Comparing Occupational Gender Bias in AI-Generated Anime-style and Realistic Illustrations

Ayano Ito Shoma Mizuno Keiichi Namikoshi Yuko Sakurai
Nagoya Institute of Technology

Satoshi Oyama
Nagoya City University

Abstract

The rapid advancement of generative AI has enabled the creation of both realistic and anime-style illustrations across diverse domains. However, such systems risk reproducing and amplifying cultural and social stereotypes embedded in their training data. This study presents a comparative analysis of occupational gender bias in AI-generated anime-style and realistic illustrations. Using Stable Diffusion 3.5, Illustrious, ChatGPT-5, and Gemini 2.5 Pro, we generated images for ten occupations, with each prompt depicting ten individuals. Our analysis examined gender distributions and expressive features such as crossed arms. Results show that anime-style illustrations amplify gender imbalances and cultural stereotypes more distinctly than realistic images. For instance, in the case of nurses, female representations are particularly emphasized. Across occupations, anime-style generation is also strongly shaped by conventions characteristic of Japanese anime. These findings underscore the critical role of stylistic domains in the reproduction of cultural conventions in generative AI and highlight the importance of evaluating bias across different visual styles.

1 Introduction

Generative AI has rapidly expanded its applications from natural language processing to image generation, attracting significant attention both socially and technologically. In particular, the automatic generation of anime-style illustrations has been widely adopted in fields such as entertainment, education, and advertising, giving rise to new forms of creative activity. However, studies have shown that generative AI can reflect, and sometimes reinforce, the social and cultural biases embedded in its training data. The possibility that generative AI could reproduce biases and inequality observed in the real world therefore poses a critical societal issue.

Existing research has addressed this issue from multiple perspectives. The seminal study by Caliskan et al. [1] demonstrated that semantic representations automatically derived from large-scale text corpora reflect human-like biases. This was the first large-scale study to empirically show that AI systems trained on human-generated data tend to inherit cultural stereotypes, thereby laying the theoretical foundation for subsequent research on algorithmic bias. In recent years, research focusing on generative AI has progressed rapidly. Zhou et al. [6] systematically organized biases in generative models for images, classifying them according to dimensions such as gender, race, and occupation. They showed that factors like data imbalance, model design, and user prompts interact in complex ways, leading generative systems not only to reproduce but also to reinforce social stereotypes. From a methodological standpoint, OpenBias proposed by D'Incà et al. [2] is a novel framework for bias detection in text-to-image generative models. This approach systematically

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: The First Workshop on Generative and Protective AI for Content Creation.

Table 1: Comparison of average number of generated individuals per image by occupation (Stable Diffusion vs. Illustrious, Anime-style vs. Realistic)

Occupation	Stable Diffusion						Illustrious					
	Anime-style			Realistic			Anime-style			Realistic		
	Male	Female	Arms	Male	Female	Arms	Male	Female	Arms	Male	Female	Arms
Childcare worker	0.5	5.0	0.5	2.4	5.0	0.2	0.0	1.1	0.0	0.1	1.1	0.0
Cook	0.4	4.0	0.3	3.1	1.1	2.0	0.2	0.9	0.1	0.0	1.1	0.2
Dancer	0.0	6.4	0.0	0.0	6.5	0.0	0.3	0.8	0.0	0.0	1.1	0.0
Doctor	2.4	3.3	1.0	4.2	1.0	3.1	0.3	0.2	0.1	0.0	0.9	0.0
Hairdresser	0.0	4.8	0.0	0.2	3.5	0.2	0.2	0.5	0.1	0.0	0.9	0.0
Maid	0.0	5.4	0.4	0.0	5.5	0.0	0.2	1.0	0.1	0.0	0.9	0.0
Nurse	0.0	5.3	0.0	0.4	5.2	2.2	0.0	0.9	0.1	0.0	1.1	0.1
Pilot	2.6	2.0	0.0	5.1	0.2	0.2	0.1	0.7	0.0	0.0	0.8	0.0
Police	0.0	3.3	0.1	1.7	0.1	0.3	0.7	0.3	0.1	0.2	1.0	0.2
Teacher	0.0	5.0	0.1	1.0	6.0	1.5	0.4	0.7	0.0	0.0	0.8	0.0

examines model outputs, enabling the detection of newly emerging or previously unanticipated forms of bias. In parallel, training data itself has been shown to play a crucial role in the formation of bias within generative models. Zeng et al. [5] analyzed widely used datasets and revealed the existence of imbalances in gender and occupational distribution. These findings indicated that bias in generative AI is strongly linked to the characteristics of the data used for training. Furthermore, in recent years, fairness-oriented generative modeling methods have been proposed to mitigate such issues. Li et al. [3] introduced fair mapping in text-to-image generation based on diffusion models, presenting a method that maintains demographic balance in generated outputs while preserving quality. Similarly, Xu et al. [4] proposed a multi-objective optimization approach within multi-agent environments aimed at reducing social bias in LLMs.

While these studies demonstrate the persistence of bias in generative AI, systematic analysis remains limited in specific stylistic domains such as anime. Anime-style illustrations, in contrast to realistic depictions of the real world, are heavily influenced by culturally and historically shaped visual conventions, making them more likely to reproduce gender and occupational stereotypes.

Therefore, this study conducts a comparative analysis between anime-style and realistic illustrations across multiple generative models. Specifically, Stable Diffusion 3.5 medium and the anime-specialized model Illustrious XL v2.0-STABLE are employed as text-to-image generative models, while ChatGPT-5 and Gemini 2.5 Pro are used as large language models. For each model, a specific occupation is designated, and ten individuals are generated within a single image, enabling comparisons across ten professions, including nurses and police officers. To capture the diversity inherent in each model, we adopt the setting of generating ten individuals per image. In addition to analyzing the gender distribution of the individuals depicted, we focus on the presence of crossed arms as an expressive feature, examining whether this expression reflects occupational context or stems from the model itself.

The results of this study suggest that anime-style illustrations produced by generative AI tend to amplify gender imbalances and cultural stereotypes more distinctly than realistic images. For example, in the case of nurses, female representations are particularly emphasized. Moreover, across various occupations, anime-style generation is strongly influenced by conventions characteristic of Japanese anime.

2 Text-to-Image Generative Models

In this section, we present the results from Stable Diffusion and Illustrious. For each of the ten occupations, we generated ten realistic-style digital illustrations and ten anime-style illustrations with both models. Table 1 summarizes the average counts of males, females, and individuals with crossed arms. As illustrative examples of characteristic outputs, Figure 1 shows images for hairdressers, nurse, and police. The images are arranged from left to right in the order of Stable Diffusion, Illustrious, Gemini, and ChatGPT. We present the images of the remaining occupations in the Appendix.

In Stable Diffusion, the anime style tends to increase the representation of females and younger individuals, typically in their twenties. The model often generates multiple figures with similar facial

Table 2: Comparison of average number of generated individuals per image by occupation (Gemini vs. chatGPT, Anime-style vs. Realistic)

Occupation	ChatGPT						Gemini					
	Anime-style			Realistic			Anime-style			Realistic		
	Male	Female	Arms	Male	Female	Arms	Male	Female	Arms	Male	Female	Arms
Childcare worker	3.5	6.9	0.0	2.5	7.6	0.2	7.0	8.1	0.1	4.1	7.8	0.5
Cook	8.3	2.0	2.3	7.4	3.1	2.9	7.0	4.7	0.0	5.9	5.0	0.3
Dancer	4.1	6.8	0.0	4.7	6.3	0.0	3.9	6.9	0.0	4.4	5.8	0.0
Doctor	6.2	3.8	0.9	6.0	4.0	1.9	7.4	5.1	1.3	5.0	4.7	1.0
Hairdresser	3.8	5.4	0.0	4.0	6.2	0.5	3.6	9.3	0.1	5.3	5.9	0.1
Maid	0.0	10.3	0.0	0.1	10.0	0.0	0.0	11.4	0.4	0.0	10.6	1.0
Nurse	0.0	10.5	0.1	1.7	8.5	0.1	1.8	11.7	0.6	2.9	8.1	1.6
Pilot	7.2	3.0	0.2	9.0	1.8	0.0	5.7	6.8	3.1	7.9	2.5	0.6
Police	7.7	3.2	0.3	7.8	2.6	0.0	6.4	5.6	2.4	7.7	2.8	0.9
Teacher	4.7	4.3	1.0	5.0	5.4	0.5	4.9	5.9	0.9	4.5	5.8	1.1

features and poses within a single image. In Illustrious, nearly all generated individuals are female, regardless of occupation, indicating a potential bias in the training data. Despite prompts specifying ten people, the model often produces only one figure, and occupational markers, such as uniforms or props, are less evident compared to other models. Nevertheless, Illustrious exhibits greater diversity in hair and eye colors, and its overall style is the most reminiscent of Japanese anime among the four models.

3 LLM Models

In this section, we present the results for the images generated by the two LLMs. The results are presented in Table 2. Both LLMs failed to consistently include ten individuals within a single image.

In Gemini, there is a strong tendency to depict occupational characteristics such as showing a cook engaged in cooking. Compared with other models, gender imbalance is less pronounced, and both realistic and anime styles generate males and females in a more balanced manner. In particular,, in the anime version of nurse, Gemini was the only model among the four to generate male figures. Furthermore, in nurse and police, the proportion of females increased in the anime style, and overall, there was a tendency to generate younger individuals, typically in their twenties. In contrast, in ChatGPT, the generated images often feature multiple individuals arranged in one to three rows. In the case of the police, the proportion of females increased in the anime style. Although these characters tend to share similar eye shapes, individual differences are expressed through features such as hair, mouth, and color.

Furthermore, In line with existing works reporting that LLMs can mitigate bias through self-reflection, we asked ChatGPT whether it perceived the generated images of nurses to exhibit bias. The response was as follows:

- This style reflects Japanese anime aesthetics (large eyes, soft facial features, small face), which may emphasize an idealized femininity.
- This may unintentionally amplify gendered expectations in caregiving roles.

This response suggests that the training dataset for anime images likely contains a large number of Japanese anime works, thereby reflecting cultural biases specific to Japan.

4 Conclusion

In this study, we conducted a comparative analysis of occupational gender bias in realistic images and anime-style illustrations generated by state-of-the-art models, including Stable Diffusion, Illustrious, Gemini, and ChatGPT. Our analysis demonstrates that bias in generative AI is not merely a technical issue but is deeply intertwined with training data and cultural contexts. Future research directions include comparing generative outputs across different cultural settings and extending bias detection methodologies to multimodal environments.

(a) Anime-Style Illustration of Hairdressers

(b) Realistic-Style Digital Illustration of Hairdressers

(c) Anime-Style Illustration of Nurses

(d) Realistic-Style Digital Illustration of Nurses

(e) Anime-Style Illustration of Polices

(f) Realistic-Style Digital Illustration of Polices

Figure 1: Generated images (from left to right: Stable Diffusion, Illustrious, Gemini, and ChatGPT)

References

[1] A. Caliskan, J. J. Bryson, and A. Narayanan. Semantics derived automatically from language corpora contain human-like biases. *Science*, 356(6334):183–186, 2017.

[2] M. D'Incà, E. Peruzzo, M. Mancini, D. Xu, V. Goel, X. Xu, Z. Wang, H. Shi, and N. Sebe. OpenBias: Open-set bias detection in text-to-image generative models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (CVPR-2024), pages 12225–12235, 2024.

[3] J. Li, L. Hu, J. Zhang, T. Zheng, H. Zhang, and D. Wang. Fair text-to-image diffusion via fair mapping. *Proceedings of the 39th AAAI Conference on Artificial Intelligence* (AAAI-2025), 39 (25):26256–26264, 2025.

[4] Z. Xu, W. Chen, Y. Tang, X. Li, C. Hu, Z. Chu, K. Ren, Z. Zheng, and Z. Lu. Mitigating social bias in large language models: A multi-objective approach within a multi-agent framework. In *Proceedings of the 39th AAAI Conference on Artificial Intelligence* (AAAI-2025), volume 39, 2025.

[5] B. Zeng, Y. Yin, and Z. Liu. Understanding bias in large-scale visual datasets. *Advances in Neural Information Processing Systems*, 37:61839–61871, 2024.

[6] M. Zhou, V. Abhishek, T. Derdenger, J. Kim, and K. Srinivasan. Bias in generative AI. *arXiv preprint arXiv:2403.02726*, 2024.

Appendix

In the Appendix, we provide a more detailed discussion along with additional generated images for the remaining occupations.

Stable Diffusion In Stable Diffusion, the anime style tends to increase the proportion of female figures across several occupations (e.g., childcare worker, cook, doctor, hairdresser, nurse, pilot, police, and teacher). Another notable characteristic is the frequent repetition of similar facial features and poses within a single image. Regardless of style, childcare workers are often depicted as children of preschool or elementary school age, while teachers are frequently portrayed as middle or high school students. This suggests either a strong sensitivity to the prompt term "child" or that teachers and students are represented as closely related in the feature space. Furthermore, anime-style outputs tend to depict young adults, typically in their twenties.

Illustrious In Illustrious, nearly all generated figures were female, regardless of occupation, suggesting that the training data may have been disproportionately female-centered. Moreover, despite prompts specifying ten people, the model often generated only a single figure. Compared with other models, occupations other than nurse, maid, and police lacked clear occupational markers such as uniforms or props, making them more difficult to identify. Moreover, despite using occupation such as cook and hairdresser, some of the generated images appeared to resemble maids. By contrast, Illustrious exhibited greater variation in hair and eye colors and, overall, produced images most closely resembling Japanese anime among the four models.

Gemini Gemini showed a strong tendency to depict occupationally relevant actions, such as portraying cooks engaged in cooking, rather than simply arranging figures side by side. Compared with other models, gender imbalance was less pronounced, and both realistic and anime styles generated males and females in a more balanced manner. Notably, in the anime version of nurse, Gemini was the only model among the four to generate male figures. Moreover, in nurse, pilot, and police, the proportion of female figures increased in the anime style, and overall, there was a stronger tendency to generate young adults, typically in their twenties.

ChatGPT In ChatGPT, the generated images often featured multiple individuals arranged in one to three rows. For pilot and police, the proportion of female figures increased in the anime style, whereas for childcare worker and cook, the proportion of male figures increased. Although many of the characters shared similar eye shapes, individual differences were expressed through other features such as hair, mouth, and color.

Specific Occupation (Pilot) For pilots, Stable Diffusion tended to generate military-style uniforms, ChatGPT depicted airline captain-style uniforms, and Gemini produced both military and captain representations. In anime style, across all models, characters were often portrayed in outfits reminiscent of robot anime, such as Evangelion or Gundam. This likely reflects the prevalence of robot anime compared to military or airline pilot portrayals in Japanese animation.

Overall Discussion Overall, adopting the anime style increased the proportion of female figures (particularly in pilot and police) and also increased the number of young adults. This tendency appears to be strongly influenced by training data. Anime works frequently feature youthful and female characters, and datasets collected from illustration-sharing platforms and other social networking services are likely skewed toward female representation.

In addition, while Stable Diffusion often generated many similar-looking figures, Gemini and ChatGPT tended to produce a broader diversity of individuals across age and gender, possibly indicating an intentional effort to represent diversity. However, in occupations such as maid, where gender is culturally fixed, the models invariably generated only female figures.

The presence of individuals with crossed arms may be intended to convey an impression of strength or authority. However, the interpretation of such gestures is itself culturally shaped, and their frequent appearance in generated images may also reflect underlying social and cultural biases in the representation of human figures.

(a) Anime-Style Illustration of Childcare Workers

(b) Realistic-Style Digital Illustration of Childcare Workers

(c) Anime-Style Illustration of Cooks

(d) Realistic-Style Digital Illustration of Cooks

(e) Anime-Style Illustration of Dancers

(f) Realistic-Style Digital Illustration of Dancers

Figure 2: Generated images (from left to right: Stable Diffusion, Illustrious, Gemini, and ChatGPT)

(a) Anime-Style Illustration of Doctors

(b) Realistic-Style Digital Illustration of Doctors

(c) Anime-Style Illustration of Maids

(d) Realistic-Style Digital Illustration of Maids

(e) Anime-Style Illustration of Pilots

(f) Realistic-Style Digital Illustration of Pilots

Figure 3: Generated images (from left to right: Stable Diffusion, Illustrious, Gemini, and ChatGPT)

(a) Anime-Style Illustration of Teachers

(b) Realistic-Style Digital Illustration of Teachers

Figure 4: Generated images (from left to right: Stable Diffusion, Illustrious, Gemini, and ChatGPT)