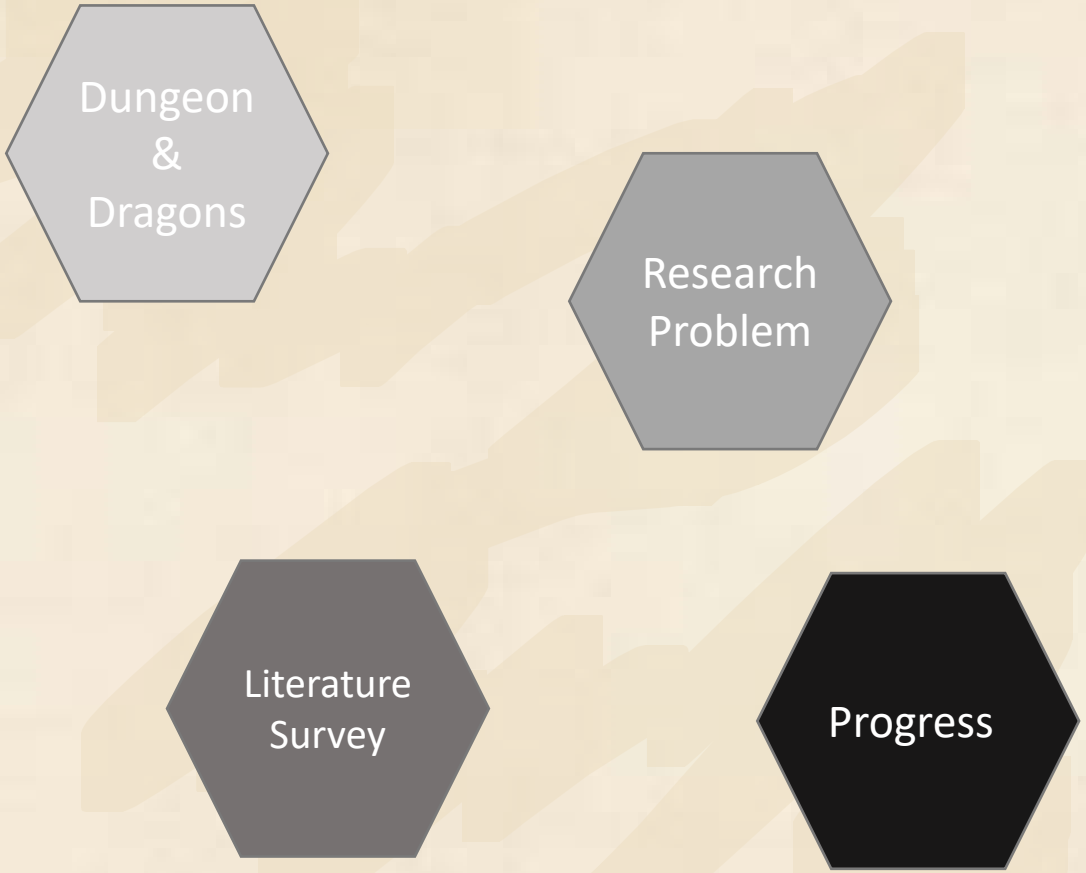
Sustained Context-Aware Image Generation for the Dungeons & Dragons Domain

229403G - WAG Weerasundara

Content


Dungeon
&
Dragons
Research
Problem
Literature
Survey
Progress

Dungeons and Dragons

- Tabletop Role playing game
- A DM designs or follow a pre existing adventure
- Use props & images for immersion

Figure 1: A group of people playing D&D from the Critical Role web series.

Research Problem

- Generating images consistent with the given narrative for a D&D adventure
  - Difficult to create art for a given narrative
  - Hard to maintain consistency of images across sessions
  - Images should be consistent with the lore
  - Pre existing or third party images needs to be incorporated

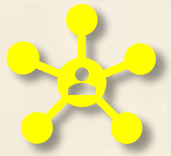
Literature Survey

Key Phrase extraction

Prompt generation

Image generation

Evaluation



Consistency

Key Phrase Extraction

- Pre-processing & Summarization
- Key phrase extraction using machine learning
- Named Entity Recognition

Key Phrase Extraction

Pre-processing & Summarization

- Text mining [1,2]
- Narrative extraction [3]
- Text summarization [4-7]

[1] E. Papagiannopoulou and G. Tsoumakas, “A review of keyphrase extraction,” Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 10, no. 2, p. e1339, 2020.
[2] S. Vijayarani, M. J. Ilamathi, M. Nithya et al., “Preprocessing techniques for text mining-an overview,” International Journal of Computer Science & Communication Networks, vol. 5, no. 1, pp. 7–16, 2015.
[3] B. Santana, R. Campos, E. Amorim, A. Jorge, P. Silvano, and S. Nunes, “A survey on narrative extraction from textual data,” Artificial Intelligence Review, pp. 143, 2023.
[4] R. Srivastava, P. Singh, K. Rana, and V. Kumar, “A topic modeled unsupervised approach to single document extractive text summarization,” Knowledge-Based Systems, vol. 246, p. 108636, 2022.
[5] S. Abdel-Salam and A. Rafea, “Performance study on extractive text summarization using bert models,” Information, vol. 13, no. 2, p. 67, 2022.
[6] H. Y. Koh, J. Ju, M. Liu, and S. Pan, “An empirical survey on long document summarization: Datasets, models, and metrics,” ACM computing surveys, vol. 55, no. 8, pp. 1–35, 2022.
[7] D. Yadav, J. Desai, and A. K. Yadav, “Automatic text summarization methods: A comprehensive review,” arXiv preprint arXiv:2204.01849, 2022.

Key Phrase Extraction

Key phrase extraction using machine learning

- Data augmentation through Domain-specific Phrase Annotation[8]
- Selectivity-based Keyword Extraction (SBKE) method[9]

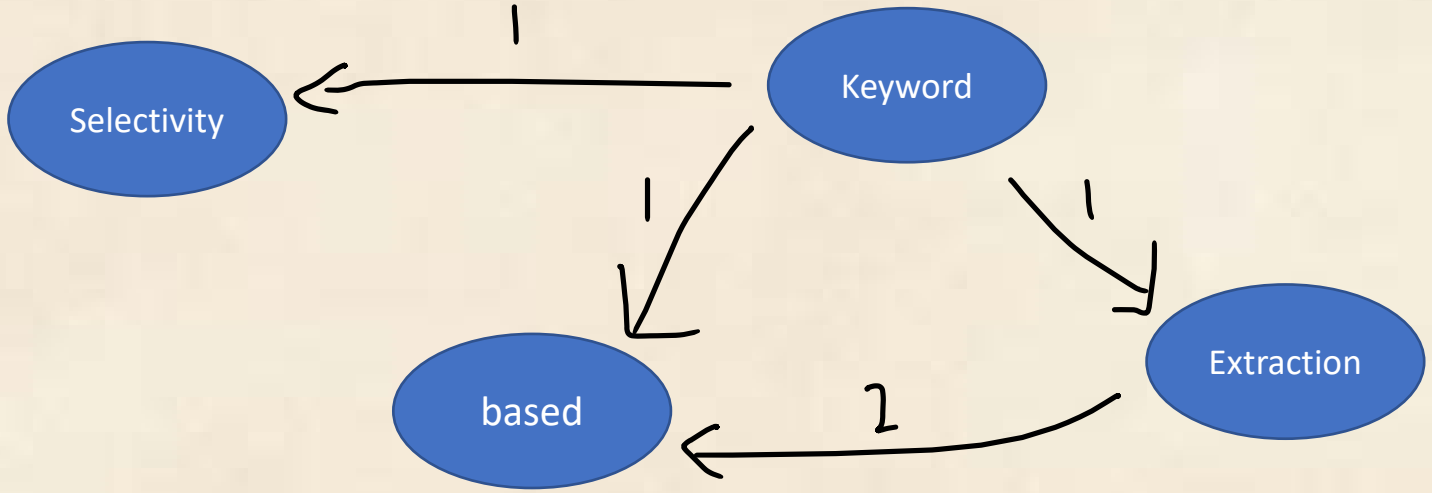


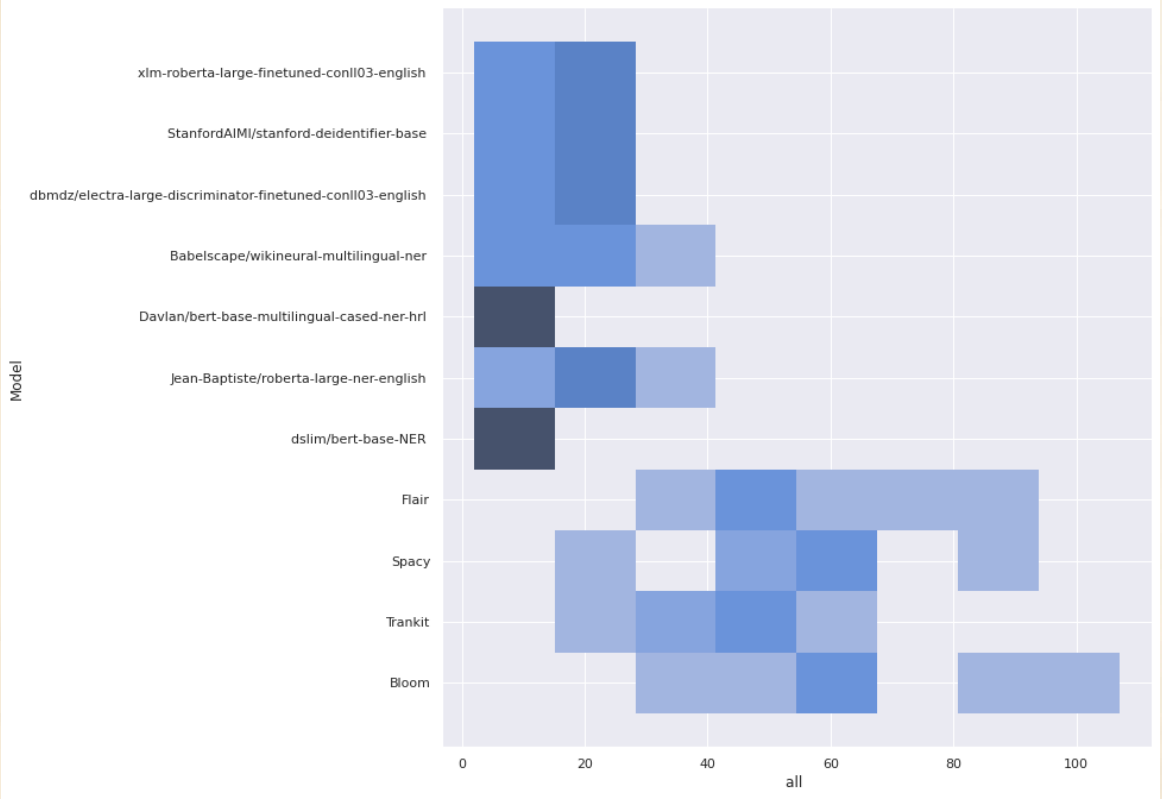
Figure 2: Diagram for SKBE process.

[8] A. P. B. Veyseh, N. Meister, F. Dernoncourt, and T. H. Nguyen, "Improving keyphrase extraction with data augmentation and information filtering," arXiv preprint arXiv:2209.04951, 2022
[9] S. Beliga, "Keyword extraction: a review of methods and approaches," University of Rijeka, Department of Informatics, Rijeka, vol. 1, no. 9, 2014.

Key Phrase Extraction

Named Entity Recognition

- BERT based NER models[8]
- NER frameworks
  - Trankit [10]
  - Flair [11]
- General NLP frameworks
  - CoreNLP
  - Spacy [12]
- LLMs
  - BARD
  - LLAMA
  - Bloom [13]



Graph 1: Comparing all identified named entities for different NER models. Tested for different source books from D&D Beyond.

[10] M. Van Nguyen, V. D. Lai, A. P. B. Veyseh, and T. H. Nguyen, “Trankit: A light weight transformer-based toolkit for multilingual natural language processing,” arXiv preprint arXiv:2101.03289, 2021. 32

[11] A. Akbik, T. Bergmann, D. Blythe, K. Rasul, S. Schweter and R. Vollgraf, “FLAIR: An Easy-to-Use Framework for State-of-the-Art NLP,” in Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: System Demonstrations, pages 54-59, Minneapolis, Minnesota, 20191.

[12] M. Honnibal and I. Montani, “spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing,” unpublished.

[13] T. Le Scao et al., “BLOOM: A 176B-Parameter Open-Access Multilingual Language Model,” arXiv preprint arXiv:2211.05100, 20221.

Prompt generation

- Prompt Engineering and Design Principles
  - Six types of prompt modifiers were identified[14]
  - Identified common descriptors for enhancing generations
- Control over Text Generation with Templates and Constraints[15]
- Generative Pretrained Transformer (GPT) Models
  - Task-specific finetuning[16]
  - Quantifying event boundaries in continuous narratives[17]

[14] J. Oppenlaender, “Prompt engineering for text-based generative art,” arXiv preprint arXiv:2204.13988, 2022.
[15] S. W. McRoy, S. Channarukul, and S. S. Ali, “An augmented template-based approach to text realization,” Natural Language Engineering, vol. 9, no. 4, pp.381–420, 2003.
[16] A. Radford, K. Narasimhan, T. Salimans, I. Sutskever et al., “Improving language understanding by generative pre-training,” 2018.
[17] S. Michelmann, M. Kumar, K. A. Norman, and M. Toneva, “Large language models can segment narrative events similarly to humans,” Jan 2023. [Online]. Available: https://arxiv.org/abs/2301.10297

What's the difference between prompts ?

Original Prompt	Original Prompt
A dream of a distant galaxy	photo of a riverbank
	

Negative prompts

- “A dream of a distant galaxy”
- Reduce artifacts
- out of frame, lowres, text, error, cropped, worst quality, low quality, jpeg artifacts, ugly, …

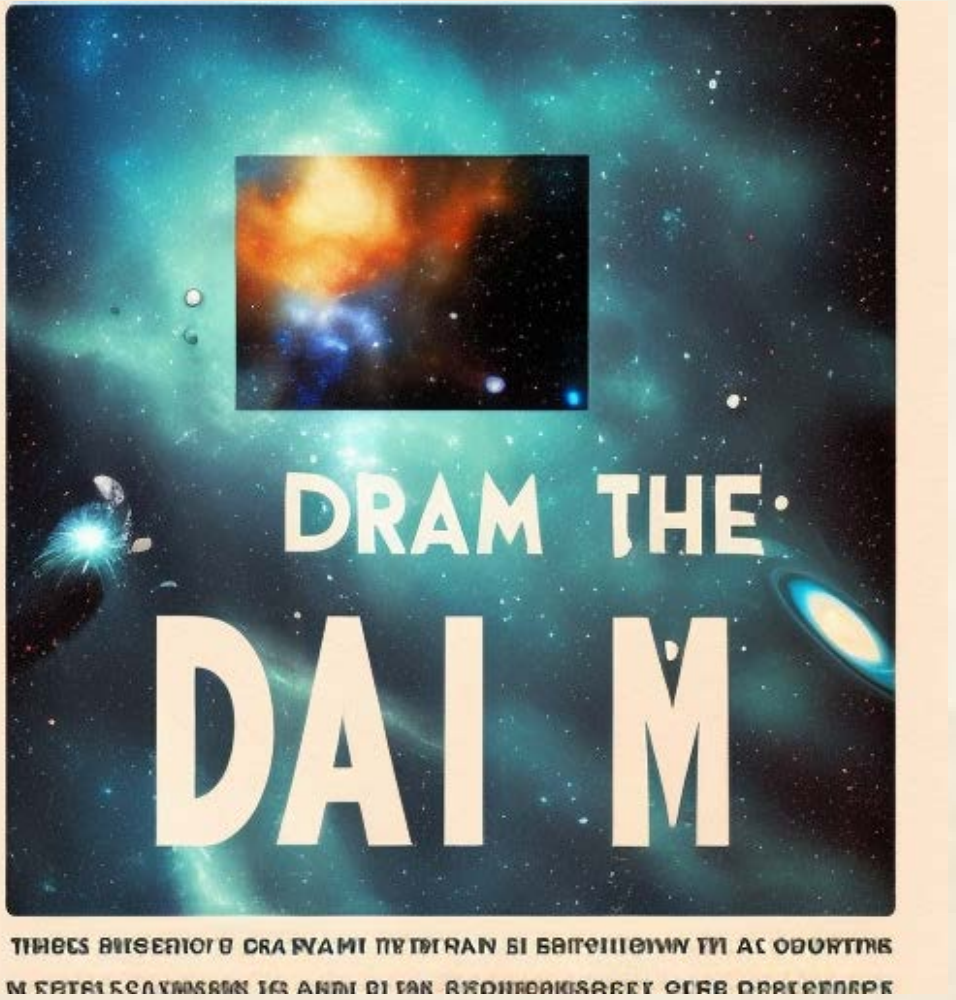


Figure 3: Example of a bad output generation.

Image Generation

- Image Synthesis with Generative Adversarial Networks
- Variational Autoencoders (VAEs) for Image Generation
- Diffusion models for Image Generation
- Image Style Transfer
- Image-to-Image Translation

Image Generation

Image Synthesis with Generative Adversarial Networks (GAN)

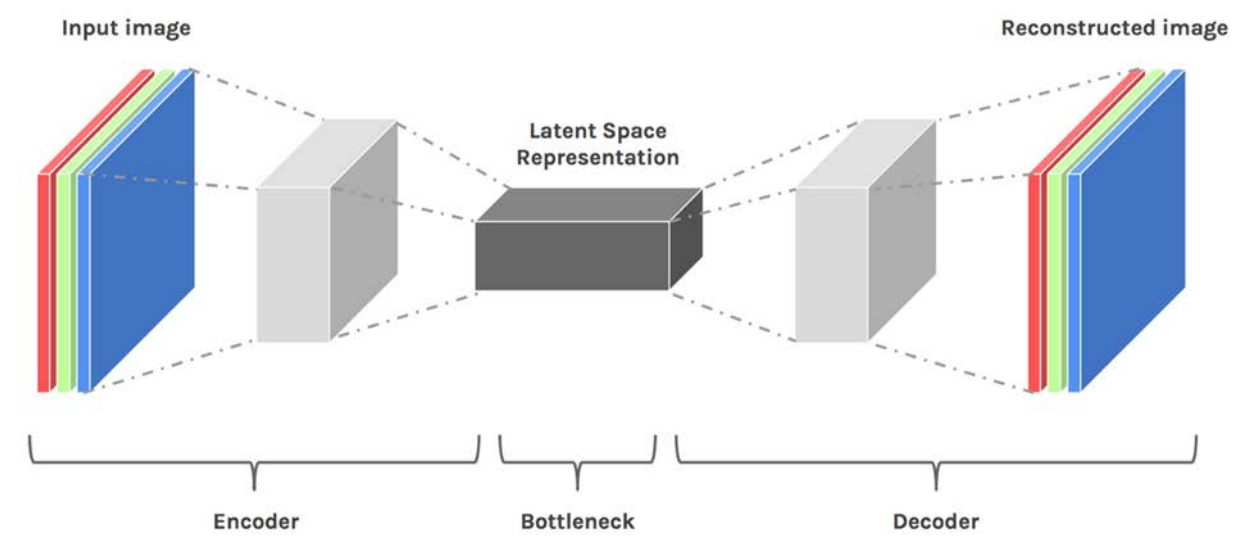
- Consist of a generator and discriminator, with the former taking random noise as input and producing synthetic images [18].
- Lu et al. [19] proposed a contextual GAN framework for sketch-to-image generation, while Li et al. [20] proposed StoryGAN for story visualization, both of which outperformed existing models in terms of image quality and consistency.
- The Parti model uses GANs to generate high-quality, photorealistic images from text prompts and shows remarkable performance [21].

[18] H. Huang, P. S. Yu, and C. Wang, "An introduction to image synthesis with generative adversarial nets," arXiv preprint arXiv:1803.04469, 2018.
[19] Y. Lu, S. Wu, Y.-W. Tai and C.-K. Tang, "Image Generation from Sketch Constraint Using Contextual GAN," in Computer Vision – ECCV 2018, Lecture Notes in Computer Science, vol 11213. Springer, Cham, 2018
[20] Y. Li, Z. Gan, Y. Shen, J. Liu, Y. Cheng, Y. Wu, L. Carin, D. Carlson, and J. Gao,"Storygan: A sequential conditional gan for story visualization," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 6329–6338.
[21] J. Yu, Y. Xu, J. Y. Koh, T. Luong, G. Baid, Z. Wang, V. Vasudevan, A. Ku, Y. Yang, B. K. Ayan et al., "Scaling autoregressive models for content-rich text-to-image generation," arXiv preprint arXiv:2206.10789, 2022.

Image Generation

Variational Autoencoders (VAEs) for Image Generation

- Combination of an encoder network and a decoder network
- Enforces prior to latent space [22].



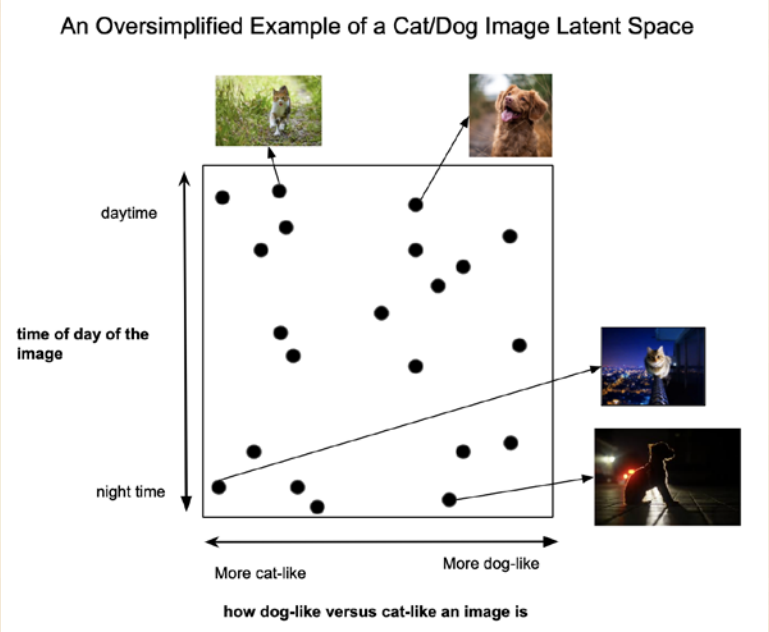


Figure 4: Explaining latent space & VAEs[22] .

[22] AutoEncoders are Essential in Deep Neural Nets. Towards Data Science. Retrieved March 21, 2023, from https://towardsdatascience.com/autoencoders-are-essential-in-deep-neural-nets-f0365b2d1d7c

Image Generation

Diffusion models for Image Generation

- Cascaded diffusion models[23]
- Text-driven image manipulation[24,25]
- Semantic diffusion guidance[26,27]

Figure 5: Semantic diffusion guidance with Instructpix2pix [27] .

[23] J. Ho, C. Saharia, W. Chan, D. J. Fleet, M. Norouzi, and T. Salimans, "Cascaded diffusion models for high fidelity image generation." J. Mach. Learn. Res., vol. 23, no. 47, pp. 1–33, 2022.

[24] A. Nichol, P. Dhariwal, A. Ramesh, P. Shyam, P. Mishkin, B. McGrew, I. Sutskever, and M. Chen, "Glide: Towards photorealistic image generation and editing with text-guided diffusion models," arXiv preprint arXiv:2112.10741,2021.

[25] G. Kim, T. Kwon, and J. C. Ye, "Diffusionclip: Text-guided diffusion models for robust image manipulation," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 2426–2435.

[26] X. Liu, D. H. Park, S. Azadi, G. Zhang, A. Chopikyan, Y. Hu, H. Shi, A. Rohrbach, and T. Darrell, "More control for free! image synthesis with semantic diffusion guidance," in Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2023, pp. 289–299.34

[27] Brooks, T., Holynski, A., & Efros, A. A. (2022). Instructpix2pix: Learning to follow image editing instructions. arXiv preprint arXiv:2211.09800.

Image Generation

Image Style Transfer

- Style transfer for GANs[28,29]
- Dreambooth[30]
- Aesthetic Gradients[31]

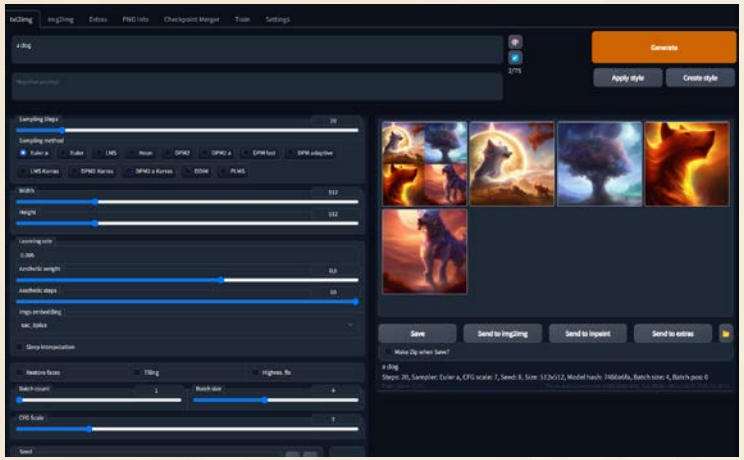


Figure 7: Example of image style through aesthetic embeddings [31] in Stable Diffusion web UI [32] .

Figure 6: Example of image style transfer from few images using dreambooth [30] .

[28] T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019, pp. 4401–4410.
[29] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila, "Analyzing and improving the image quality of stylegan," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2020, pp. 8110–8119.
[30] N. Ruiz, Y. Li, V. Jampani, Y. Pritch, M. Rubinstein, and K. Aberman, "Dreambooth: Fine tuning text-to-image diffusion models for subject-driven generation," arXiv preprint arXiv:2208.12242, 2022.
[31] V. Gallego, "Personalizing text-to-image generation via aesthetic gradients," arXiv preprint arXiv:2209.12330, 2022.
[32] GIGAZINE. (2022). Stable Diffusion Web UI that allows you to easily generate images from text. [online] Available at: https://gigazine.net/gsc_news/en/20220904-stable-diffusion-webui/ [Accessed 24 March 2023].

Image Generation

Image-to-Image Translation

- Prompt + image as input[33]
- Image + mask as input[34]
- Image as character[35]
- Natural language guidance[36]

Figure 8: Example of image to image translation [33] in Stable Diffusion web UI [32] .

[33] J. Ho, C. Saharia, W. Chan, D. J. Fleet, M. Norouzi, and T. Salimans, “Cascaded diffusion models for high fidelity image generation.” J. Mach. Learn. Res., vol. 23, no. 47, pp. 1–33, 2022.

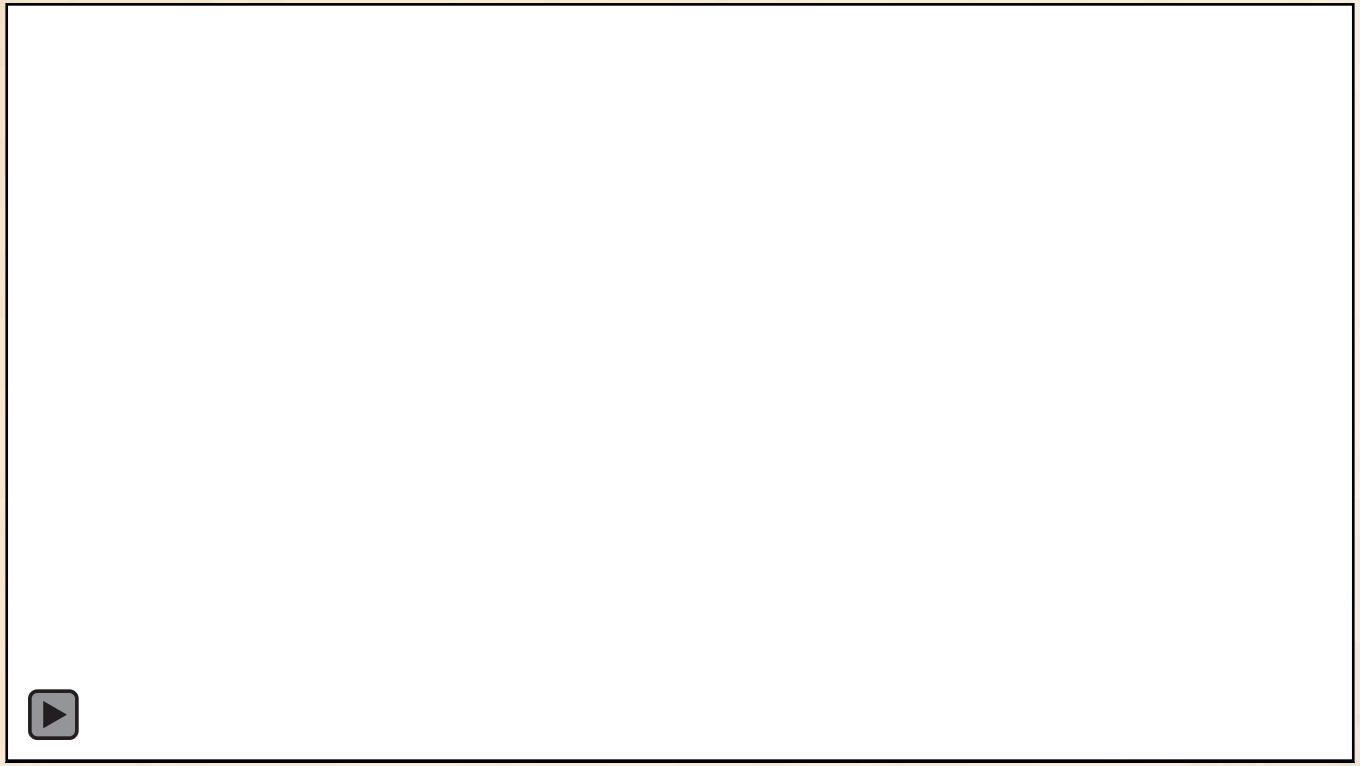
[34] R. Rombach, A. Blattmann, D. Lorenz, P. Esser and B. Ommer, “High-Resolution Image Synthesis With Latent Diffusion Models,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2022, pp. 10684-10695.

[35] X. Pan, P. Qin, Y. Li, H. Xue and W. Chen, “Synthesizing Coherent Story with Auto-Regressive Latent Diffusion Models,” arXiv preprint arXiv:2211.10950, 2022

[36] Brooks, T., Holynski, A., & Efros, A. A. (2022). Instructpix2pix: Learning to follow image editing instructions. arXiv preprint arXiv:2211.09800.

Consistency

- Document-Level Sentiment Analysis[37]
- Document-Level Relation Extraction[38]
- Image Temporal Consistency[39,40]



Video 1: Example of using temporal consistent images to form a video[41] .

[37] S. Behdenna, F. Barigou, and G. Belalem, “Document level sentiment analysis: a survey,” EAI Endorsed Transactions on Context-aware Systems and Applications, vol. 4, no. 13, pp. e2–e2, 2018.
[38] X. Han and L. Wang, “A novel document-level relation extraction method based on bert and entity information,” IEEE Access, vol. 8, pp. 96 912–96 919, 2020.
[39] G. Kim, H. Shim, H. Kim, Y. Choi, J. Kim, and E. Yang, “Diffusion video autoencoders: Toward temporally consistent face video editing via disentangled video encoding,” arXiv preprint arXiv:2212.02802, 2022.
[40] B. Kim, I. Han, and J. C. Ye, “Diffusemorph: Unsupervised deformable image registration along continuous trajectory using diffusion models,” arXiv preprint arXiv:2112.05149, 2021.36
[41] “Corridor Crew. (2023, 03 07). VFX Reveal Before & After - Anime Rock, Paper, Scissors [Video]. YouTube. https://www.youtube.com/watch?v=ljBSmQdL_Ow”

Evaluation

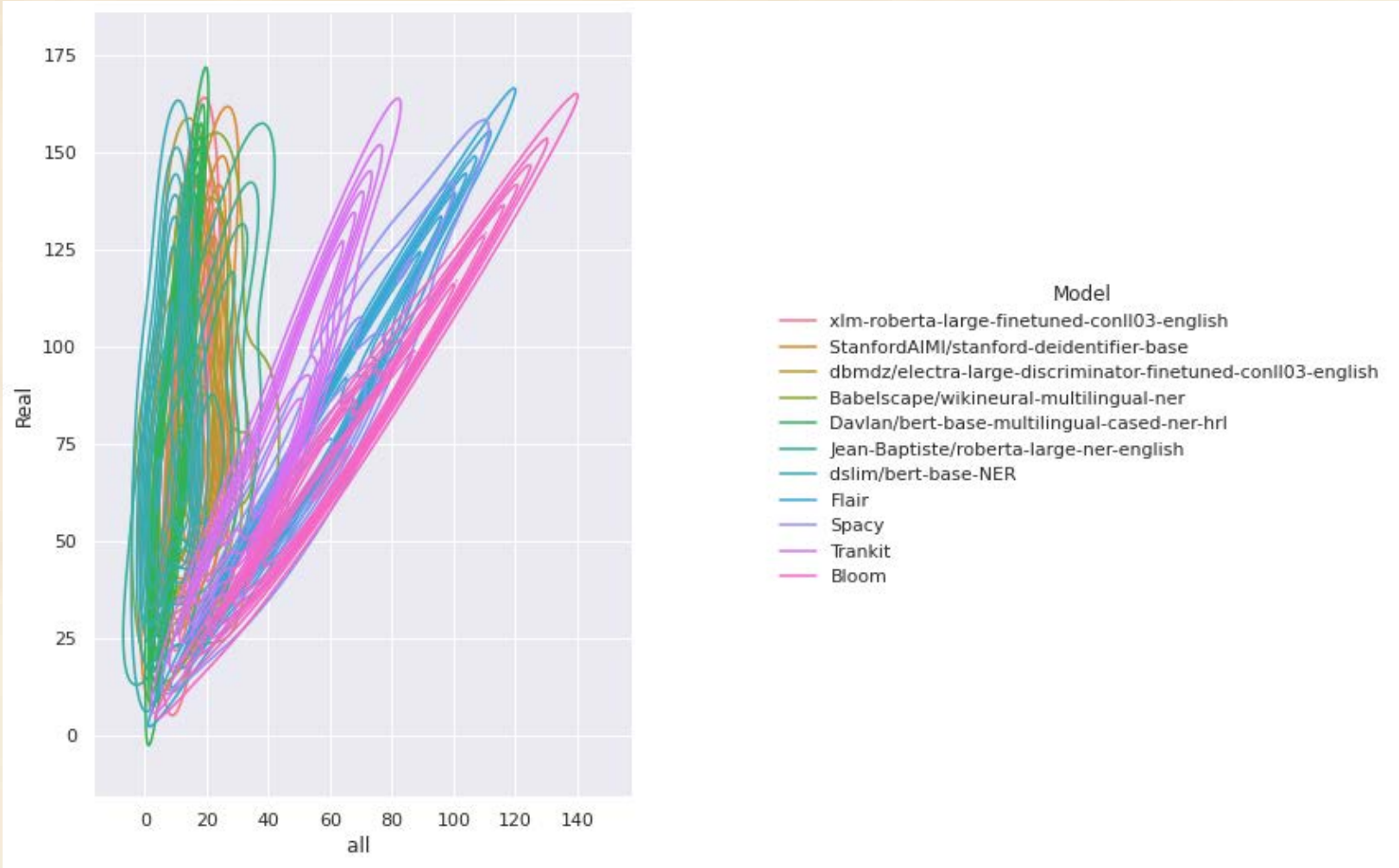
- Evaluation Metrics for Text Generation[41]
- Image Evaluation[42]
- Consistency evaluation[43]

[42] T. Zhang, V. Kishore, F. Wu, K. Q. Weinberger, and Y. Artzi, “Bertscore: Evaluating text generation with bert,” arXiv preprint arXiv:1904.09675, 2019.
[43] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” IEEE transactions on image processing, vol. 13, no. 4, pp. 600–612, 2004.
[44] A. Cherian and A. Sullivan, “Sem-gan: Semantically-consistent image-to-image translation,” in 2019 ieee winter conference on applications of computer vision (wacv). IEEE, 2019, pp. 1797–1806.

Progress

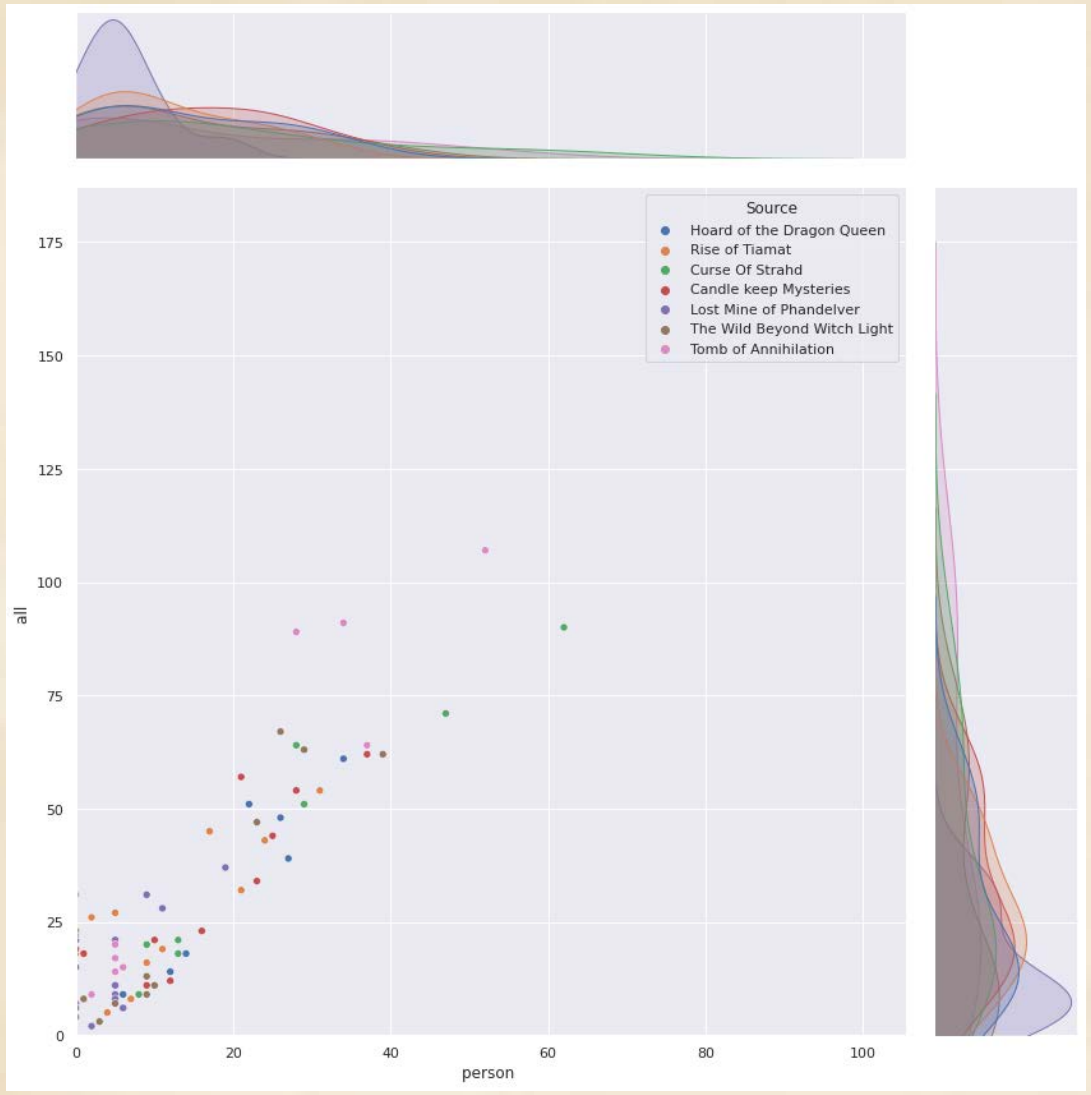
- Testing NER models
- Applying style to images
- Testing Progressive changes on images

Testing NER models



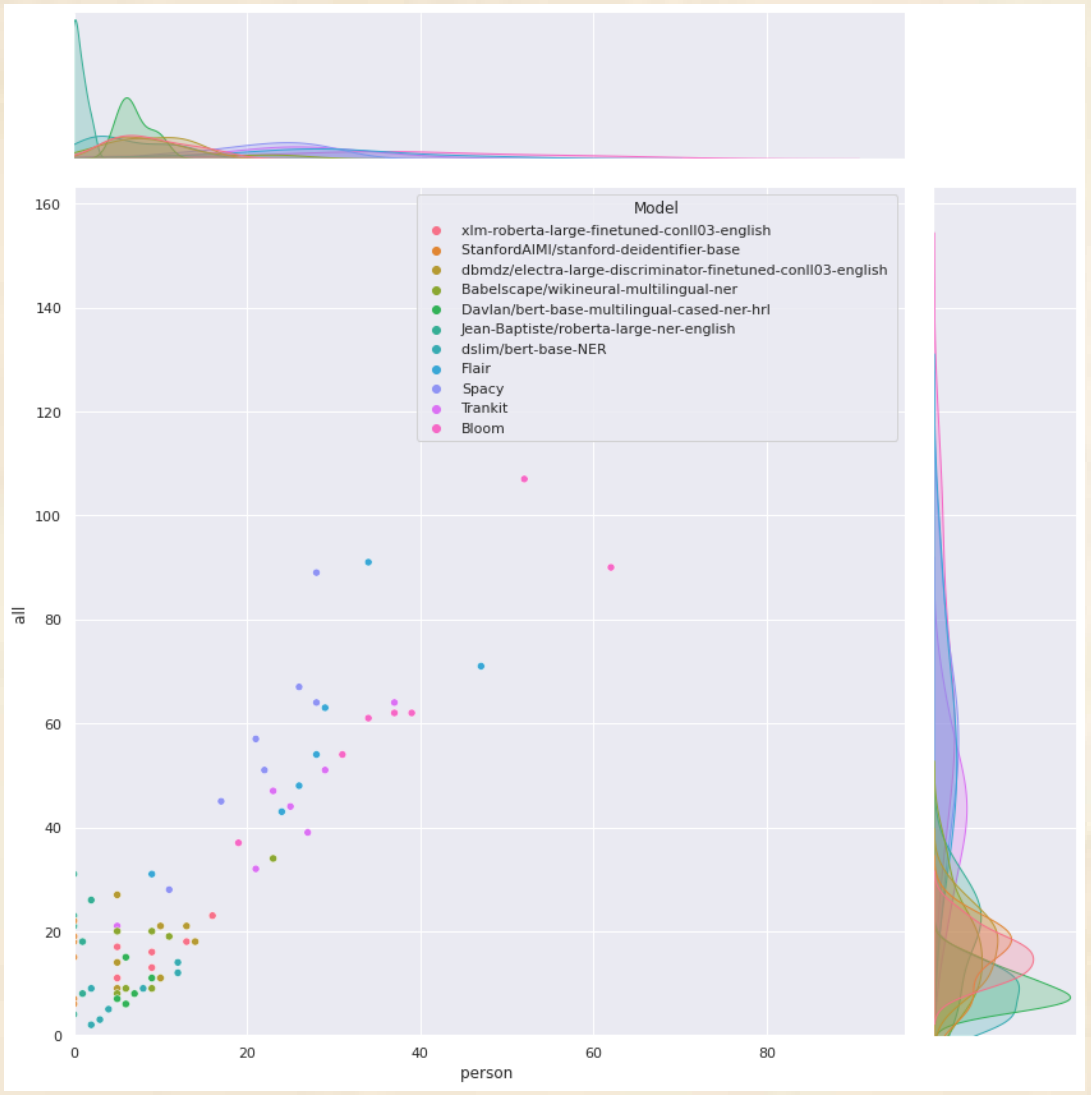
Graph 2: Comparing all identified named entities vs real named entities for different NER models.

By Source



Graph 3: A plot of all named entities vs characters identified by source books.

By Model



Graph 4: A plot of all named entities vs characters identified by NER models.

Applying style to images

portrait (painting) of tabaxi, de Rivia closeup, suit, collar, formal attire, D&D, fantasy, intricate, elegant, highly detailed, artstation, concept art, matte, sharp focus, (brush strokes), (oil on canvas), hearthstone, art by Titian and Greg Rutkowski and Rembrandt van Rijn and Alphonse Mucha

No style transfer

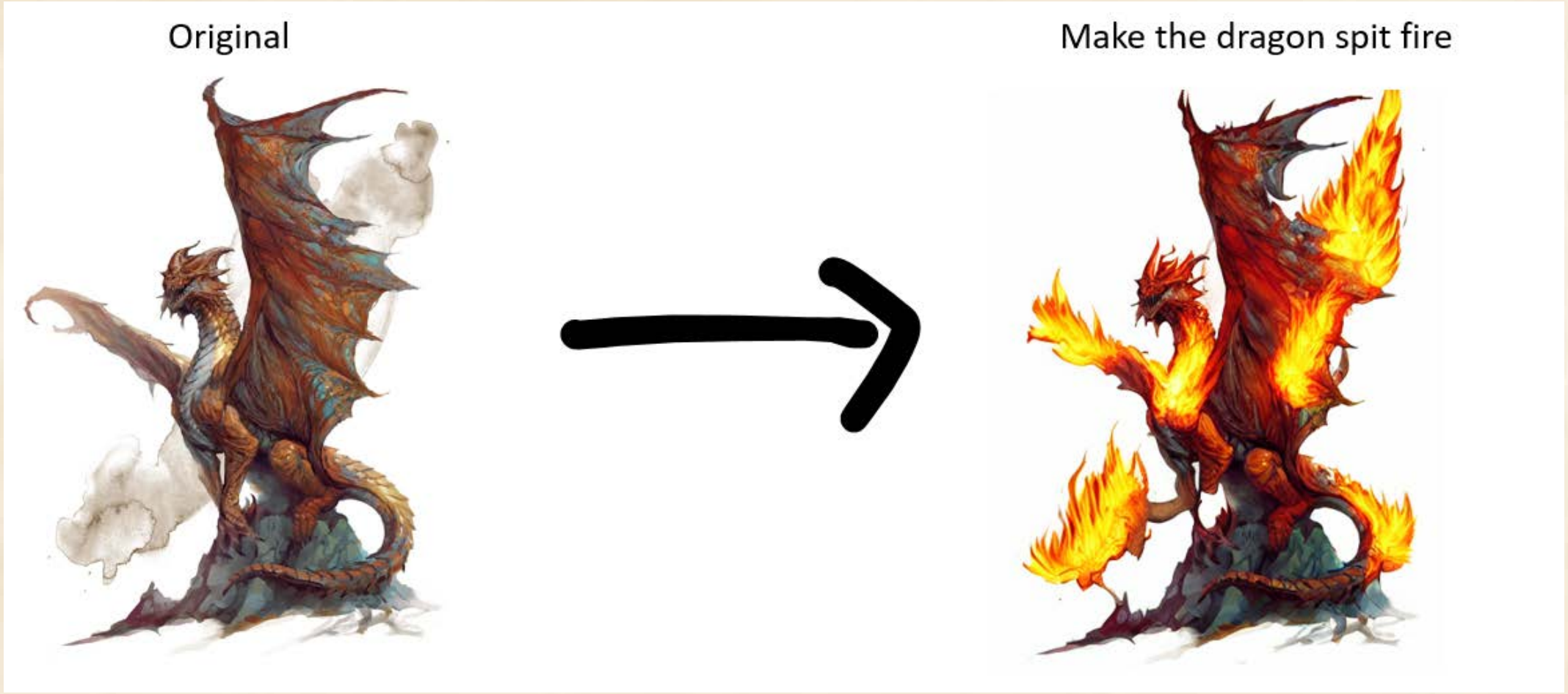
With style transfer

Testing Progressive changes on images

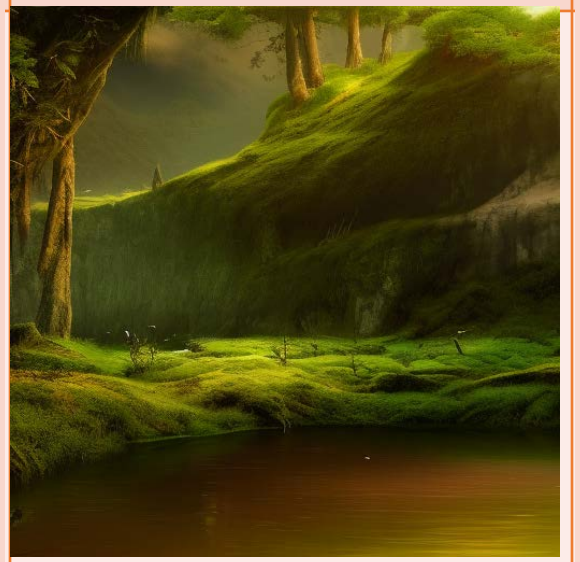
Original

Change the clothing to armor

Testing Progressive changes on images



Testing Progressive changes on images

A dream of a distant galaxy, concept art, matte painting, HQ, 4k		**photo of a riverbank, concept art, matte painting, HQ, 4k**	
Original		Original	
			

Iterative changes

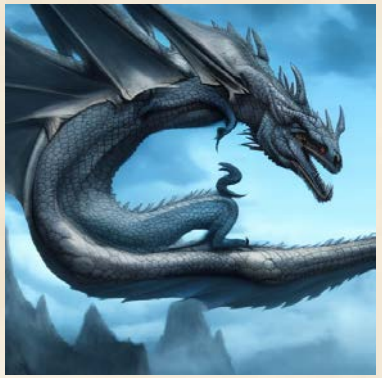
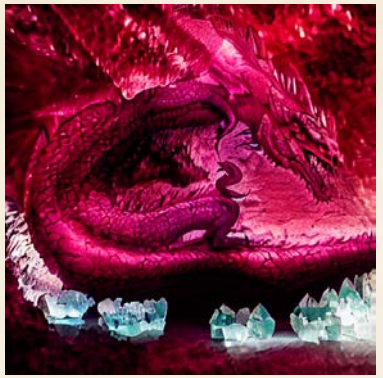


Image of a dragon, concept art, matte painting, HQ, 4k (Stable diffusion [1])

make only the dragon red (InstructPix2Pix [2])

dragon is inside a cave (InstructPix2Pix [2])

cave is a crystal cave (InstructPix2Pix [2])



make only the dragon red (InstructPix2Pix [2])

Comparison

Image of a red dragon in a crystal cave, concept art, matte painting, HQ, 4k

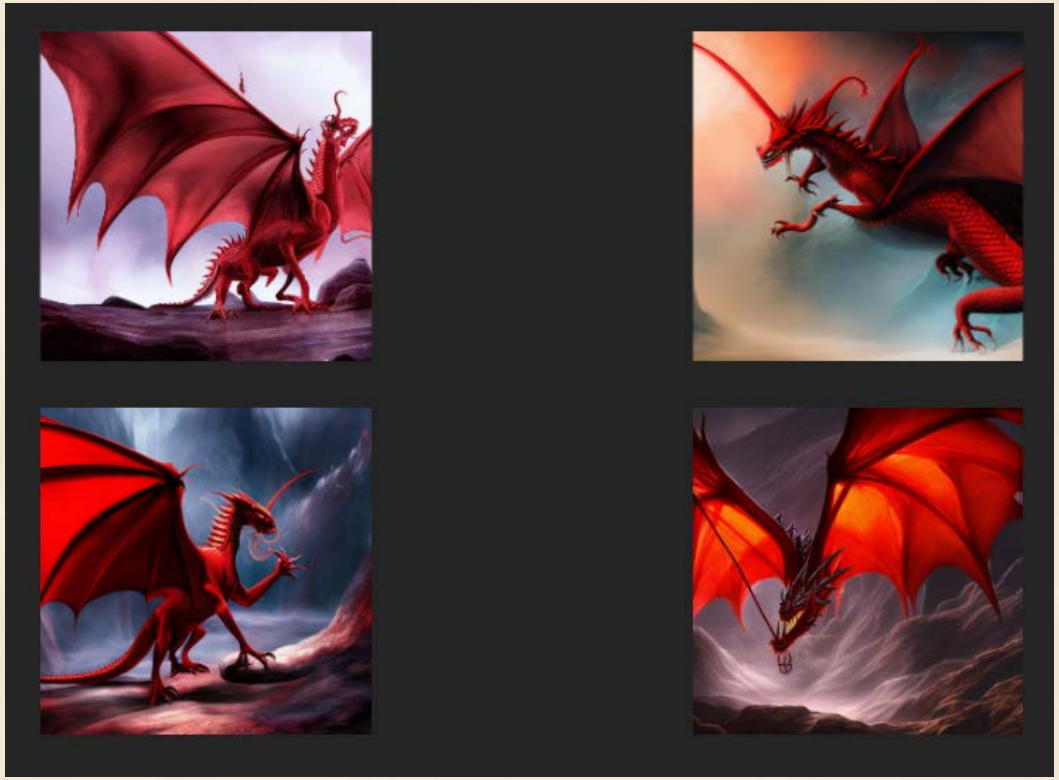
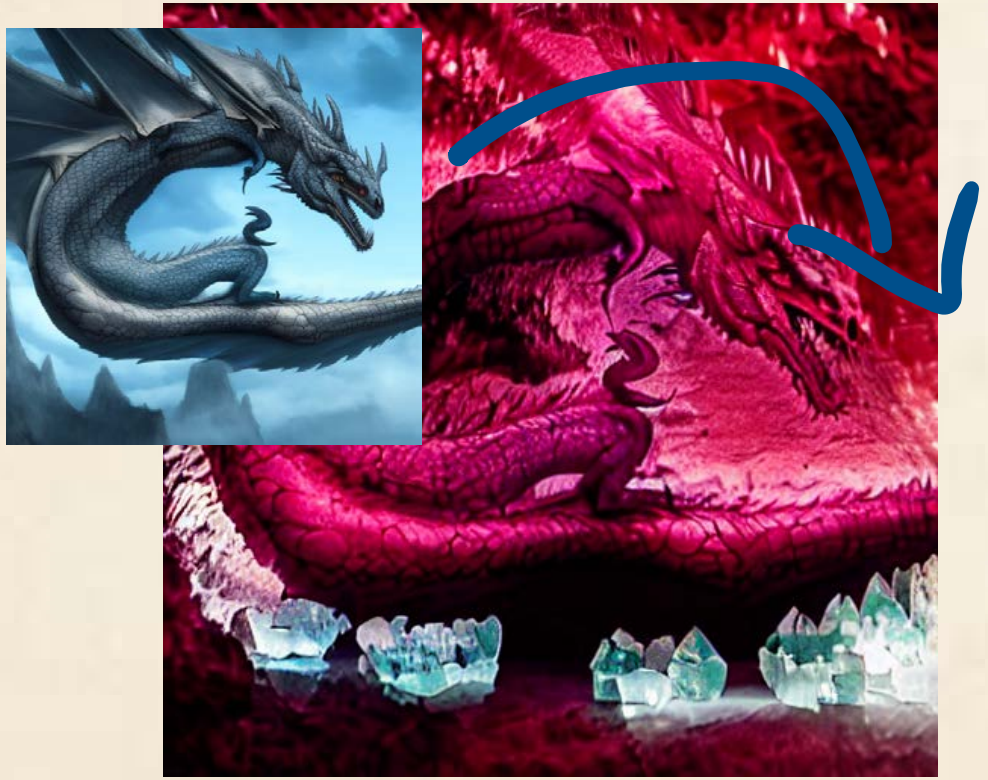


Image of a dragon, concept art, matte painting, HQ, 4k



Planned methodology

Figure 9: A page from Lost Mine Of Phandelver [45].

Figure 10: Images from D&Dbeyond [45].

[45] Wizards of the Coast, (2023, March 28). LOST MINE OF PHANDELVER. D&D Beyond. [Online]. Available: https://www.dndbeyond.com/sources/lmop

Reference

[1] E. Papagiannopoulou and G. Tsoumakas, “A review of keyphrase extraction,” Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 10, no. 2, p. e1339, 2020.

[2] S. Vijayarani, M. J. Ilamathi, M. Nithya et al., “Preprocessing techniques for text mining-an overview,” International Journal of Computer Science & Communication Networks, vol. 5, no. 1, pp. 7–16, 2015.

[3] B. Santana, R. Campos, E. Amorim, A. Jorge, P. Silvano, and S. Nunes, “A survey on narrative extraction from textual data,” Artificial Intelligence Review, pp. 143, 2023.

[4] R. Srivastava, P. Singh, K. Rana, and V. Kumar, “A topic modeled unsupervised approach to single document extractive text summarization,” Knowledge-Based Systems, vol. 246, p. 108636, 2022.

[5] S. Abdel-Salam and A. Rafea, “Performance study on extractive text summarization using bert models,” Information, vol. 13, no. 2, p. 67, 2022.

[6] H. Y. Koh, J. Ju, M. Liu, and S. Pan, “An empirical survey on long document summarization: Datasets, models, and metrics,” ACM computing surveys, vol. 55, no. 8, pp. 1–35, 2022.

[7] D. Yadav, J. Desai, and A. K. Yadav, “Automatic text summarization methods: A comprehensive review,” arXiv preprint arXiv:2204.01849, 2022.

[8] A. P. B. Veyseh, N. Meister, F. Dernoncourt, and T. H. Nguyen, “Improving keyphrase extraction with data augmentation and information filtering,” arXiv preprint arXiv:2209.04951, 2022.[2] S. Vijayarani, M. J. Ilamathi, M. Nithya et al., “Preprocessing techniques for text mining-an overview,” International Journal of Computer Science & Communication Networks, vol. 5, no. 1, pp. 7–16, 2015.

[9] S. Beliga, “Keyword extraction: a review of methods and approaches,” University of Rijeka, Department of Informatics, Rijeka, vol. 1, no. 9, 2014.

[10] M. Van Nguyen, V. D. Lai, A. P. B. Veyseh, and T. H. Nguyen, “Trankit: A light weight transformer-based toolkit for multilingual natural language processing,” arXiv preprint arXiv:2101.03289, 2021. 32

Reference

[11] A. Akbik, T. Bergmann, D. Blythe, K. Rasul, S. Schweter and R. Vollgraf, “FLAIR: An Easy-to-Use Framework for State-of-the-Art NLP,” in Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: System Demonstrations, pages 54-59, Minneapolis, Minnesota, 20191.

[12] M. Honnibal and I. Montani, “spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing,” unpublished.

[13] T. Le Scao et al., “BLOOM: A 176B-Parameter Open-Access Multilingual Language Model,” arXiv preprint arXiv:2211.05100, 20221.

[14] J. Oppenlaender, “Prompt engineering for text-based generative art,” arXiv preprint arXiv:2204.13988, 2022.

[15] S. W. McRoy, S. Channarukul, and S. S. Ali, “An augmented template-based approach to text realization,” Natural Language Engineering, vol. 9, no. 4, pp.381–420, 2003.

[16] A. Radford, K. Narasimhan, T. Salimans, I. Sutskever et al., “Improving language understanding by generative pre-training,” 2018.

[17] S. Michelmann, M. Kumar, K. A. Norman, and M. Toneva, “Large language models can segment narrative events similarly to humans,” Jan 2023. [Online]. Available: https://arxiv.org/abs/2301.10297

[18] H. Huang, P. S. Yu, and C. Wang, “An introduction to image synthesis with generative adversarial nets,” arXiv preprint arXiv:1803.04469, 2018.

[19] Y. Lu, S. Wu, Y.-W. Tai and C.-K. Tang, “Image Generation from Sketch Constraint Using Contextual GAN,” in Computer Vision – ECCV 2018, Lecture Notes in Computer Science, vol 11213. Springer, Cham, 2018

[20] Y. Li, Z. Gan, Y. Shen, J. Liu, Y. Cheng, Y. Wu, L. Carin, D. Carlson, and J. Gao,“Storygan: A sequential conditional gan for story visualization,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 6329–6338.

Reference

[21] J. Yu, Y. Xu, J. Y. Koh, T. Luong, G. Baid, Z. Wang, V. Vasudevan, A. Ku, Y. Yang, B. K. Ayan et al., “Scaling autoregressive models for content-rich text-to-image generation,” arXiv preprint arXiv:2206.10789, 2022.

[22] AutoEncoders are Essential in Deep Neural Nets. Towards Data Science. Retrieved March 21, 2023, from https://towardsdatascience.com/autoencoders-are-essential-in-deep-neural-nets-f0365b2d1d7c

[23] J. Ho, C. Saharia, W. Chan, D. J. Fleet, M. Norouzi, and T. Salimans, “Cascaded diffusion models for high fidelity image generation.” J. Mach. Learn. Res., vol. 23, no. 47, pp. 1–33, 2022.

[24] A. Nichol, P. Dhariwal, A. Ramesh, P. Shyam, P. Mishkin, B. McGrew, I. Sutskever, and M. Chen, “Glide: Towards photorealistic image generation and editing with text-guided diffusion models,” arXiv preprint arXiv:2112.10741,2021.

[25] G. Kim, T. Kwon, and J. C. Ye, “Diffusionclip: Text-guided diffusion models for robust image manipulation,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 2426–2435.

[26] X. Liu, D. H. Park, S. Azadi, G. Zhang, A. Chopikyan, Y. Hu, H. Shi, A. Rohrbach, and T. Darrell, “More control for free! image synthesis with semantic diffusion guidance,” in Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2023, pp. 289–299.34

[27] Brooks, T., Holynski, A., & Efros, A. A. (2022). Instructpix2pix: Learning to follow image editing instructions. arXiv preprint arXiv:2211.09800.

[28] T. Karras, S. Laine, and T. Aila, “A style-based generator architecture for generative adversarial networks,” in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019, pp. 4401–4410.

[29] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila, “Analyzing and improving the image quality of stylegan,” in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2020, pp. 8110–8119.

[30] N. Ruiz, Y. Li, V. Jampani, Y. Pritch, M. Rubinstein, and K. Aberman, “Dreambooth: Fine tuning text-to-image diffusion models for subject-driven generation,” arXiv preprint arXiv:2208.12242, 2022.

Reference

[31] V. Gallego, “Personalizing text-to-image generation via aesthetic gradients,” arXiv preprint arXiv:2209.12330, 2022.

[32] GIGAZINE. (2022). Stable Diffusion Web UI that allows you to easily generate images from text. [online] Available at: https://gigazine.net/gsc_news/en/20220904-stable-diffusion-webui/ [Accessed 24 March 2023].

[33] J. Ho, C. Saharia, W. Chan, D. J. Fleet, M. Norouzi, and T. Salimans, “Cascaded diffusion models for high fidelity image generation.” J. Mach. Learn. Res., vol. 23, no. 47, pp. 1–33, 2022.

[34] R. Rombach, A. Blattmann, D. Lorenz, P. Esser and B. Ommer, “High-Resolution Image Synthesis With Latent Diffusion Models,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2022, pp. 10684-10695.

[35] X. Pan, P. Qin, Y. Li, H. Xue and W. Chen, “Synthesizing Coherent Story with Auto-Regressive Latent Diffusion Models,” arXiv preprint arXiv:2211.10950, 2022

[36] Brooks, T., Holynski, A., & Efros, A. A. (2022). Instructpix2pix: Learning to follow image editing instructions. arXiv preprint arXiv:2211.09800.

[37] S. Behdenna, F. Barigou, and G. Belalem, “Document level sentiment analysis: a survey,” EAI Endorsed Transactions on Context-aware Systems and Applications, vol. 4, no. 13, pp. e2–e2, 2018.

[38] X. Han and L. Wang, “A novel document-level relation extraction method based on bert and entity information,” IEEE Access, vol. 8, pp. 96 912–96 919, 2020.

[39] G. Kim, H. Shim, H. Kim, Y. Choi, J. Kim, and E. Yang, “Diffusion video autoencoders: Toward temporally consistent face video editing via disentangled video encoding,” arXiv preprint arXiv:2212.02802, 2022.

[40] B. Kim, I. Han, and J. C. Ye, “Diffusemorph: Unsupervised deformable image registration along continuous trajectory using diffusion models,” arXiv preprint arXiv:2112.05149, 2021.36

[41] “Corridor Crew. (2023, 03 07). VFX Reveal Before & After - Anime Rock, Paper, Scissors [Video]. YouTube. https://www.youtube.com/watch?v=ljBSmQdL_Ow”

[42] T. Zhang, V. Kishore, F. Wu, K. Q. Weinberger, and Y. Artzi, “Bertscore: Evaluating text generation with bert,” arXiv preprint arXiv:1904.09675, 2019.

[43] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” IEEE transactions on image processing, vol. 13, no. 4, pp. 600–612, 2004.

[44] A. Cherian and A. Sullivan, “Sem-gan: Semantically-consistent image-to-image translation,” in 2019 ieee winter conference on applications of computer vision (wacv). IEEE, 2019, pp. 1797–1806.

[45] Wizards of the Coast, (2023, March 28). LOST MINE OF PHANDELVER. D&D Beyond. [Online]. Available: https://www.dndbeyond.com/sources/lmop

Thank You