Sprite Sheet Generation using a Diffusion Model

Hei Lit Wong, Marco Volino
{hw01605, m.volino}@surrey.ac.uk
University of Surrey

Motivation

- Sprite sheets contain a series of frames to form 2D animations in games, which require extensive time and labour for manual creation.
- Image generation models perform badly in sprite sheet creation due to the lack of temporal coherence.
- A significant gap exists in generating game characters directly from realistic human portraits despite the introduction of the Sprite Sheet Diffusion model [2].

Contribution

- This paper presents an end-to-end pipeline for automating the creation of 2D pixel-art sprite sheet animations from realistic human portraits, which adapts the AnimateAnyone [1] architectures with Stable Diffusion v1.5 [3] pretrained weights and fine-tuned by a custom-built dataset.
- The paper proves that the optimal workflow is to first generate animations from realistic portraits and apply the pixel-art stylisation afterwards.

Methodology

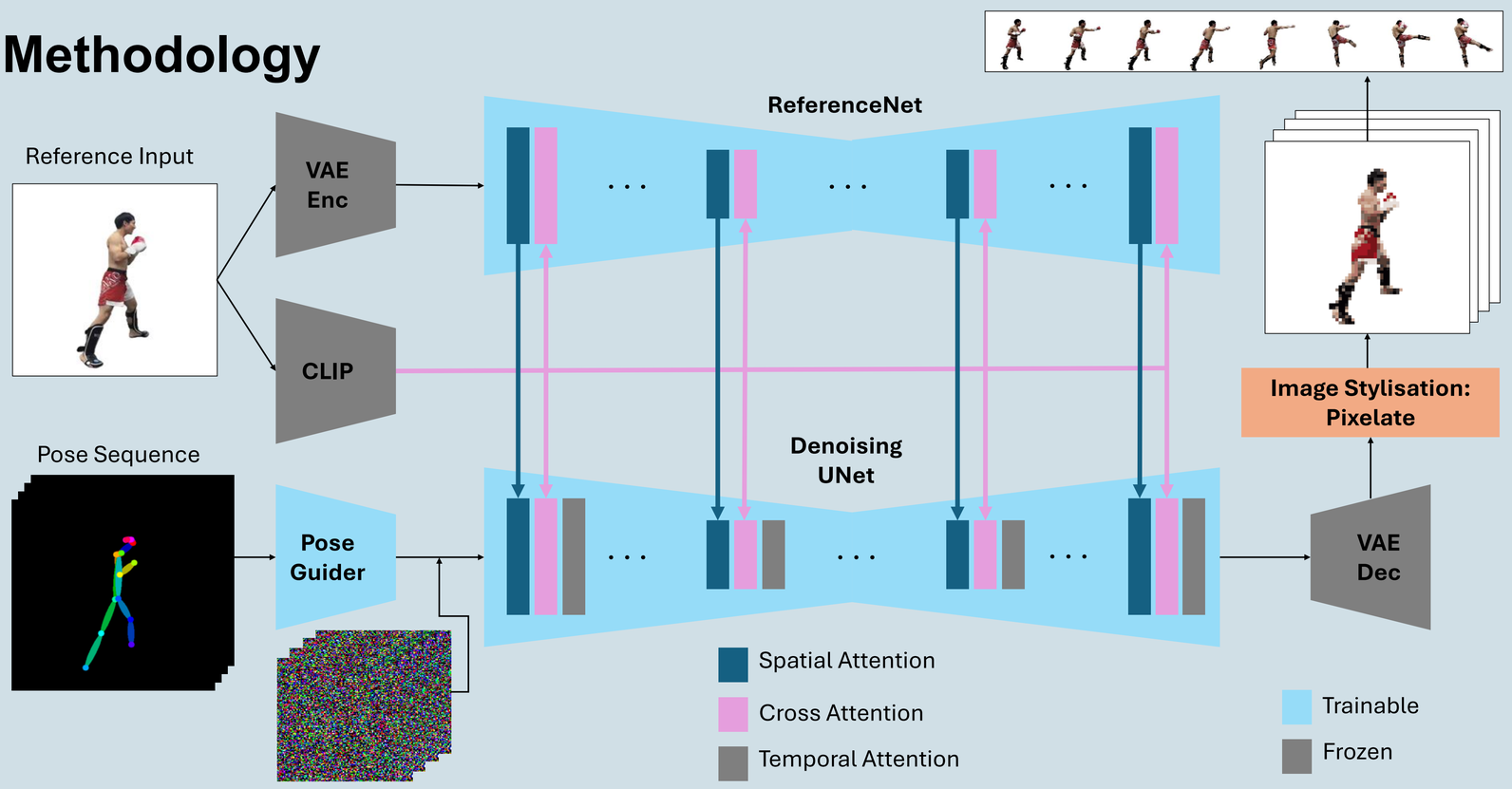

Figure 1: Model Architecture

Ablation Study & Results

Table 1: Performance against baseline model

Model	SSIM ↑	PSNR ↑	LPIPS ↓	SC ↑
Sprite Sheet Diffusion [2]	0.6544	15.7420	0.2103	0.8547
Approach A	**0.7450**	**24.8354**	**0.0366**	**0.8716**
Approach B	0.7414	22.5298	0.0642	0.8167

Approach A: Applying the pixelate stylisation post-generation
Approach B: Generate animations directly on pixelated portraits

- Both fine-tuned approach yields superior results than the baseline Sprite Sheet Diffusion [2] model
- The approach of applying pixelate stylisation after animation generation achieved the best result in all four SSIM, PSNR, LPIPS and Subject Consistency (SC) metrics.

Output Sample

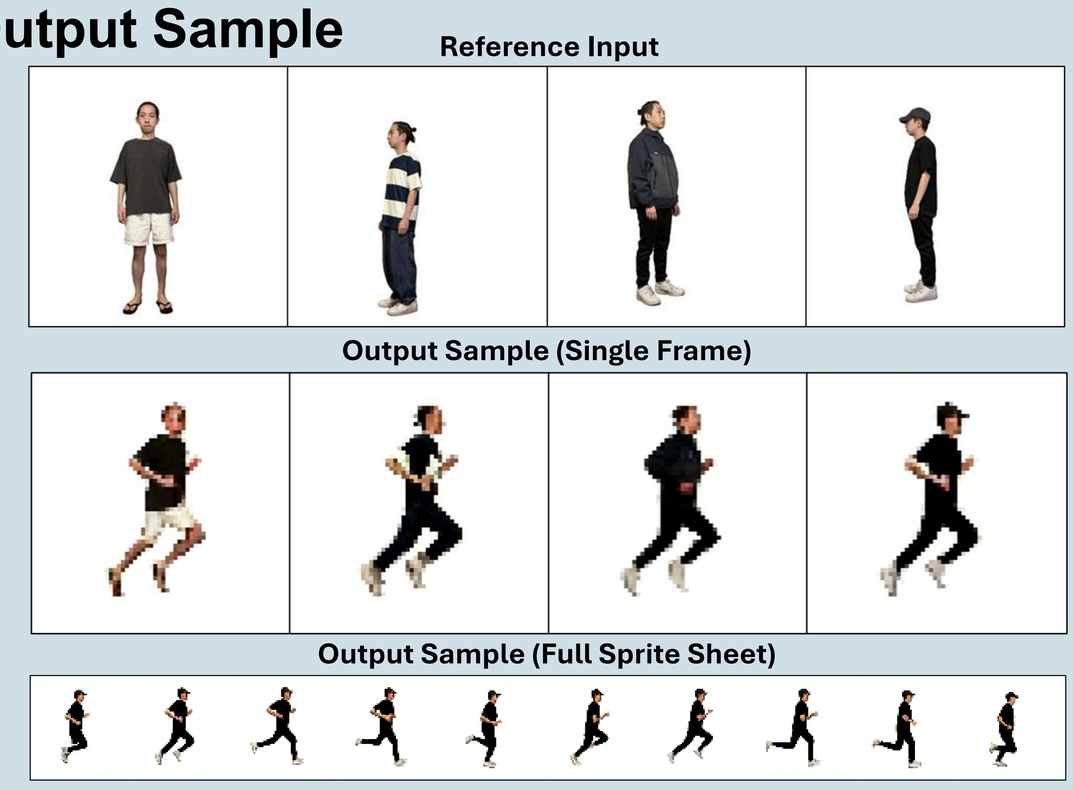

Figure 3: Generalisation test with personal portraits in different angles and outfits

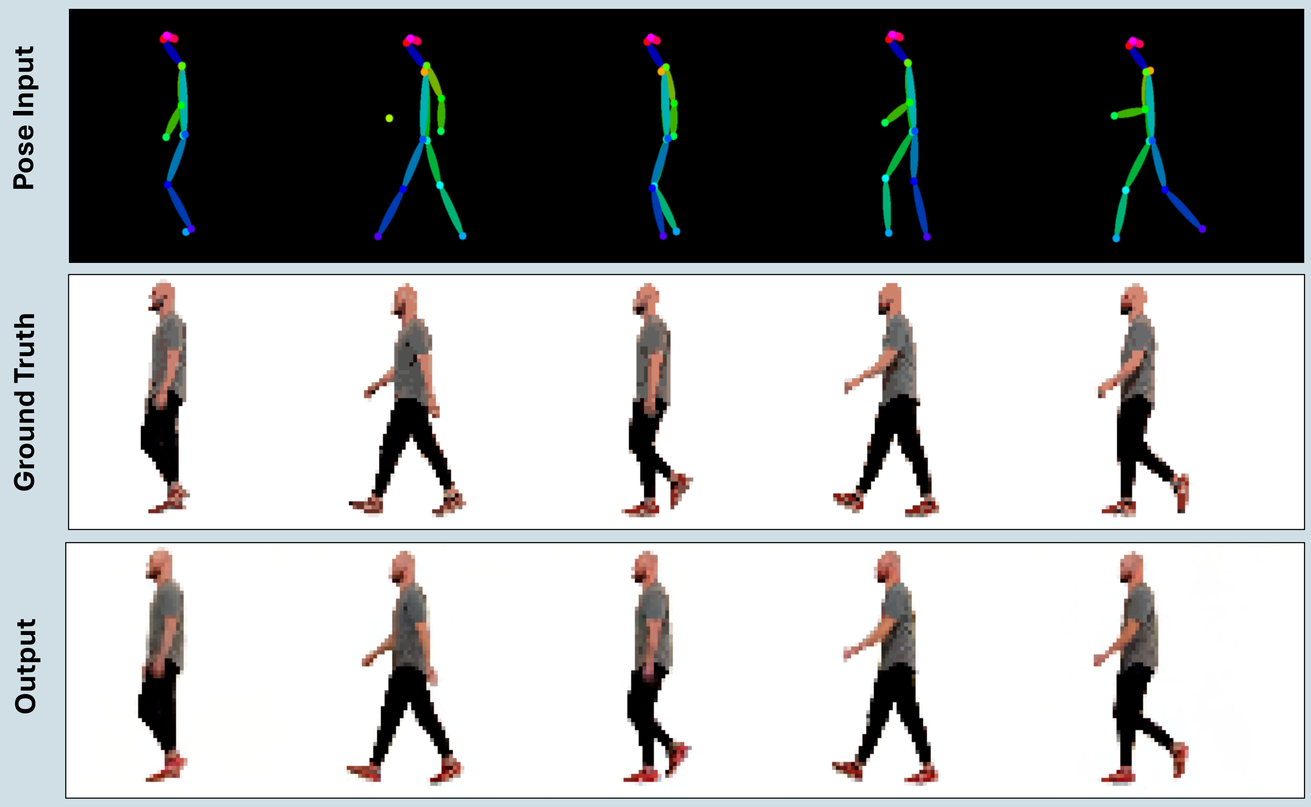

Figure 4: Model output sample for testing data with related pose input and GT

Conclusions

- This work successfully developed and validated an end-to-end pipeline to transform realistic human portraits into 2D pixel-art sprite sheet animations using Generative AI.
- The model is proven to generalise well across different camera angles and outfits that showcase practical utility.
- The proposed method tackles a time-consuming problem in sprite sheet creation, enabling rapid and personalised creation of avatars for use in games and the metaverse.

References

[1] Li Hu, Xin Gu, Zhedong Chen, Wei Huang, Sanyuan Liu, Yujun Wang, Long Chen, Yao Wang, Yiliang Wang, and Lu Wang. Animate anyone: Consistent and controllable image-to-video synthesis for character animation. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 8153–8163, 2024.
[2] Chien-An Hsieh, Jialu Zhang, and Alice Yan. Sprite sheet diffusion: Generate game character for animation. arXiv preprint, 2025.
[3] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, High-resolution image synthesis with latent diffusion models. pp. 10 684–10 695, 2022.