Journal of Student Research Volume 13 Issue 3 (2024)

Comparing Demographic Representation in AI-Generated and Stock Images of Occupations

Kabir Goklani[1] and Tom Flanagan[1#]

[1]Singapore American School, Singapore
[#]Advisor

ABSTRACT

This paper analyzes how AI-generated images from DALL·E 3 and stock images from Getty Images compare in their representation of race, gender, and age across various occupations. The researcher randomly selected ten occupations from the O*NET Database, generating 30 images from DALL·E 3 and selecting 30 images from Getty Images for each occupation. Following this, the researcher categorized each image based on its race, gender, and age. The results showed significant difference in the portrayal of race and gender between the image sources, with DALL·E 3 depicting more white male individuals than Getty Images. However, the researcher found little difference in the representation of age between DALL·E 3 and Getty Images. This study highlights the extent to which AI models like DALL·E 3 might reflect or amplify societal biases present in their training data–and what steps are needed to mitigate this issue.

Introduction

Artificial Intelligence (AI) is revolutionizing the way we interact with the world, embedding itself into daily tasks and transforming industries with unprecedented speed. AI carries out tasks that humans typically believe require cognition, such as perceiving, learning, interacting with the environment, problem-solving, and even exercising creativity (“What Is AI?”, 2023). Generative AI–the most widely-used form of AI–has various functions such as generating articles, essays, code, and images, making it one of the most powerful tools in modern technology. Software such as ChatGPT, a generative AI chatbot developed by AI research company OpenAI, boasts 180.5 million users, and in January 2024 alone, the website attracted 1.6 billion visits, making it one of the most influential technologies in the AI landscape today (Mortensen, 2024). In fact, this cutting-edge technology could potentially outpace human production capacity by 2030 (“Is Generative AI a Game Changer?”, 2024).

It is important, however, to understand and acknowledge the potential drawbacks of this technology. AI models such as ChatGPT are trained through processing input data, using algorithms to uncover patterns and correlations, and then subsequently using what they have discovered to act upon future inputs (“What Are AI Models?”, 2021). However, these models can produce biased content when the data used to train them is not representative of the reality they are meant to model. ImageNet, a 2009 training set of 14 million images, was used for more than a decade before researchers found disturbing content involving sexual images in AI-generated images, where women were clearly present (Birhane & Prabhu, 2021). Some images were labeled into problematic categories such as “spastic,” “mulatto,” or “redneck” (Crawford & Paglen, 2019). Research has also found that even ordinary prompts can produce images that reinforce harmful stereotypes related to race, gender, age, and class (Bianchi et al., 2022). For instance, when the model was prompted with a basic descriptor such as “an attractive person”, it generated images that conform to a ‘white ideal’ of attractiveness, featuring characteristics such as blue eyes, pale skin, and long straight hair. With users generating over two million images a day with DALL·E (“DALL·E Now Available without Waitlist”, 2022), an AI system developed by OpenAI that can create realistic images and art from a description in human-like language, the widespread use of such biased portrayals can greatly influence public perception by reinforcing narrow standards of various communities.

The tendency of AI to amplify biases in its generated images highlights a parallel challenge in the media, where the underrepresentation of diverse demographics continues to skew societal perceptions and reinforce stereotypes. A study conducted by researchers at Osaka University addresses societal biases present in large image-text datasets. The researchers found that all demographic attributes (i.e. age, gender, skintone, ethnicity) showed significant imbalances, with an overrepresentation of 'man' and 'lighter' skin tones (Garcia et al., 2023). The media plays an important role in how people view themselves–especially in children, who are still developing a sense of who they are. A lack of representation can lead children from minority ethnic backgrounds to feel invisible or unimportant (Braga, 2022).

This study compares the demographic representation held in various occupations from O*NET occupation database between two image sources, DALL·E 3 and Getty Images. This research will provide a comprehensive analysis of the portrayal of race, age, and gender across these platforms aiming to identify and quantify differences in representation. Hence, it is the researcher's goal to answer the research question: *How do AI-generated images from DALL·E 3 and stock images from Getty Images compare in their representation of racial diversity, gender, and age across various occupations?*

Method

This research study serves to understand how AI-generated images from DALL·E 3 and stock images from Getty Images compare in their representation of racial diversity, gender, and age across various occupations. Given the nature of this study, the researcher used quantitative analysis, where descriptive statistics analyzing the demographics for each occupation was compared between the AI-generated images and stock images to draw conclusions about training bias in AI image generation.

Inclusion Criteria

In order to determine which occupations would be used in this study, the researcher selected 10 occupations from an occupational prestige ratings database. In a research study, participants were recruited from Amazon's Mechanical Turk platform to provide ratings for specific occupations. The ratings were collected using a sliding scale from 0 to 100, with participants instructed to rate each occupation based on their opinion about its prestige (Hughes et al., 2022). Furthermore, the study utilized bootstrapping techniques to calculate standard errors and derive 95% confidence intervals around the mean prestige scores. See Appendix A for the list of occupations, ranked from least to most prestigious. Amongst the list of occupations, the researcher randomly selected five occupations from the 20 most prestigious occupations (highest rated) and five occupations from the 20 least prestigious occupations (lowest rated). This was to ensure a diverse representation of occupational prestige levels, when considering factors such as racial diversity, gender, and age. Repeated occupations were excluded to establish that no one occupation has an advantage of being randomly selected. The researcher used ChatGPT-4 to randomly select five occupations from the 20 lowest-ranked occupations and 20 highest-ranked occupations (see Appendix B and C for the code used, respectively). Table 1 depicts the randomly selected occupations, followed by the search term for each platform that was used for generating each image.

Table 1. List of randomly selected occupations, followed by DALL·E 3 and Getty Images search term

Occupation	DALL·E 3 Search Term	Getty Images Search Term
Sewer Pipe Cleaner	"Show me a photo of a Sewer Pipe Cleaner"	"Sewer Pipe Cleaner"
Food Preparation Worker	"Show me a photo of a Food Preparation Worker"	"Food Preparation Worker"
Door-to-Door Sales Worker	"Show me a photo of a Door-to-Door Sales Worker"	"Door-to-Door Sales Worker"
Dishwasher	"Show me a photo of a Dishwasher"	"Dishwasher"
Maid	"Show me a photo of a Maid"	"Maid"
Space Scientist	"Show me a photo of a Space Scientist"	"Space Scientist"
Nuclear Engineer	"Show me a photo of a Nuclear Engineer"	"Nuclear Engineer"
Judge	"Show me a photo of a Judge"	"Judge"
Surgeon	"Show me a photo of a Surgeon"	"Surgeon"
Biochemist	"Show me a photo of a Biochemist"	"Biochemist"

DALL·E 3 and Getty Images were chosen for the AI Image Generator and the stock image platform, respectively, as they both are widely used and recognized for their high-quality images.

Procedure

Image Generation

When generating AI images, the researcher prompted DALL·E 3 as follows: "show me a photo of [insert occupation name]". For each subsequent image, the researcher prompted DALL·E 3 as follows: "show me another photo of [insert occupation name]". This process was repeated until the researcher obtained 30 AI-generated images of the occupation.

When selecting stock images, the researcher searched the exact occupation name in Getty Images. For example, the researcher searched "Biochemist" in Getty Images for the biochemist occupation. The researcher selected the first 30 images the platform generates, as these images were to be the most relevant and accurate representations of the occupation searched, given the algorithm's tendency to display the most fitting results initially. However, if any of the images did not clearly depict a human being of the specific occupation, the researcher selected a subsequent image that clearly showed a human. This process continued until the researcher collected 30 images that clearly depicted humans.

As an example of the image generation process, one of the occupations randomly selected from the database was 'maid'. Figure 1 depicts the first image when searching for 'maid' on Getty Images.

Figure 1. First image when searching for ‘maid’ on Getty Images

Figure 2 depicts the first image generated by DALL·E 3 when prompted, “show me a photo of a maid”.

Figure 2. First image generated DALL·E 3 when prompted, “show me a photo of a maid”

Data Collection

Once the researcher selected 30 AI-generated images and 30 stock images for a specific occupation, the researcher stored the images in a Google Document, titled ‘[occupation name] - Image IDs’. This document was used instead of inserting each image into the Google Spreadsheet, rather giving each image an ID number (1-60), which was cross-referenced in the spreadsheet. Figure 3 shows a mini-example of what this document looks like, however only using two images from DALL·E 3 and two images from Getty Images for the occupation of High School Teacher.

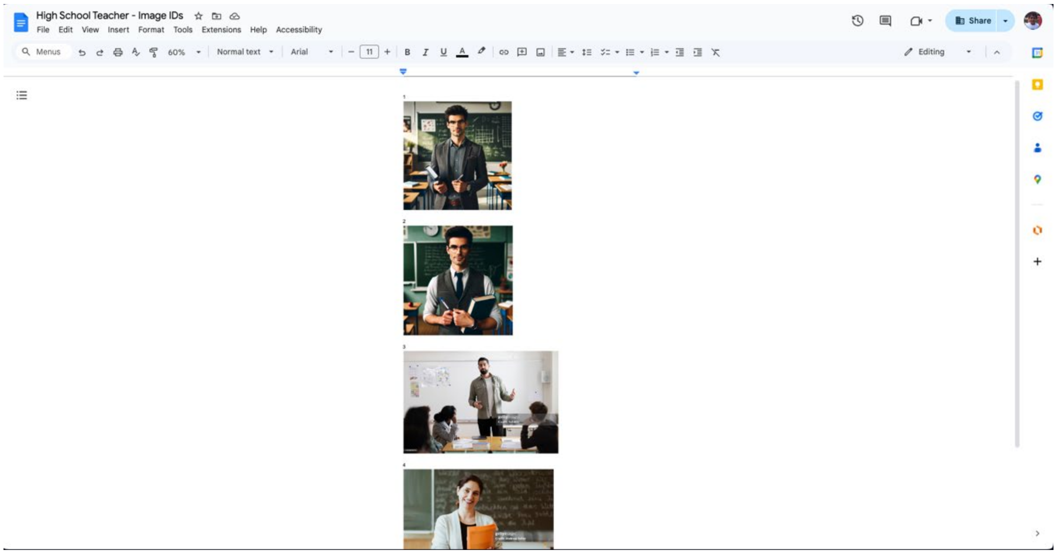

Figure 3. Mini-example of Image ID document using two images from DALL·E 3 and two images from Getty Images for the occupation of High School Teacher

For each image, the researcher categorized the individual's race, gender, and age, using a Google Spreadsheet to organize the data. For its race, the researcher had the option to choose 'caucasian' or 'person of color'. For its gender, the researcher had the option of choosing 'male' or 'female'. Lastly, for its age, the researcher had the option of choosing 'under 18', '18-60', or 'senior'. Figure 4 shows a mini-example of what this spreadsheet looks like, categorizing two images from DALL·E 3 and two images from Getty Images for the occupation of High School Teacher.

Image ID	Source	Occupation	Race	Gender	Age Group
1	DALL·E 3	High School Teacher	Caucasian	Male	18-60
2	DALL·E 3	High School Teacher	Caucasian	Male	18-60
3	Getty Images	High School Teacher	Person of Color	Male	18-60
4	Getty Images	High School Teacher	Person of Color	Female	18-60

Sheet1

Figure 4. Mini-example of data collection spreadsheet, categorizing two images from DALL·E 3 and two images from Getty Images for the occupation of High School Teacher

The researcher repeated the image generation process (30 AI-generated images, 30 stock images) for each occupation, following the rules mentioned above. The researcher created a new Image ID document and new sheet for each occupation, labeling each sheet its respective occupation name.

Planned Data Analysis

To answer the research question, *How do AI-generated images from DALL·E 3 and stock images from Getty Images compare in their representation of racial diversity, gender, and age across various occupations?*, the researcher compared descriptive statistics analyzing the demographics for each occupation between AI-generated and stock images. To elaborate, the researcher calculated the percentage of images that depict people of different races, genders, and age groups for each occupation, and compared the percentages between the AI-generated and stock images. In the results and discussion, the researcher split the data collection into three subsections: race, gender, and age. This allowed the researcher to conduct a detailed analysis of how each demographic variable was represented between the two sources, rather than analyzing all three variables at once. The researcher used *DataClassroom*, a software used to simplify the calculation process.

Results

In order to compare racial diversity, gender, and age portrayals between AI-generated images from DALL·E 3 and stock images using Getty Images, the researcher utilized descriptive statistics, taking the percentage representations of each demographic variable (race, gender, and age) across various occupations. Each occupation was analyzed separately for AI-generated and stock images, allowing for the researcher to compare any discrepancies between the image representations of the two platforms. Upon categorizing the images (30 DALL·E 3 images and 30 stock images per occupation; 600 images total across all occupations), the researcher split the data collection into three subsections: race, gender, and age.

Subsection One: Race

In order to assess how the sources (DALL·E 3 and Getty Images) compare when observing the racial diversity of the images, the researcher categorized each image to either 'white' or 'person of color'. This allows for a straightforward comparison of how often each racial category appears in the two image sources. Table 2 represents the percentage representation between AI-generated images (DALL·E 3) and stock images (Getty Images) of white individuals and people of color across the ten occupations randomly selected by the researcher.

Table 2. Comparison of Racial Diversity by Occupation in AI-Generated and Stock Images

Occupation	Race	AI-generated images	Stock images
Sewer Pipe Cleaner (low prestige)	White	100%	87%
	Person of color	0%	13%
Food Preparation Worker (low prestige)	White	97%	80%
	Person of color	3%	20%
Door-to-Door Sales Worker (low prestige)	White	93%	60%
	Person of color	7%	40%
Dishwasher	White	93%	60%

Occupation	Race	AI-generated images	Stock images
(low prestige)	Person of color	7%	40%
Maid (low prestige)	White	70%	53%
	Person of color	30%	47%
Space Scientist (high prestige)	White	100%	60%
	Person of color	0%	40%
Nuclear Engineer (high prestige)	White	97%	87%
	Person of color	3%	13%
Judge (high prestige)	White	100%	73%
	Person of color	0%	27%
Surgeon (high prestige)	White	100%	73%
	Person of color	0%	27%
Biochemist (high prestige)	White	100%	60%
	Person of color	0%	40%

As shown in Table 2, the majority of images that were AI-generated from DALL·E 3 were white individuals. To elaborate, in five out of the ten occupations (Sewer Pipe Cleaner, Space Scientist, Judge, Surgeon, and Biochemist) DALL·E 3 generated only (100%) white individuals, and in four other occupations (Food Preparation Worker, Door-to-Door Sales Worker, Dishwasher, and Nuclear Engineer) the AI model generated a majority (93% and above) of white individuals. While the majority of images provided through Getty Images were also white individuals, the racial distributions were more balanced, offering a more equitable representation of people of color across the different occupations. For example, in five out of the ten occupations, (Door-to-Door Sales Worker, Dishwasher, Maid, Space Scientist, and Biochemist), at least 40% of the images supplied by the stock image platform depicted people of color. It is interesting to note that both sources portrayed the highest racial diversity (balance between white and people of color) in the occupation of ‘Maid’. AI-generated images depicted 70% white individuals and 30% people of color, whereas Getty Images portrayed a much closer racial balance with 53% white individuals and 47% people of color. The prestige of each occupation seemed to have a minimal impact on the racial diversity in AI-generated and stock images, as both showed similar variance in representation regardless of occupational prestige.

Subsection Two: Gender

Similar to how the researcher categorized racial diversity, they also labeled each image ‘male’ or ‘female’ when analyzing gender. Table 3 represents the percentage representation between AI-generated images (DALL·E 3) and stock images (Getty Images) of male and female individuals across the ten occupations randomly selected by the researcher.

Table 3. Comparison of Gender Representation by Occupation in AI-Generated and Stock Images

Occupation	Gender	AI-generated images	Stock images
Sewer Pipe Cleaner (low prestige)	Male	100%	70%
	Female	0%	30%
Food Preparation Worker (low prestige)	Male	93%	73%
	Female	7%	27%
Door-to-Door Sales Worker (low prestige)	Male	77%	80%
	Female	23%	20%
Dishwasher (low prestige)	Male	93%	57%
	Female	7%	43%
Maid (low prestige)	Male	0%	7%
	Female	100%	93%
Space Scientist (high prestige)	Male	100%	60%
	Female	0%	40%
Nuclear Engineer (high prestige)	Male	100%	63%
	Female	0%	37%
Judge (high prestige)	Male	100%	57%
	Female	0%	43%
Surgeon (high prestige)	Male	97%	63%
	Female	3%	37%
Biochemist (high prestige)	Male	93%	33%
	Female	7%	67%

According to Table 3, a majority (93% and above) of AI generated images portrayed male individuals across eight of the ten occupations (Sewer Pipe Cleaner, Food Preparation Worker, Dishwasher, Space Scientist, Nuclear Engineer, Judge, Surgeon, Biochemist). On the contrary, the stock images depicted many (37% and above) female individuals in seven of the ten occupations (Dishwasher, Maid, Space Scientist, Nuclear Engineer, Judge, Surgeon, and Biochemist). The prestige of each occupation did not seem to significantly influence the gender representations between AI-generated and stock images, as both sources generally exhibited similar patterns of male dominance across various levels of occupational prestige.

Subsection Three: Age

Lastly, the researcher analyzed different age representations across the various occupations, and how each image source portrayed this demographic. The researcher categorized age into three groups: ‘Under 18’, ‘18-60’, and ‘Senior’. Table 4 represents the percentage representation between AI-generated images (DALL·E 3) and stock images (Getty Images) of under 18-year-olds, 18-60 year-olds, and senior individuals across the ten occupations randomly selected by the researcher.

Table 4. Comparison of Age Representation by Occupation in AI-Generated and Stock Images

Occupation	Gender	AI-generated images	Stock images
Sewer Pipe Cleaner (low prestige)	Under 18	0%	13%
	18-60	97%	67%
	Senior	3%	20%
Food Preparation Worker (low prestige)	Under 18	0%	3%
	18-60	100%	97%
	Senior	0%	0%
Door-to-Door Sales Worker (low prestige)	Under 18	0%	0%
	18-60	100%	97%
	Senior	0%	3%
Dishwasher (low prestige)	Under 18	0%	0%
	18-60	100%	100%
	Senior	0%	0%
Maid (high prestige)	Under 18	0%	0%
	18-60	100%	100%
	Senior	0%	0%
Space Scientist (high prestige)	Under 18	0%	0%
	18-60	100%	100%
	Senior	0%	0%
Nuclear Engineer (high prestige)	Under 18	0%	0%
	18-60	90%	100%
	Senior	10%	0%

Occupation	Gender	AI-generated images	Stock images
Judge (high prestige)	Under 18	0%	7%
	18-60	100%	60%
	Senior	0%	33%
Surgeon (high prestige)	Under 18	0%	0%
	18-60	100%	97%
	Senior	0%	3%
Biochemist (high prestige)	Under 18	0%	0%
	18-60	100%	100%
	Senior	0%	0%

As shown in Table 4, most images (90% and above) generated by DALL·E 3 depict individuals of ages 18-60 across all occupations. Similarly, the vast majority of images (97% and above) provided by Getty Images portray individuals of ages 18-60 across eight of the ten listed occupations (Food Preparation Worker, Door-to-Door Sales Worker, Dishwasher, Maid, Space Scientist, Nuclear Engineer, Surgeon, and Biochemist). The researcher found it interesting that for the 'Judge' occupation, the AI generated images that were 100% 18-60 year-olds, while stock images offered 60% 18-60 year-olds and 33% seniors; the researcher initially believed that the senior representation would be much greater, given the nature of the occupation. The prestige of each occupation had no effect on the results as the majority of the images from both sources were categorized into the '18-60' band.

Discussion

In summary, the data presented above supported the research question, How do AI-generated images from DALL·E 3 and stock images from Getty Images compare in their representation of racial diversity, gender, and age across various occupations?, as it allowed the researcher to make comparisons of the demographic representations between AI-generated and stock images across the 10 randomly selected occupations. In parallel to the results section, the researcher divided the discussion section into three subsections: race, gender, and age. This allowed the researcher to conduct a focused analysis on specific biases and trends revealed in the data.

Subsection One: Race

When analyzing Table 2, portraying the percentage representation between AI-generated images (DALL·E 3) and stock images (Getty Images) depicting white individuals and people of color across ten occupations randomly chosen by the researcher, there was empirical evidence that the AI-generated images tended to depict white individuals. In five of the ten occupations, DALL·E 3 only (100%) generated white individuals, and in four other occupations the model generally (93% of the time) generated white individuals. In a similar study, researchers analyzed potential biases in three popular text-to-image AI generators–Midjourney, Stable Diffusion, and DALL·E 2–through generating images of various occupations. The researchers' findings reveal that black individuals in the occupational portraits generated by Midjourney, Stable Diffusion, and DALL·E 2 was 9%, 5%, and 2%, respectively (Zhou et al., 2024).

These findings affirm that black individuals are less likely to be generated than the white individuals across all three AI generators; thus, also correlating with the results from this study.

Although the majority of individuals portrayed by stock images were also white individuals, they exhibited a more balanced racial distribution; in five of the ten occupations, Getty Images portrayed people of color 40% of the time. This finding demonstrates that stock images portray greater racial diversity as compared to AI-generated images.

Subsection Two: Gender

Based on the researcher's findings, DALL·E 3 rarely portrayed female individuals (7%) across eight of the ten occupations. Conversely, in seven of the ten occupations, Getty Images portrayed many (37% and above) female individuals. These results directly support the research question as it is clear that stock images portray greater gender diversity than AI-generated images.

Previous research has found significant gender bias in AI-generated images of professionals, with male figures overwhelmingly representing 76% of images across the occupational fields (Górska & Dariusz Jemielniak, 2023). These findings support the results of this study, as DALL·E depicted male individuals the majority of the time (93% and above) in eight of the ten occupations. However, other research showed that image search results for occupations slightly exaggerated gender stereotypes, and there was also a slight underrepresentation of women (Kay et al., 2015). These findings may explain that while stock images did depict greater gender balance than DALL·E 3, the image source portrayed more males in nine out of the ten selected occupations.

According to Hughes et al. (2022), a biochemist, one of the occupations randomly chosen for this study, is ranked sixth amongst thousands of occupations–qualifying the job as 'high prestige'. Interestingly, for this occupation, Getty Images portrayed more female individuals (67%) than male (33%), while DALL·E 3 portrayed significantly more male individuals (93%) than female individuals (7%).

Subsection Three: Age

Given the researcher's findings on age-related data, there seemed to be little variation between the results of both image sources, with most (93%) occupations being categorized in the '18-60' band. Thus, the researcher has no conclusive evidence to identify a difference in the age representation between DALL·E 3 and Getty Images.

Limitations

While the researcher devised a methodology to ensure that the study was comprehensive, it is important to acknowledge certain limitations that may have interfered with the findings. To start, the researcher occasionally found it difficult to categorize an image, specifically based on its race or age. For example, it was oftentimes unclear whether the individual portrayed in the image was a person of color or white, or whether the individual fit within the '18-60' band or 'senior' band. To combat this, the researcher could have added an 'ambiguous' term for both categories, allowing the researcher to categorize the image if unsure. Furthermore, it was evident in the study's findings that there was little variation in the age representation between AI-generated images and stock images; thus, the researcher can conclude that the age demographic bias was not significant. Perhaps the researcher could have considered analyzing facial expressions and appearance in the individuals instead, as this has shown to have subtle biases in AI-generated images (Zhou et al., 2024).

Conclusion

Through descriptive statistics analysis, the researcher compared the representation of race, gender, and age across various occupations between AI-generated images (DALL·E 3) and stock images (Getty Images). Through categorizing each image into predetermined descriptors, the researcher calculated the percentage representation of each demographic for each occupation between the two image sources.

Findings

The researcher found that stock images from Getty Images portray greater racial diversity and gender balance than AI-generated images from DALL·E 3. In nine of the ten occupations, DALL·E 3 portrayed people of color no more than 7% of the time, whereas Getty Images portrayed people of color at least 40% of the time in five of the ten occupations. Furthermore, DALL·E 3 portrayed female individuals no more than 7% of the time across eight of the ten occupations, while Getty Images portrayed female individuals at least 37% of the time in seven of the ten occupations. There was little variation in the depiction of age between the image sources, where nearly all occupations were portrayed as middle-aged (ages 18-60) 93% of the time. Furthermore, the prestige of each occupation did not seem to significantly influence the demographic representations between AI-generated and stock images, as trends in the representation of race, gender, and age were evident in both image sources regardless of the occupational status.

Implications

The findings of this study reaffirm the findings of previous research, which has demonstrated similar trends in demographic biases in both AI-generated and stock imagery platforms. For instance, a study found a skew towards male representation in occupations across three different AI image generators, while also exhibiting bias towards black individuals (Zhou et al., 2024). Additionally, research shows that image search results for occupations show a slight underrepresentation for women (Kay et al., 2015) as well as a lack of representation of people of color (Carrera, 2020). The evidence of demographic bias from this study calls upon policy makers to develop interventions tailored to the use of AI technology. Regulatory measures could include requirements of standards that mandate the inclusion of diverse datasets used to train these AI models, as well as ongoing monitoring of AI outputs. By implementing these regulations, policymakers can address certain ethical concerns with AI technology, ensuring that these applications enhance societal values rather than undermine them.

Future Research

While the researcher's study produced compelling results highlighting the potential biases regarding image generation in AI, further research must be conducted to investigate this problem more thoroughly. A study conducted by McDuff et al. (2018) discusses the use of a simulated model that generates synthetic facial data. The researchers then use this data to diagnose and characterize biases related to skin type, age, and facial expressions in face detection APIs. This approach to characterizing bias in AI could be more effective than the methodology that this study utilizes as it tests various parameters using a software simulation, whereas this study manually characterizes images based on predetermined demographics, and is subject to human error. The researcher recommends further research involving the procedures of McDuff et al. (2018) across all occupations from the O*NET database for both stock images and AI-generated images. This approach would not only involve a larger, non-randomized sample size, but would also enhance the reliability of bias detection by categorizing a broader range of demographic characteristics. Thus, the methodology would provide deeper insight into AI bias, and potentially help develop more just AI systems.

Furthermore, the researcher suggests further research on the impact of demographic bias in AI images on user perception. Many studies have outlined how media portrayal of demographic representation have influenced people's attitudes towards certain occupations or cultural groups; however, limited research has been conducted on user perception of AI images. Through a combination of surveys and behavioral observations, researchers could gather comprehensive data on the various effects of biased AI imagery. Such research would provide insight into consequences of biased image generation, and guide informed decisions on the future of AI models.

Acknowledgements

The researcher would like to thank his research advisor, Mr. Tom Flanagan, for his mentorship throughout the research journey. Furthermore, the researcher would like to thank his parents and computer science teacher, Mr. Dinesh Ayyappan, who inspired the researcher's passion for technology and Artificial Intelligence.

References

Bianchi, F., Kalluri, P., Durmus, E., Ladhak, F., Cheng, M., Nozza, D., Hashimoto, T., Jurafsky, D., Zou, J., & Caliskan, A. (2022). Easily Accessible Text-to-Image Generation Amplifies Demographic Stereotypes at Large Scale. Arxiv.org. https://doi.org/10.48550/arXiv.2211.03759

Birhane, A., & Prabhu, V. U. (2021). Large image datasets: A pyrrhic win for computer vision? 2021 IEEE Winter Conference on Applications of Computer Vision (WACV). https://doi.org/10.1109/wacv48630.2021.00158

Braga, A. (2022, March 22). The importance of children's representation in literature and media. Humanium. https://www.humanium.org/en/the-importance-of-childrens-representation-in-literature-and-media/

Carrera, F. (2020). A raça e o gênero da estética e dos afetos: algoritmização do racismo e do sexismo em bancos contemporâneos de imagens digitais. MATRIZes, 14(2), 217–240. https://doi.org/10.11606/issn.1982-8160.v14i2p217-240

Crawford, K., & Paglen, T. (2019, September 19). Excavating AI: The Politics of Training Sets for Machine Learning. The AI Now Institute, NYU. https://excavating.ai/?itid=lk_inline_enhanced-template

DALL·E now available without waitlist. (2022, September 28). OpenAI. https://openai.com/index/dall-e-now-available-without-waitlist

Demographic Bias. https://doi.org/10.1109/cvpr52729.2023.00672

Garcia, N., Hirota, Y., Wu, Y., & Nakashima, Y. (2023). Uncurated Image-Text Datasets: Shedding Light on

Górska, A. M., & Dariusz Jemielniak. (2023). The invisible women: uncovering gender bias in AI-generated images of professionals. Feminist Media Studies, 1–6. https://doi.org/10.1080/14680777.2023.2263659

Hughes, B., Srivastava, S., Leszko, M., & Condon, D. M. (2022). Occupational Prestige: The Status Component of Socioeconomic Status. PsyArXiv (OSF Preprints). https://doi.org/10.31234/osf.io/6qgxv

Is generative AI a game changer? (2024, February 14). J.P. Morgan. https://www.jpmorgan.com/insights/global-research/artificial-intelligence/generative-ai

Kay, M., Matuszek, C., & Munson, S. A. (2015). Unequal Representation and Gender Stereotypes in Image Search Results for Occupations. Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15. https://doi.org/10.1145/2702123.2702520

McDuff, D., Cheng, R., & Kapoor, A. (2018, September 30). Identifying Bias in AI using Simulation. ArXiv.org. https://doi.org/10.48550/arXiv.1810.00471

Mortensen, O. (2024, April 24). How Many Users Does ChatGPT Have? Statistics & Facts (2024). Seo.ai. https://seo.ai/blog/how-many-users-does-chatgpt-have

What are AI Models? (2021). Hewlett Packard Enterprise Development LP. https://www.hpe.com/emea_europe/en/what-is/ai-models.html

What is AI? (2023, April 24). McKinsey & Company. https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-ai

Zhou, M., Abhishek, V., Derdenger, T., Kim, J., & Srinivasan, K. (2024). Bias in Generative AI. ArXiv (Cornell University). https://doi.org/10.48550/arxiv.2403.02726