AI-driven automatic generation and rendering of game characters

Yaopeng Xie

Faculty of Science, Western University, ON, Canada

yxie447@uwo.ca

Abstract. This paper provides a comprehensive review of AI-driven techniques for the automatic generation and rendering of game characters, with a particular focus on Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs). GANs, using adversarial training approaches, have revolutionized character development by producing incredibly realistic and aesthetically pleasing gaming characters. VAEs, despite frequently encountering issues like image blurriness, provide an alternate strategy that emphasizes diversity and originality in the generated content. Additionally, conditional models that enable more individualized and regulated character generation are explored, as well as hybrid models that combine the best features of both GANs and VAEs. The difficulties with mode collapse in GANs and the requirement for big datasets for both GANs and VAEs are also covered, along with some possible fixes like transfer learning and semi-supervised learning strategies. This analysis emphasizes the growing significance of AI-driven game character generation in the gaming industry by highlighting its current state, problems, and future directions.

Keywords: Game, generative adversarial networks, variational autoencoders.

1. Introduction

The creative industries have been greatly impacted by the rapid growth of artificial intelligence (AI), with the gaming industry being one of the most severely affected. The autonomous creation and rendering of game characters stands out as a revolutionary advancement among the many uses of AI in gaming. Historically, game character development has been a labor-intensive process, requiring experienced designers to create distinctive and visually captivating characters. To get the right amount of realism and detail, this procedure frequently required several iterations and adjustments. But the paradigm for character development has completely changed with the advent of AI-driven methods like Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs).

Since its introduction by Goodfellow et al. GANs have become a potent tool in generative modeling, especially for tasks that need a high degree of detail and realism [1]. GANs have proven highly effective in creating lifelike graphics through the adversarial interaction of their discriminator and generator components. For this reason, they are perfect for creating game characters that closely resemble real people or imaginary animals. This feature has given game creators more creative options in addition to cutting down on the time and expense of character production.

However, VAEs-which were initially put forth by Kingma and Welling-offer an alternative strategy to generative modeling [2]. VAEs are very useful for investigating the wide range of potential character

designs, which enables the construction of a variety of characters that can be highly stylized or realistic. VAEs have the capacity to encode and decode intricate data distributions. This ability enables the creation of distinctive and flexible characters, offering a crucial degree of artistic freedom in the current game production industry. Additionally, the use of hybrid models that combine VAEs and GANs has demonstrated promise in overcoming the drawbacks of each technique, resulting in character generating systems that are more reliable and adaptable.

This paper provides a thorough analysis of AI-driven game character generation. It examines the uses of GANs and VAEs, their challenges, and future directions in this rapidly evolving field. By understanding these dynamics, academics and game developers can better harness AI's potential to innovate in game character creation.

2. Generative Adversarial Networks (GANs)

Generative Adversarial Networks (GANs), introduced by Goodfellow et al. have transformed generative modeling in many fields, including game character generation. Two competing neural networks make up GANs: the discriminator (D), which assesses the new data instances, and the generator (G), which generates them. This procedure can be compared to the efforts of a counterfeiter (the generator) to create counterfeit money and the police (the discriminator) to find it. As both the discriminator and generator improve over time, they collectively produce increasingly lifelike artificial data [1].

GANs have undergone numerous iterations and improvements to overcome various challenges and enhance their effectiveness. The original GAN framework, known as the classical GAN, suffered from issues such as mode collapse and training instability, but also laid the foundation for adversarial training. It used basic multi-layer perceptrons as both the discriminator and the generator [3]. Deep convolutional GAN (DCGAN) was proposed by Radford et al. DCGAN incorporates convolutional layers into both the generator and the discriminator. This significantly improves the quality and stability of the generated images, making them more suitable for high-resolution tasks [4]. Arjovsky et al. proposed the Wasserstein distance as a new loss function to overcome the instability in GAN training, which is known as the Wasserstein GAN (WGAN). By addressing the mode collapse issue, WGAN stabilizes the training process and produces more realistic images [5]. Spectral normalization or spectral normalized GAN (SNGAN) was proposed by Miyato et al. It stabilizes the training of the discriminator by normalizing the spectral norm of the weight matrix. By using this approach, the training process becomes more persistent and reliable [6]. With the introduction of BigGAN by Brock et al., larger batch sizes and deeper networks were added to the GAN architecture. This allows for the generation of high-fidelity graphics with more variation and complexity, making it ideal for developing complex game characters [7]. The automatic creation of characters for role-playing games and Dungeons & Dragons is a notable use of GANs in gaming. GANs were used by Aydin Emekligil and Öksüz to automate the process of creating visual characters for video games.

Figure 1. Examples of two datasets including original game characters [3]

They used RPG and DnD datasets for training, each containing thousands of internet-sourced photos (Figure 1). With these datasets, they trained six distinct GAN models. SNGAN performed the best out of all the models that were tested. The study also emphasized the benefits of transfer learning strategies, such BigGAN and WGAN-GP, especially in situations when data availability is constrained. These models demonstrated the effectiveness of transfer learning in situations with limited datasets, outperforming those trained from scratch [3]. In another study, Shi et al. created a technique for employing GANs to convert facial characteristics to gaming character qualities. Based on input parameters, this technique, called Face-to-Parameter Translation, enables the production of realistic and varied character faces.
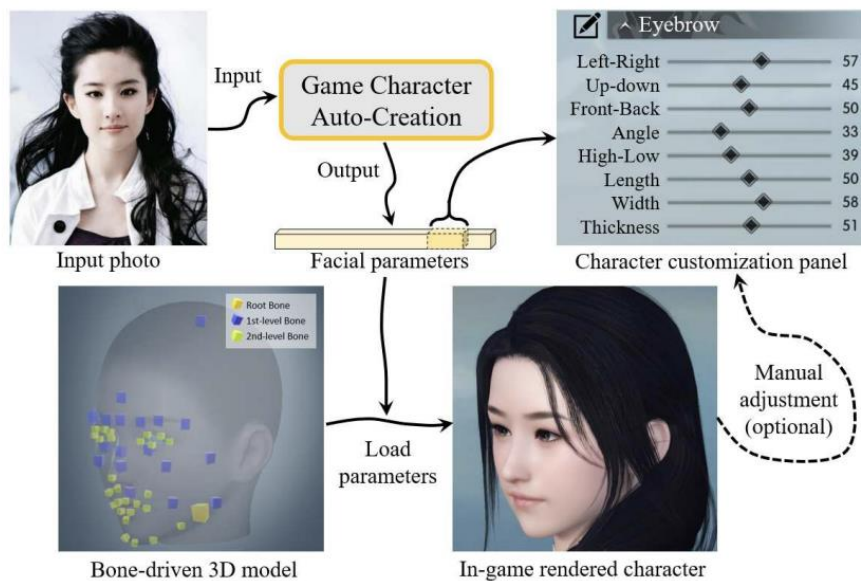
Figure 2. Face-to-Parameter Translation technique [8].

By training on a collection of character faces, the GAN learns to map latent variables to specific facial features. This provides game creators with a powerful tool for character generation, as shown in Figure 2 [8]. By creating Zero-Shot Text-to-Parameter Translation for gaming character auto-creation, Zhao et al. expanded on the idea even further. This method makes it possible to create game characters using textual descriptions, making the process of creating characters more flexible and intuitive. Based on user input, the model is trained to translate textual descriptions into character parameters, generating a varied range of high-quality characters [9].

3. Variational Autoencoders (VAEs)

Apart from GANs, another generative model that has been employed in game character creation is the variational autoencoder (VAE). In order for VAEs—which Kingma and Welling first proposed—to function, input data must first be encoded into a latent space and then decoded back into the original data space. With the help of this method, VAEs can produce new samples that closely resemble the original dataset and discover the underlying distribution of the data [2].

The creation of Pokémon characters is an intriguing use case for VAEs. Pokémon characters have been created by researchers using a Convolutional VAE (CVAE) based on their type. The model was able to create new characters that matched the attributes and style of the original dataset by training the CVAE on a collection of photos of Pokémon. This application demonstrates how VAEs may be used to generate a wide range of game character styles, which makes them an invaluable tool for game designers [10]. VAEs have been extensively used for character face generation and customization. By training on large datasets of human faces, VAEs can learn complex facial trait distributions to create new, realistic

faces. VAEs have been used, for instance, to produce intricate and programmable video game avatars. These models enable users to modify various aspects of generated faces, such as age, gender, and expression, offering a highly customized gaming experience [2][11].

VAEs have been used to create a vast range of gaming assets, such as environments, objects, and textures, in addition to characters. Emekligil demonstrated the use of VAEs in creating game environments. Their approach produces environments that complement the game's overall aesthetic in a cohesive and artistic manner. This program is especially helpful for procedurally generated games, where it's essential to automatically create interesting and varied content [11].

To improve their capabilities, VAEs have been coupled with other generative models in addition to being used alone. Combining VAEs with GANs or other deep learning methods allows hybrid models to take advantage of the advantages of each method. For example, Combining GANs' ability to generate high-quality images with VAEs' encoding of complex data distributions can result in more robust and adaptable models. These hybrid models work especially well for producing varied and high-resolution game characters and objects [1][2]. Many improvements have been suggested to alleviate some of the drawbacks of VAEs, such as the production of hazy images. Methods like conditional VAEs (cVAEs) and hierarchical VAEs have demonstrated promise in raising the caliber and precision of generated information. Multiple levels of latent variables are introduced by hierarchical VAEs, enabling more intricate and organized representations. Conversely, conditional VAEs produce samples according to particular features or situations, allowing for more focused and regulated creation of game characters and materials [2][11].

4. Comparative Analysis of GANs and VAEs

When it comes to creating game characters, both GANs and VAEs have advantages and disadvantages that are specific to them. Although GANs are renowned for generating realistic, high-quality images, they frequently have problems like mode collapse, in which the generator only generates a small number of output variations. On the other hand, because VAEs are probabilistic in nature, they can result in fuzzy images even if they often yield more diverse outputs [1][8]. The need for large datasets is one of the main obstacles to applying these models. Extensive training data is usually required for high-quality character production; however, this data is not always available. Transfer learning becomes essential at this point. Even with insufficient data, high-quality results can be obtained by fine-tuning models that have been pre-trained on larger datasets. Research has demonstrated that transfer learning greatly improves the performance of VAEs and GANs in situations where there is limited data availability [1][3][8][9].

Scholars have begun investigating hybrid models that incorporate features of both GANs and VAEs in order to take advantage of their respective strengths. The goal of these models is to produce high-quality images that have the diversity that VAE outputs are known for. One strategy is to apply adversarial training techniques from GANs to enhance the realism of the generated images while encoding and decoding data using the VAE architecture. This combination may be able to get beyond the drawbacks of both models and offer a reliable method for creating gaming characters [1][8][11].

The creation of conditional models, such as Conditional VAEs (cVAEs) and Conditional GANs (cGANs), is another exciting avenue. Character development can be more precisely controlled and tailored with the help of these models, which can produce characters depending on particular traits or conditions. To provide game producers more control over the generated content, a cGAN might be trained to create characters with particular hair colors, outfit designs, or facial traits [1][8][9].

The creation of gaming characters is likely to benefit from future developments in unsupervised and semi-supervised learning methodologies. These methods can lessen the reliance on substantial labeled datasets, making it possible to produce high-quality characters with little oversight. These techniques can improve the scalability and efficiency of the character generation process by utilizing unlabeled data [1][2][11]. Maintaining a balance between diversity and quality in generated characters remains a significant challenge. Finding the ideal balance between the trade-offs of GANs and VAEs is essential. To solve this problem, hybrid models and enhanced training strategies like hierarchical VAEs and

progressive GANs are being investigated [1][2][11]. GAN and VAE training requires substantial computational resources, including high-performance GPUs and large memory capacities. For freelance creators and smaller studios, this can be a hindrance. These issues can be lessened by cloud-based solutions and optimization strategies like model reduction and quantization, which provide accessibility to sophisticated generative models [1][2][11].

5. Conclusion

The creation of game characters has been completely transformed by AI-driven methods, especially those that make use of GANs and VAEs. These methods offer previously unheard-of levels of realism, diversity, and efficiency. Although GANs are excellent at creating incredibly complex and lifelike characters, they can have issues with mode collapse, which can limit the range of characters that can be produced. VAEs present a new set of issues because, although they enable a greater variation in character generation, they frequently yield images that lack crispness and detail. A promising approach to getting over these restrictions is the construction of hybrid models that combine the best features of VAEs and GANs, allowing for the production of diverse and aesthetically pleasing characters. Furthermore, game developers can create characters that adhere to particular design specifications by using conditional models, such as Conditional GANs (cGANs) and Conditional VAEs (cVAEs), which afford them additional control over the character generation process. These developments raise the bar for produced characters while also streamlining the character design process and increasing accessibility for independent and smaller teams. The quality, diversity, and control of created characters in video games will probably continue to increase in the future as AI technology develops. The use of big labeled datasets will be lessened by unsupervised and semi-supervised learning approaches, making high-quality character generation possible even with sparse data. The future of the gaming business will be greatly influenced by the further advancement of these AI-driven techniques, which will make it possible to create more immersive and captivating gaming experiences.

References

[1] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Sherjil O, Aaron C and Yoshua B Generative adversarial nets 2014 Adv. Neural Inf. Process. Syst. 27 2672-80

[2] Shi T, Zuo Z, Yuan Y and Fan C Fast and robust face-to-parameter translation for game character auto-creation 2020 Proc. AAAI Conf. Artif. Intell. 34(02) 1733-1740

[3] Emekligil F G A and Öksüz İ Game character generation with generative adversarial networks 2022 In: 30th Signal Processing and Communications Applications Conference (SIU) 1-4

[4] Radford A, Metz L and Chintala S Unsupervised representation learning with deep convolutional generative adversarial networks 2016 arXiv:1511.06434

[5] Gulrajani I, Ahmed F, Arjovsky M, Dumoulin V and Courville A C Improved training of Wasserstein GANs 2017 Adv. Neural Inf. Process. Syst. 30 5769-79

[6] Miyato T, Kataoka T, Koyama M and Yoshida Y Spectral normalization for generative adversarial networks 2018 arXiv:1802.05957

[7] Brock A, Donahue J and Simonyan K Large scale GAN training for high fidelity natural image synthesis 2019 arXiv:1809.11096

[8] Shi T, Yuan Y, Fan C, Zou Z, Shi Z and Liu Y Face-to-parameter translation for game character auto-creation 2019 In: IEEE/CVF International Conference on Computer Vision (ICCV) 161-170

[9] Zhao R, Li W, Hu Z, Li L, Zou Z, Shi Z and Fan C Zero-shot text-to-parameter translation for game character auto-creation 2023 In: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 21013-21023

[10] Wang S, Zeng W, Wang X, Yang H, Chen L, Yuan Y, Zeng Y, Zheng M, Zhang C and Wu M SwiftAvatar: Efficient auto-creation of parameterized stylized character on arbitrary avatar engines 2023 arXiv:2301.08153

[11] Emekligil F G A Generative models for game character generation 2023