Literature Review for Ingur Thesis References (Farzane Part)

Reza

September 14, 2024

Contents

Bibliography

[1] P.-H. C. Le and X. Li, "Binaryvit: pushing binary vision transformers towards convolutional models," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 4664–4673.

[2] S. H. Lee, S. Lee, and B. C. Song, "Vision transformer for small-size datasets," *CoRR*, vol. abs/2112.13492, 2021. [Online]. Available: https://arxiv.org/abs/2112.13492

[3] A. Steiner, A. Kolesnikov, X. Zhai, R. Wightman, J. Uszkoreit, and L. Beyer, "How to train your vit? data, augmentation, and regularization in vision transformers," *CoRR*, vol. abs/2106.10270, 2021. [Online]. Available: https://arxiv.org/abs/2106.10270

[4] H. Touvron, M. Cord, M. Douze, F. Massa, A. Sablayrolles, and H. Jégou, "Training data-efficient image transformers & distillation through attention," *CoRR*, vol. abs/2012.12877, 2020. [Online]. Available: https://arxiv.org/abs/2012.12877

[5] H. Touvron, M. Cord, A. Sablayrolles, G. Synnaeve, and H. Jégou, "Going deeper with image transformers," *CoRR*, vol. abs/2103.17239, 2021. [Online]. Available: https://arxiv.org/abs/2103.17239

[6] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," *CoRR*, vol. abs/1706.03762, 2017. [Online]. Available: http://arxiv.org/abs/1706.03762

[7] D. Wodajo and S. Atnafu, "Deepfake video detection using convolutional vision transformer," *CoRR*, vol. abs/2102.11126, 2021. [Online]. Available: https://arxiv.org/abs/2102.11126

[8] M. Zhu, K. Han, Y. Tang, and Y. Wang, "Visual transformer pruning," *CoRR*, vol. abs/2104.08500, 2021. [Online]. Available: https://arxiv.org/abs/2104.08500

[9] R.-J. Zhu, Y. Zhang, E. Sifferman, T. Sheaves, Y. Wang, D. Richmond, P. Zhou, and J. K. Eshraghian, "Scalable matmul-free language modeling," 2024. [Online]. Available: https://arxiv.org/abs/2406.02528

[10] R.-J. Zhu, Q. Zhao, G. Li, and J. K. Eshraghian, "Spikegpt: Generative pre-trained language model with spiking neural networks," 2024. [Online]. Available: https://arxiv.org/abs/2302.13939

[11] Y. J. Heo, Y. J. Choi, Y. Lee, and B. Kim, "Deepfake detection scheme based on vision transformer and distillation," *CoRR*, vol. abs/2104.01353, 2021. [Online]. Available: https://arxiv.org/abs/2104.01353

[12] P. Korshunov and S. Marcel, "Deepfakes: a new threat to face recognition? assessment and detection," *CoRR*, vol. abs/1812.08685, 2018. [Online]. Available: http://arxiv.org/abs/1812.08685

[13] W. Maass, "Networks of spiking neurons: the third generation of neural network models," *Neural networks*, vol. 10, no. 9, pp. 1659–1671, 1997.

[14] R. Tolosana, S. Romero-Tapiador, J. Fiérrez, and R. Vera-Rodríguez, "Deepfakes evolution: Analysis of facial regions and fake detection performance," *CoRR*, vol. abs/2004.07532, 2020. [Online]. Available: https://arxiv.org/abs/2004.07532

[15] X. Yang, Y. Li, and S. Lyu, "Exposing deep fakes using inconsistent head poses," *CoRR*, vol. abs/1811.00661, 2018. [Online]. Available: http://arxiv.org/abs/1811.00661

[16] V. L. L. Thing, "Deepfake detection with deep learning: Convolutional neural networks versus transformers," 2023. [Online]. Available: https://arxiv.org/abs/2304.03698

[17] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," *IEEE Signal Processing Letters*, vol. 23, no. 10, p. 1499–1503, Oct. 2016. [Online]. Available: http://dx.doi.org/10.1109/LSP.2016.2603342

[18] W. Huang, Y. Liu, H. Qin, Y. Li, S. Zhang, X. Liu, M. Magno, and X. Qi, "Billm: Pushing the limit of post-training quantization for llms," 2024. [Online]. Available: https://arxiv.org/abs/2402.04291

[19] Z. Li, X. Liu, J. Zhang, and Q. Gu, "Repquant: Towards accurate post-training quantization of large transformer models via scale reparameterization," 2024. [Online]. Available: https://arxiv.org/abs/2402.05628

[20] Z. Liu, Y. Wang, K. Han, S. Ma, and W. Gao, "Post-training quantization for vision transformer," *CoRR*, vol. abs/2106.14156, 2021. [Online]. Available: https://arxiv.org/abs/2106.14156

[21] S. Ma, H. Wang, L. Ma, L. Wang, W. Wang, S. Huang, L. Dong, R. Wang, J. Xue, and F. Wei, "The era of 1-bit llms: All large language models are in 1.58 bits," 2024. [Online]. Available: https://arxiv.org/abs/2402.17764

[22] B. Zhang and R. Sennrich, "Root mean square layer normalization," 2019. [Online]. Available: https://arxiv.org/abs/1910.07467

[23] H. Wang, S. Ma, L. Dong, S. Huang, H. Wang, L. Ma, F. Yang, R. Wang, Y. Wu, and F. Wei, "Bitnet: Scaling 1-bit transformers for large language models," 2023. [Online]. Available: https://arxiv.org/abs/2310.11453

[24] L. Jiang, R. Li, W. Wu, C. Qian, and C. C. Loy, "Deeperforensics-1.0: A large-scale dataset for real-world face forgery detection," 2020. [Online]. Available: https://arxiv.org/abs/2001.03024

[25] Y. Li, X. Yang, P. Sun, H. Qi, and S. Lyu, "Celeb-df: A large-scale challenging dataset for deepfake forensics," 2020. [Online]. Available: https://arxiv.org/abs/1909.12962