Structured Learning and Inference in Computer Vision

Matthew B. Blaschko
http://www.robots.ox.ac.uk/~blaschko/

Visual Geometry Group,
University of Oxford

Learning methods have been widely applied in computer vision to learn how to solve tasks such as image classification. While this is highly simplified from the original goal of enabling computers to process visual data with similar sophistication to humans, the high variability of visual data and intractability of complex models has slowed attempts to more directly solve vision problems. This course will cover an exciting new development in machine learning that has enabled systems to better model the visual world: structured output learning. Kernel methods are most commonly known through the popular support vector machine (SVM) classification algorithm, which is capable of predicting binary outputs. One view of structured output learning is as a generalisation of the SVM algorithm to the direct prediction of more complex and interdependent output spaces. Many problems in computer vision are more appropriately framed as the prediction of an interdependent structured representation, such as an image segmentation, or scene layout. Though these learning tasks may seem disparate, we will see that through an appropriate task representation the same underlying learning algorithms can be employed. The approaches presented here are unified by a common set of assumptions that ensure tractability for a large number of problems, while ensuring a sufficiently rich representation for non-trivial learning, aided by the use of kernels. Examples primarily focus on object detection and scene analysis, with a discussion of graphical models, kernel design, and inference strategies [1, 2, 3, 4, 5, 6].

[1] M.B. Blaschko and C. H. Lampert. Object localization with global and local context kernels. In *Proceedings of the British Machine Vision Conference (BMVC)*, 2009.

[2] M.B. Blaschko, A. Vedaldi, and A. Zisserman. Simultaneous object detection and ranking with weak supervision. In *Proceedings of the Twenty-Fourth Annual Conference on Neural Information Processing Systems (NIPS)*, 2010.

[3] C. H. Lampert and M. B. Blaschko. Structured prediction by joint kernel support estimation. *Machine Learning*, 77(2-3):249–269, 2009.

[4] E. Rahtu, J. Kannala, and M.B. Blaschko. Learning a category independent object detection cascade. In *International Conference on Computer Vision (ICCV)*, 2011.

[5] T. Tuytelaars, C. H. Lampert, M. B. Blaschko, and W. Buntine. Unsupervised object discovery: A comparison. *International Journal of Computer Vision (IJCV)*, 88(2):284–302, 2010.

[6] A. Vedaldi, M. B. Blaschko, and A. Zisserman. Learning equivariant structured output svm regressors. In *International Conference on Computer Vision (ICCV)*, 2011.

Matthew Blaschko Biography: Matthew B. Blaschko holds a B.S. from Columbia University and a M.S. from the University of Massachusetts Amherst. He was awarded a doctorate in Computer Science, summa cum laude, from the Technische Universität Berlin for work done at the Max Planck Institute for Biological Cybernetics in Tübingen, Germany. He is currently a Newton International Fellow in the Department of Engineering Science at the University of Oxford, where he works with Professor Andrew Zisserman. He has been the recipient of the Main Prize of the German Association for Pattern Recognition, as well as best paper awards at CVPR and ECCV. His research interests are in machine learning techniques applied to visual data.