









- Proceedings of the IEEE conference on computer vision and pattern recognition.* 580–587.
- [9] Forrest Iandola, Matt Moskewicz, Sergey Karayev, Ross Girshick, Trevor Darrell, and Kurt Keutzer. 2014. Densenet: Implementing efficient convnet descriptor pyramids. *arXiv preprint arXiv:1404.1869* (2014).
- [10] Gunhee Kim, Seungwhan Moon, and Leonid Sigal. 2015. Joint photo stream and blog post summarization and exploration. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 3081–3089.
- [11] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. 2012. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*. 1097–1105.
- [12] Kuang-Huei Lee, Xi Chen, Gang Hua, Houdong Hu, and Xiaodong He. 2018. Stacked cross attention for image-text matching. In *Proceedings of the European Conference on Computer Vision (ECCV)*. 201–216.
- [13] Liudmila Prokhorenkova, Gleb Gusev, Aleksandr Vorobev, Anna Veronika Dorogush, and Andrey Gulin. 2018. CatBoost: unbiased boosting with categorical features. In *Advances in neural information processing systems*. 6638–6648.
- [14] Yao Qin, Dongjin Song, Haifeng Chen, Wei Cheng, Guofei Jiang, and Garrison W Cottrell. 2017. A Dual-Stage Attention-Based Recurrent Neural Network for Time Series Prediction. In *IJCAI*.
- [15] Weijie Su, Xizhou Zhu, Yue Cao, Bin Li, Lewei Lu, Furu Wei, and Jifeng Dai. 2019. VL-BERT: Pre-training of Generic Visual-Linguistic Representations. In *International Conference on Learning Representations*.
- [16] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. 2015. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 1–9.
- [17] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in neural information processing systems*. 5998–6008.
- [18] Bo Wu, Wen-Huang Cheng, Yongdong Zhang, and Tao Mei. 2016. Time Matters: Multi-scale Temporalization of Social Media Popularity. In *Proceedings of the 2016 ACM on Multimedia Conference (ACM MM)* (Amsterdam, The Netherlands).
- [19] Bo Wu, Wen-Huang Cheng, Yongdong Zhang, Huang Qiushi, Li Jintao, and Tao Mei. 2017. Sequential Prediction of Social Media Popularity with Deep Temporal Context Networks. In *International Joint Conference on Artificial Intelligence (IJCAI)* (Melbourne, Australia).
- [20] Bo Wu, Tao Mei, Wen-Huang Cheng, and Yongdong Zhang. 2016. Unfolding Temporal Dynamics: Predicting Social Media Popularity Using Multi-scale Temporal Decomposition. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI)* (Phoenix, Arizona).
- [21] Qidi Xu, Haocheng Xu, Weilong Chen, Chaojun Han, Haoyang Li, Wenxin Tan, Fumin Shen, and Heng Tao Shen. 2019. Time-aware Session Embedding for Click-Through-Rate Prediction. In *Proceedings of the 27th ACM International Conference on Multimedia*. 2617–2621.
- [22] Zhenguo Yang, Qing Li, Wenyin Liu, Yun Ma, and Min Cheng. 2017. Dual graph regularized NMF model for social event detection from Flickr data. *World Wide Web* 20, 5 (2017), 995–1015.
- [23] Zhenguo Yang, Qing Li, Zheng Lu, Yun Ma, Zhiguo Gong, and Wenyin Liu. 2017. Dual structure constrained multimodal feature coding for social event detection from flickr data. *ACM Transactions on Internet Technology (TOIT)* 17, 2 (2017), 1–20.
- [24] Sheng Yu and Subhash Kak. [n.d.]. A Survey of Prediction Using Social Media. ([n. d.]).