

Acknowledgement

We thank the anonymous reviewers for valuable comments and suggestions. This work is partly supported by the US NSF Grant IIS-1218156 and the US NSF CAREER Award IIS-1350521.

References

- Argyriou, A.; Evgeniou, T.; and Pontil, M. 2008. Convex multi-task feature learning. *Machine Learning* 73(3):243–272.
- Arica, N., and Yarman-Vural, F. 2001. An overview of character recognition focused on off-line handwriting. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on* 31(2):216–233.
- Caruana, R. 1997. Multitask learning. *Machine Learning* 28(1):41–75.
- Dai, W.; Yang, Q.; Xue, G.-R.; and Yu, Y. 2007. Boosting for transfer learning. In *Proceedings of the 24th International Conference on Machine Learning, ICML '07*, 193–200. New York, NY, USA: ACM.
- Freund, Y., and Schapire, R. E. 1997. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences* 55(1):119 – 139.
- Friedman, J.; Hastie, T.; and Tibshirani, R. 2000. Additive logistic regression: a statistical view of boosting. *Annals of Statistics* 28(2):337–407.
- Frigyik, B.; Srivastava, S.; and Gupta., M. 2008. An introduction to functional derivatives. *Technical Report (University of Washington)*.
- Haralick, R.; Shanmugam, K.; and Dinstein, I. 1973. Textural features for image classification. *Systems, Man and Cybernetics, IEEE Transactions on* SMC-3(6):610–621.
- Hastie, T., and Tibshirani, R. 1990. Generalized additive model. *Chapman Hall, London*.
- Hwang, S. J.; Grauman, K.; and Sha, F. 2011. Learning a tree of metrics with disjoint visual features. In *Shawe-Taylor, J.; Zemel, R. S.; Bartlett, P. L.; Pereira, F. C. N.; and Weinberger, K. Q., eds., NIPS*, 621–629.
- Hwang, S. J.; Sha, F.; and Grauman, K. 2011. Sharing features between objects and their attributes. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, 1761–1768.
- Lyons, M.; Akamatsu, S.; Kamachi, M.; and Gyoba, J. 1998. Coding facial expressions with gabor wavelets. In *Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on*, 200–205.
- Ojala, T.; Pietikainen, M.; and Maenpaa, T. 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 24(7):971–987.
- Ojansivu, V., and Heikkil, J. 2008. A method for blur and affine invariant object recognition using phase-only bispectrum. In *Campilho, A., and Kamel, M., eds., Image Analysis and Recognition*, volume 5112 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg. 527–536.
- Pan, S. J., and Yang, Q. 2010. A survey on transfer learning. *Knowledge and Data Engineering, IEEE Transactions on* 22(10):1345–1359.
- Romera-Paredes, B.; Argyriou, A.; Berthouze, N.; and Pontil, M. 2012. Exploiting unrelated tasks in multi-task learning. In *Lawrence, N. D., and Girolami, M., eds., AISTATS*, volume 22 of *JMLR Proceedings*, 951–959. JMLR.org.
- Schomaker, L. 2007. Advances in writer identification and verification. In *Document Analysis and Recognition, 2007. ICDAR 2007. Ninth International Conference on*, volume 2, 1268–1273.
- Shalev-Shwartz, S.; Wexler, Y.; and Shashua, A. 2011. Shareboost: Efficient multiclass learning with feature sharing. In *Shawe-Taylor, J.; Zemel, R. S.; Bartlett, P. L.; Pereira, F. C. N.; and Weinberger, K. Q., eds., NIPS*, 1179–1187.
- Torralba, A.; Murphy, K.; and Freeman, W. 2007. Sharing visual features for multiclass and multiview object detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 29(5):854–869.
- Viola, P., and Jones, M. 2004. Robust real-time face detection. *International Journal of Computer Vision* 57(2):137–154.
- Wang, X.; Zhang, C.; and Zhang, Z. 2009. Boosted multi-task learning for face verification with applications to web image and video search. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, 142–149.
- Xiao, L.; Zhou, D.; and Wu, M. 2011. Hierarchical classification via orthogonal transfer. In *Getoor, L., and Scheffer, T., eds., ICML*, 801–808. Omnipress.
- Yao, Y., and Doretto, G. 2010. Boosting for transfer learning with multiple sources. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, 1855–1862.
- Zeng, Z.; Pantic, M.; Roisman, G.; and Huang, T. 2009. A survey of affect recognition methods: Audio, visual, and spontaneous expressions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 31(1):39–58.
- Zhang, X.-Y., and Liu, C.-L. 2013. Writer adaptation with style transfer mapping. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 35(7):1773–1787.
- Zhao, W.; Chellappa, R.; Phillips, P. J.; and Rosenfeld, A. 2003. Face recognition: A literature survey. *ACM Comput. Surv.* 35(4):399–458.
- Zhou, Y.; Jin, R.; and Hoi, S. C. H. 2010. Exclusive lasso for multi-task feature selection. In *Teh, Y. W., and Titterton, D. M., eds., AISTATS*, volume 9 of *JMLR Proceedings*, 988–995. JMLR.org.