens.*, vol. 57, pp. 3552–3565, 2018.

[98] S. Parrilli et al., "A nonlocal sar image denoising algorithm based on llmmse wavelet shrinkage," *IEEE Trans Geosci Remote Sens*, vol. 50, no. 2, pp. 606–616, 2012.

[99] M. J. D. Powell, "An efficient method for finding the minimum of a function of several variables without calculating derivatives," *The Computer Journal*, vol. 7, no. 2, pp. 155–162, Jan. 1964.

[100] R. P. Brent, "An algorithm with guaranteed convergence for finding a zero of a function," *The Computer Journal*, vol. 14, no. 4, pp. 422–425, Jan. 1971.

[101] Yurii Nesterov et al., *Lectures on convex optimization*, vol. 137, Springer, 2018.

[102] I. Zavorin and J. Le Moigne, "Use of multiresolution wavelet feature pyramids for automatic registration of multisensor imagery," *IEEE Transactions on Image Processing*, vol. 14, pp. 770–782, 2005.

[103] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2223–2232.

[104] X. Zhu, D. Tuia, L. Mou, G. Xia, L. Zhang, F. Xu, and F. Fraundorfer, "Deep learning in remote sensing: A comprehensive review and list of resources," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 4, pp. 8–36, 2017.

[105] M. Volpi and D. Tuia, "Dense semantic labeling of subdecimeter resolution images with convolutional neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 2, pp. 881–893, 2017.

[106] E. Maggiori, Y. Tarabalka, G. Charpiat, and P. Alliez, "Convolutional neural networks for large-scale remote-sensing image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 2, pp. 645–657, 2017.

[107] X. Zhu, "Semi-supervised learning literature survey," *Comput. Sci., University of Wisconsin-Madison*, vol. 2, 07 2008.

[108] O. Chapelle, B. Scholkopf, and A. Zien, "Semi-supervised learning," *Cambridge: MIT Press*, 2006.

[109] T. Shen, G. Lin, L. Liu, C. Shen, and I. Reid, "Weakly supervised semantic segmentation based on co-segmentation," *CoRR*, vol. abs/1705.09052, 2017.

[110] A. Bearman, O. Russakovsky, V. Ferrari, and L. Fei-Fei, "What's the point: Semantic segmentation with point supervision," in *Proc.*

European Conference on Computer Vision. Springer, 2016, pp. 549–565.

[111] D. Lin, J. Dai, J. Jia, K. He, and J. Sun, "Scribblesup: Scribble-supervised convolutional networks for semantic segmentation," in *Proc. Computer Vision and Pattern Recognition*. 2016, pp. 3159–3167, IEEE/CVF.

[112] L. Maggiolo, D. Marcos, G. Moser, and D. Tuia, "Improving maps from cnns trained with sparse, scribbled ground truths using fully connected crfs," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Valencia, Spain, Jul 2018, pp. 2099–2102.

[113] L. Maggiolo, D. Marcos, G. Moser, S. B. Serpico, and D. Tuia, "A semisupervised crf model for cnn-based semantic segmentation with sparse ground truth," *IEEE Transactions on Geoscience and Remote Sensing*, 2021.

[114] A. Krizhevsky, I. Sutskever, and G. E Hinton, "Imagenet classification with deep convolutional neural networks," in *Adv. Neural. Inf. Process. Syst.*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds., pp. 1097–1105. Curran Associates, Inc., 2012.

[115] M. Campos-Taberner, A. Romero-Soriano, C. Gatta, G. Camps-Valls, A. Lagrange, B. Le Saux, A. Beaupère, A. Boulch, A. Chan-Hon-Tong, S. Herbin, H. Randrianarivo, M. Ferecatu, M. Shimoni, G. Moser, and D. Tuia, "Processing of extremely high-resolution LiDAR and RGB data: Outcome of the 2015 IEEE GRSS data fusion contest–part a: 2-d contest," *IEEE J. Sel. Top. Appl. Earth. Obs. Remote. Sens.*, vol. 9, no. 12, pp. 5547–5559, Dec 2016.

[116] K. Makantasis, K. Karantzalos, A. Doulamis, and N. Doulamis, "Deep supervised learning for hyperspectral data classification through convolutional neural networks," in *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2015, pp. 4959–4962.

[117] A. Lagrange, B. Le Saux, A. Beaupère, A. Boulch, A. Chan-Hon-Tong, S. Herbin, H. Randrianarivo, and M. Ferecatu, "Benchmarking classification of earth-observation data: From learning explicit features to convolutional networks," in *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2015, pp. 4173–4176.

[118] F. P. S. Luus, B. P. Salmon, F. van den Bergh, and B. T. J. Maharaj, "Multiview deep learning for land-use classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 12, pp. 2448–2452, 2015.

[119] J. Sherrah, "Fully convolutional networks for dense semantic labelling of high-resolution aerial imagery," *CoRR*, vol. abs/1606.02585, 2016.

[120] N. Audebert, B. Le Saux, and S. Lefèvre, "Semantic segmentation of earth observation data using multimodal and multi-scale deep networks," in *Computer Vision – ACCV 2016*, SH. Lai, V. Lepetit, K. Nishino, and Y. Sato, Eds., Cham, 2017, pp. 180–196, Springer International Publishing.

[121] D. Marmanis, K. Schindler, J.D. Wegner, S. Galliani, M. Datcu, and U. Stilla, "Classification with an edge: Improving semantic image segmentation with boundary detection," *ISPRS J. Photo. Remote Sens.*, vol. 135, pp. 158–172, 2018.

[122] E. Maggiori, Y. Tarabalka, G. Charpiat, and P. Alliez, "High-resolution semantic labeling with convolutional neural networks," *CoRR*, vol. abs/1611.01962, 2016.

[123] D. Marcos, M. Volpi, B. Kellenberger, and D. Tuia, "Land cover mapping at very high resolution with rotation equivariant CNNs: towards small yet accurate models," *ISPRS J. Int. Soc. Photo. Remote Sens.*, in press.

[124] Z. Zhou, "A brief introduction to weakly supervised learning," *National science review*, vol. 5, no. 1, pp. 44–53, 2018.

[125] A. Nivaggioli and H. Randrianarivo, "Weakly supervised semantic segmentation of satellite images," in *2019 Joint Urban Remote Sensing Event (JURSE)*. IEEE, 2019, pp. 1–4.

[126] W. Wu, H. Qi, Z. Rong, L. Liu, and H. Su, "Scribble-supervised segmentation of aerial building footprints using adversarial learning," *IEEE Access*, vol. 6, pp. 58898–58911, 2018.

[127] Y. Hua, D. Marcos, L. Mou, X. Xiang Zhu, and D. Tuia, "Semantic segmentation of remote sensing images with sparse annotations," *CoRR*, vol. abs/2101.03492, 2021.

[128] J. Wang, C. H. Q. Ding, S. Chen, C. He, and B. Luo, "Semi-supervised remote sensing image semantic segmentation via consistency regularization and average update of pseudo-label," *Remote Sensing*, vol. 12, no. 21, 2020.

[129] X. Sun, A. Shi, H. Huang, and H. Mayer, "Bas^{4}net: Boundary-aware semi-supervised semantic segmentation network for very high resolution remote sensing images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 5398–5413, 2020.

[130] K. Schindler, "An overview and comparison of smooth labeling methods for land-cover classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 11, pp. 4534–4545, 2012.

[131] C. Sutton, A. McCallum, and F. Pereira, ""An introduction to conditional random fields," *Found. Trends Mach. Learn.*, vol. 4, no. 4, pp. 267–373, 2011.

[132] X. He, R. S. Zemel, and M. A. Carreira-Perpinan, "Multiscale conditional random fields for image labeling.," in *Proc. Computer Vision and Pattern Recognition*. 2004, IEEE/CVF.

[133] S. Kumar and H. Hebert, "A hierarchical field framework for unified context-based classification," in *Proc. International Conference on Compute Vision*, 2005.

[134] P. Kohli, L. Ladicky, and P. H. S. Torr, "Robust higher order potentials for enforcing label consistency.," *Int. J. Comput. Vision*, 2009.

[135] L. Ladicky, C. Russell, P. Kohli, and P. H. S. Torr, "Associative hierarchical CRFs for object class image segmentation.," in *Proc. International Conference on Compute Vision*, 2009.

[136] L. Ladicky, C. Russell, P. Kohli, and P. H. S. Torr, "Graph cut based inference with co-occurrence statistics.," in *Proc. European Conference on Computer Vision*, 2010.

[137] D. Tuia, M. Volpi, and G. Moser, "Decision fusion with multiple spatial supports by conditional random fields," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 6, pp. 3277–3289, 2018.

[138] M. Volpi and D. Tuia, "Deep multi-task learning for a geographically-regularized semantic segmentation of aerial images," *CoRR*, vol. abs/1808.07675, 2018.

[139] A. Rabinovich, A. Vedaldi, C. Galleguillos, E. Wiewiora, and S. Belongie, "Objects in context.," in *Proc. International Conference on Compute Vision*, Oct 2007, pp. 1–8.

[140] T. Toyoda and O. Hasegawa, "Random field model for integration of local information and global information," *IEEE Trans. Pattern Anal. Machine Intell.*, 2008.

[141] C. Galleguillos, A. Rabinovich, and S. Belongie, "Object categorization using co-occurrence, location and appearance," in *Proc. Computer Vision and Pattern Recognition*. 2008, IEEE/CVF.

[142] N. Payet and S. Todorovic, "RF^2- random forest random field.,," *Proc. Neural Information Processing Systems*, 2010.

[143] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-6, no. 6, pp. 721–741, 1984.

[144] S. Z. Li, *Markov Random Field Modeling in Image Analysis*, Springer Publishing Company, Incorporated, 3rd edition, 2009.

[145] I. Hedhli, G. Moser, J. Zerubia, and S. B. Serpico, "A new cascade model for the hierarchical joint classification of multitemporal and multiresolution remote sensing data," *IEEE Trans. Geosci. Rem. Sens.*, vol. 54, no. 11, pp. 6333–6348, 2016.

[146] Y. Boykov, O. Veksler, and R. Zabih, "Efficient approximate energy minimization via graph cuts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 12, pp. 1222–1239, 2001.

[147] B. Fulkerson, A. Vedaldi, and S. Soatto, "Class segmentation and object localization with superpixel neighborhoods," in *Proc. Computer Vision and Pattern Recognition*. 2009, IEEE/CVF.

[148] V. Kolmogorov and R. Zabih, "What energy functions can be minimized via graph cuts?," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 2, pp. 147–159, February 2004.

[149] Y. Boykov and V. Kolmogorov, “An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision.,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 9, pp. 1124–1137, 2004.

[150] I. Demir, K. Koperski, D. Lindenbaum, G. Pang, J. Huang, S. Basu, F. Hughes, D. Tuia, and R. Raskar, “Deepglobe 2018: A challenge to parse the earth through satellite images,” *CoRR*, vol. abs/1805.06561, 2018.

[151] P. Soille, *Morphological Image Analysis: Principles and Applications*, Springer-Verlag, Berlin, Heidelberg, 2 edition, 2003.

[152] Z. Wang and S. Ji, “Smoothed dilated convolutions for improved dense prediction,” *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Jul 2018.

[153] Luis Gómez-Chova, Devis Tuia, Gabriele Moser, and Gustau Camps-Valls, “Multimodal classification of remote sensing images: A review and future directions,” *Proceedings of the IEEE*, vol. 103, no. 9, pp. 1560–1584, 2015.

[154] Jon A Benediktsson, Gabriele Cavallaro, Nicola Falco, Ihsen Hedhli, Vladimir A Krylov, Gabriele Moser, Sebastiano B Serpico, and Josiane Zerubia, “Remote sensing data fusion: Markov models and mathematical morphology for multisensor, multiresolution, and multiscale image classification,” in *Mathematical Models for Remote Sensing Image Processing*, pp. 277–323. Springer, 2018.

[155] Andreas Merentitis and Christian Debes, “Many hands make light work-on ensemble learning techniques for data fusion in remote sensing,” *IEEE Geoscience and Remote Sensing Magazine*, vol. 3, no. 3, pp. 86–99, 2015.

[156] Jon Atli Benediktsson and Ioannis Kanellopoulos, “Classification of multisource and hyperspectral data based on decision fusion,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 3, pp. 1367–1377, 1999.

[157] Behnaz Bigdeli, Farhad Samadzadegan, and Peter Reinartz, “A decision fusion method based on multiple support vector machine system for fusion of hyperspectral and lidar data,” *International Journal of Image and Data Fusion*, vol. 5, no. 3, pp. 196–209, 2014.

[158] Xavier Ceamanos, Björn Waske, Jon Atli Benediktsson, Jocelyn Chanussot, Mathieu Fauvel, and Johannes R Sveinsson, "A classifier ensemble based on fusion of support vector machines for classifying hyperspectral data," *International Journal of Image and Data Fusion*, vol. 1, no. 4, pp. 293–307, 2010.

[159] Moumen T El-Melegy and Safaa M Ahmed, "Neural networks in multiple classifier systems for remote-sensing image classification," in *Soft Computing in Image Processing*, pp. 65–94. Springer, 2007.

[160] Mathieu Fauvel, Jocelyn Chanussot, and Jón Atli Benediktsson, "Decision fusion for the classification of urban remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 10, pp. 2828–2838, 2006.

[161] Karoly Livius Bakos and Paolo Gamba, "Hierarchical hybrid decision tree fusion of multiple hyperspectral data processing chains," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 1, pp. 388–394, 2010.

[162] Andrey V Bogdanov, "Neuroinspired architecture for robust classifier fusion of multisensor imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 5, pp. 1467–1487, 2008.

[163] J.A. Benediktsson and P.H. Swain, "Consensus theoretic classification methods," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 22, no. 4, pp. 688–704, 1992.

[164] Jon Atli Benediktsson, "Hybrid consensus theoretic classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, no. 4, pp. 833–843, July 1997.

[165] Norman Crolee Dalkey, "An Impossibility Theorem for Group Probability Functions.," 1972.

[166] J.A. Benediktsson and I. Kanellopoulos, "Classification of multisource and hyperspectral data based on decision fusion," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 3, pp. 1367–1377, May 1999.

[167] G. Storvik, R. Fjortoft, and A.H.S. Solberg, "A bayesian approach to classification of multiresolution remote sensing data," *IEEE Transac-*

tions on Geoscience and Remote Sensing, vol. 43, no. 3, pp. 539–547, Mar. 2005.

[168] R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother, "A Comparative Study of Energy Minimization Methods for Markov Random Fields with Smoothness-Based Priors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 6, pp. 1068–1080, June 2008.

[169] "CCI+ - High Resolution Land cover," `https://climate.esa.int/en/projects/high-resolution-land-cover/`, [].

[170] Rainer Hollmann, Chris J Merchant, Roger Saunders, Catherine Downy, Michael Buchwitz, Anny Cazenave, Emilio Chuvieco, Pierre Defourny, Gerrit de Leeuw, René Forsberg, et al., "The esa climate change initiative: Satellite data records for essential climate variables," *Bulletin of the American Meteorological Society*, vol. 94, no. 10, pp. 1541–1552, 2013.

[171] Philip H Swain, "Bayesian classification in a time-varying environment," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 8, no. 12, pp. 879–883, 1978.

[172] Ben Poulter, Natasha MacBean, Andrew Hartley, Iryna Khlystova, Olivier Arino, Richard Betts, Sophie Bontemps, Martin Boettcher, Carsten Brockmann, Pierre Defourny, et al., "Plant functional type classification for earth system models: results from the european space agency's land cover climate change initiative," *Geoscientific Model Development*, vol. 8, no. 7, pp. 2315–2328, 2015.

[173] Wei Li, Philippe Ciais, Natasha MacBean, Shushi Peng, Pierre Defourny, and Sophie Bontemps, "Major forest changes and land cover transitions based on plant functional types derived from the esa cci land cover product," *International journal of applied earth observation and geoinformation*, vol. 47, pp. 30–39, 2016.