[5] A. Krizhevsky, I. Sutskever, and G. E. Hinton, ‘‘ImageNet classification

with deep convolutional neural networks,’’ in Proc. Adv. Neural Inf. Pro-

cess. Syst., 2012, pp. 1097–1105.

[13] M. E. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee,

and L. Zettlemoyer, ‘‘Deep contextualized word representations,’’ 2018,

arXiv:1802.05365. [Online]. Available: https://arxiv.org/abs/1802.05365

[14] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever. (2018). Improv-

ing Language Understanding by Generative Pre-Training. [Online].

Available: https://s3-us-west-2.amazonaws.com/openai-assets/research-

covers/language-unsupervised/language\_understanding\_paper.pdf

[15] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, ‘‘BERT: Pre-training

of deep bidirectional transformers for language understanding,’’ 2018,

arXiv:1810.04805. [Online]. Available: https://arxiv.org/abs/1810.04805

[16] Z. Yang, Z. Dai, Y. Yang, J. Carbonell, R. Salakhutdinov, and Q. V. Le,

‘‘XLNet: Generalized autoregressive pretraining for language under-

standing,’’ 2019, arXiv:1906.08237. [Online]. Available: https://arxiv.

org/abs/1906.08237

[18] S. J. Pan, X. Ni, J.-T. Sun, Q. Yang, and Z. Chen, ‘‘Cross-domain sentiment

classification via spectral feature alignment,’’ in Proc. 19th Int. Conf.

WWW, 2010, pp. 751–760.

[19] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette,

M. Marchand, and V. Lempitsky, ‘‘Domain-adversarial training of neural

networks,’’ J. Mach. Learn. Res., vol. 17, no. 59, pp. 1–35, 2016.

[20] B. Sun, J. Feng, and K. Saenko, ‘‘Return of frustratingly easy domain

adaptation,’’ in Proc. 30th AAAI, 2016, pp. 1–8.

[21] W. Zellinger, T. Grubinger, E. Lughofer, T. Natschlger, and

S. Saminger-Platz, ‘‘Central moment discrepancy (CMD) for domain-

invariant representation learning,’’ 2017, arXiv:1702.08811. [Online].

Available: https://arxiv.org/abs/1702.08811

[22] C. Chen, Z. Chen, B. Jiang, and X. Jin, ‘‘Joint domain alignment and

discriminative feature learning for unsupervised deep domain adaptation,’’

in Proc. 33rd AAAI, 2019, pp. 3296–3303.

[23] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and

T. Darrell, ‘‘DeCAF: A deep convolutional activation feature for genericvisual recognition,’’ in Proc. ICML, 2014, pp. 1–9.

[24] A. S. Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, ‘‘CNN features

off-the-shelf: An astounding baseline for recognition,’’ in Proc. IEEE Conf.

CVPR Workshops, Jun. 2014, pp. 806–813.

[25] J. Howard and S. Ruder, ‘‘Universal language model fine-tuning

for text classification,’’ 2018, arXiv:1801.06146. [Online]. Available:

https://arxiv.org/abs/1801.06146