VidMagic: Prompt to Image Clips

Ms. Nayana Ghuikar[1], Mr. Tejas Moon[2], Mr. Prasad Bhagat[3],
Mr. Sushant Chafale[4], Mr. Gaurav Sabe[5]
Assistant Professor, Department of Information Technology[1]
Student, Department of Information Technology[2,3,4,5]
Shri Sant Gajanan Maharaj College of Engineering Shegaon, Maharashtra, India

Abstract*: Prompt to image clips is a modern innovation that uses artificial intelligence to make videos from composed descriptions. This think about looks at the most recent strategies in this range. It talks around how analysts are working with the information and what sorts of computer programs they are utilizing. It moreover looks at how they check if the recordings are any great. The paper too talks around the issues with this innovation and what we still require to figure out. In general, these instruments have a part of potential for making recordings, making substance, and making a difference with education.*

REFERENCES

[1] T. Zia, S. Arif, S. Murtaza, and M. A. Ullah, "Prompt to-Image Generation with Attention Based Recurrent Neural Networks," arXiv preprint arXiv:2001.06658, 2020.

[2] S. Reed, Z. Akata, X. Yan, L. Logeswaran, B. Schiele, and H. Lee, "Generative adversarial Prompt to image synthesis," in Proceedings of the 33rd International Conference on Machine Learning, New York, NY, USA, 2016, pp. 1060-1069

[3] N. A. Fotedar and J. H. Wang, "Bumblebee: Prompt to-Image Generation with Transformers," in Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, 2019, pp. 3465-3469.

[4] H. Chang, H. Zhang, J. Barber, A. J. Maschinot, J. Lezama, L. Jiang, M. -H. Yang, K. Murphy, W. T. Freeman, M. Rubinstein, Y. Li, and D. Krishnan, "Muse: Prompt to-

Image Generation via Masked Generative Transformers," arXiv preprint arXiv:2301.00704, 2023.

[5] A. Haleem, M. Javaid, M. A. Qadri, R. P. Singh, and R. Suman, "Artificial intelligence (AI) applications for marketing: A literaturebased study," International Journal of Intelligent Networks, vol. 3, pp. 119-132, 2022. doi: 10.1016/j.ijin.2022.08.005

[6] S. Aktay, "The usability of Images Generated by Artificial Intelligence (AI) in Education," International Technology and Education Journal, vol. 6, no. 2, pp. 51-62, 2022.

[7] E. Cetinic and J. She, "Understanding and Creating Art with AI: Review and Outlook," ACM Trans. Multimedia Comput. Commun. Appl., vol. 18, no. 2, Article 66, May 2022, pp. 1-22, doi: 10.1145/3475799.

[8] M. Ding, W. Zheng, W. Hong, and J. Tang, "CogView2: Faster and Better Prompt to-Image Generation via Hierarchical Transformers," arXiv, 2022. [Online]. Available: https://arxiv.org/abs/2204.14217. [Accessed: March 18, 2023].

[9] M. Ding, Z. Yang, W. Hong, W. Zheng, C. Zhou, D. Yin, J. Lin, X. Zou, Z. Shao, H. Yang, and J. Tang, "CogView: Mastering Prompt toImage Generation via Transformers," arXiv:2105.13290 [cs.CV], 2021.

[10] A. Ramesh, P. Dhariwal, A. Nichol, C. Chu, and M. Chen, "Hierarchical Prompt-Conditional Image Generation with CLIP Latents," in arXiv preprint arXiv:2202.10775, 202

[11] C. Saharia, W. Chan, S. Saxena, L. Li, J. Whang, E. Denton, S. K. Seyed Ghasemipour, B. Karagol Ayan, S. S. Mahdavi, R. G. Lopes, T. Salimans, J. Ho, D. J. Fleet, and M. Norouzi, "Photorealistic Prompt toImage Diffusion Models with Deep Language Understanding," arXiv:2205.11487 [cs.CV], May 2022.

[12] U. Singer, A. Polyak, T. Hayes, X. Yin, J. An, S. Zhang, Q. Hu, H. Yang, O. Ashual, O. Gafni, D. Parikh, S. Gupta, and Y. Taigman, "Make-A-Video: Prompt to image clips Generation without Prompt-Video Data," arXiv:2209.14792 [cs.CV], Sep. 2022.

[13] J. Ho, W. Chan, C. Saharia, J. Whang, R. Gao, A. Gritsenko, D. P. Kingma, B. Poole, M. Norouzi, D. J. Fleet, and T. Salimans, "Imagen Video: High Definition Video Generation with Diffusion Models," arXiv preprint arXiv:2210.02303, Oct. 2022. [Online]. Available: https://arxiv.org/abs/2210.02303.

[14] R. Villegas, M. Babaeizadeh, P.-J. Kindermans, H. Moraldo, H. Zhang, M. T. Saffar, S. Castro, J. Kunze, and D. Erhan, "Phenaki: Variable Length Video Generation from Open Domain Promptual Description," arXiv:2210.02399 [cs.CV], Oct. 2022.

[15] W. Hong, M. Ding, W. Zheng, X. Liu, and J. Tang, "CogVideo: Largescale Pretraining for Prompt to image clips Generation via Transformers," arXiv:2205.15868 [cs.CV], May 2022.

[16] C. Wu, L. Huang, Q. Zhang, B. Li, L. Ji, F. Yang, G. Sapiro, and N. Duan, "GODIVA: Generating Open-DomaIn Videos from nAtural Descriptions," arXiv:2104.14806, Apr. 2021.

[17] C. Wu, J. Liang, L. Ji, F. Yang, Y. Fang, D. Jiang, and N. Duan, "NÜWA: Visual Synthesis Pre-training for Neural visUal World creAtion," arXiv:2111.12417 [cs.CV], Nov. 2021.

[18] B. Bordia and S. R. Bowman, "Identifying and Reducing Gender Bias in Word-Level Language Models," arXiv:1904.03035 [cs.CL], 2019.

[19] A. Birhane, V. U. Prabhu, and E. Kahembwe, "Multimodal Datasets: Misogyny, Pornography, and Malignant Stereotypes," arXiv:2110.01963, 2021.

[20] E. M. Bender, T. Gebru, A. McMillan-Major, and S. Shmitchell, "On the dangers of stochastic parrots: Can language models be too big?" in Proc. FAccT, 2021

Copyright to IJARSCT
www.ijarsct.co.in

DOI: 10.48175/IJARSCT-17473

ISSN
2581-9429
IJARSCT

427