

A Nudez como Estereótipo de Beleza através das Lentes do Instagram

Breno Matos¹, Fabrício Benevenuto¹, Ana Paula Couto da Silva¹

¹Departamento de Ciência da Computação
Universidade Federal de Minas Gerais

{brenomatos, fabricio, ana.coutosilva}@dcc.ufmg.br

Abstract. *Social media platforms take part in the process of defining beauty standards, whether reinforcing real-world standards or creating new definitions to be replicated outside the digital environment. However, this process is conducive to the surfacing of biases. Previous studies focus on assessing the impact of gender and ethnicity on beauty standards and biases that may arise from them. However, little is discussed about the impact of nudity on the perception of beauty in images on social media. In this work, we explore this gap, analyzing photos posted on Instagram.*

Resumo. *Redes sociais são parte do processo de construção de padrões de beleza, seja reforçando padrões do mundo real ou criando novas definições, replicadas fora do ambiente digital. Entretanto, este processo é propício para o surgimento de vieses. Na literatura, grande parte dos estudos focam em avaliar impacto de gênero e etnia na construção de padrões e vieses. Entretanto, pouco é discutido sobre o impacto de nudez sobre a percepção de beleza em imagens publicadas em redes sociais. Neste trabalho, exploramos essa lacuna, utilizando fotos publicadas no Instagram.*

1. Introdução

Na atualidade, redes sociais são peças fundamentais na construção de estereótipos e padrões de beleza, muitas vezes replicando padrões do mundo real. Entretanto, através do engajamento de seus usuários, o processo inverso pode ocorrer, em que as redes sociais definem padrões que são replicados fora do ambiente digital. Essa dinâmica pode favorecer a criação de vieses, como de gênero ou etnia, propagando estereótipos danosos através dos algoritmos que decidem o que é apresentado aos usuários.

Padrões de beleza humanos podem ser capturados por algoritmos de aprendizado de máquina [Eisenthal et al. 2006]. No entanto, o estudo e conceituação de beleza é complexo e está presente em diferentes áreas de pesquisa: [Grammer et al. 2003, Fink et al. 2006] discutem que padrões de beleza reforçam características físicas que simbolizam saúde, além de denotar parceiros mais aptos para reprodução. Há também diferenças culturais na conceituação de beleza, com povos diferentes favorecendo características distintas para classificar pessoas como bonitas [Coetzee et al. 2014].

Na literatura, a maioria dos esforços para estudo de vieses é centrado em etnia ou gênero, como por exemplo em [Zehlike et al. 2017], com análises realizadas com base na rede social profissional *Xing*. Existem também esforços para analisar estereótipos sobre atração física no ambiente online, como realizado em [Araújo et al. 2016], em que

os autores propõem um estudo de resultados de buscas de imagem no Google, identificando claro viés racial. Em [Noble 2013], é avaliada a representação de gênero e raça em sistemas de busca na Web e seus resultados. Em [Kay et al. 2015], são apresentados os resultados de estudos que caracterizam o viés de gênero presente em resultados de pesquisas de imagens para diferentes ocupações. [Sweeney 2013] apresenta um estudo de viés nos resultados de busca no Google para sobrenomes mais presentes na população negra americana. O trabalho apresentado em [Noble 2012] avalia o uso comercial de máquinas de busca com foco na sua representação de mulheres, em especial, mulheres negras.

Apesar de um conjunto rico de trabalhos que avaliam diferentes características associadas à beleza, estudos que investigam a relação entre beleza e nudez em redes sociais ainda são raros na literatura. Esse tipo de análise se mostra uma tarefa complexa, uma vez que a definição de nudez é subjetiva, gerando controvérsias dos limites impostos por diferentes plataformas para classificação do nível de nudez (quantidade de pele à mostra, por exemplo) e do tipo de nudez (campanhas sobre câncer de mama¹ ou fotos eróticas, por exemplo).

Neste trabalho, apresentamos um estudo de nudez e beleza em redes sociais, avaliando a relação entre estes dois fatores, oferecendo uma primeira visão sobre como a nudez está embutida em padrões de beleza muitas vezes favorecidos por algoritmos de redes sociais. Para tal, coletamos postagens do Instagram, rede social de compartilhamento de fotos e vídeos com milhões de usuários ativos, que contêm *hashtags* que descrevem homens e mulheres no intervalo de Dezembro de 2020, selecionando as 100 postagens com maior número de comentários e curtidas. Em seguida, realizamos a extração de características que quantificam a beleza e nudez presentes nas imagens. Apesar de nudez total não ser permitida no Instagram,² há indícios de que a rede social privilegia postagens com fotos ou vídeos de pessoas com muita pele à mostra,³ como fotos trajando roupas íntimas, de banho ou cobrindo-se utilizando outras partes do corpo.

Nossos resultados mostram que os valores médios de nudez para postagens com mulheres são maiores que suas contrapartidas masculinas. Valores de nudez para as fotos de *hashtags* que denotam pessoas mais jovens são maiores quando comparadas com fotos de adultos mais velhos. Por fim, há aumento significativo de nudez em fotos com *hashtags* que descrevem jovens quando são adicionadas adjetivos que denotam beleza.

O restante do trabalho está dividido desta forma: a Seção 2 apresenta nossa metodologia, descrevendo como foi feita a coleta de dados. Nossas análises e resultados são apresentados na Seção 3 e a Seção 4 apresenta as conclusões do trabalho.

2. Metodologia

Coleta e Tratamento de Dados. Com o objetivo de caracterizar a presença de nudez e a possível definição de estereótipos de beleza nos compartilhamentos de fotos que retratam homens e mulheres, foram coletados dados históricos da rede social Instagram,⁴ a partir da plataforma Crowdtangle,⁵ ferramenta do Facebook que facilita a obtenção e

¹<https://ncac.org/news/facebook-oversight-board-decisions-2021>

²<https://help.instagram.com/477434105621119>

³<https://algorithmwatch.org/en/instagram-algorithm-nudity/>

⁴<https://instagram.com/>

⁵<https://apps.crowdtangle.com/>

	#boy	#handsomeboy	#man	#handsomeman	#girl	#cutegirl	#woman	#beautifulwoman
Total de Posts	10000	976	8389	384	10000	1354	10000	1198

Tabela 1. Total de postagens coletadas para cada hashtag

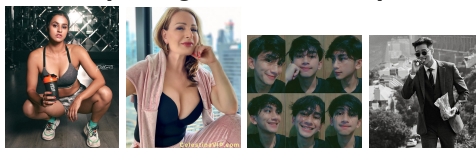


Figura 1. Exemplos de postagens das hashtags: #cutegirl, #beautifulwoman, #handsomeboy e #handsomeman, respectivamente

análise de dados de suas redes sociais, especificamente, Instagram e Facebook. Utilizar o Crowdtangle proporciona vantagens por permitir diversas customizações nas buscas, como selecionar dados históricos e filtrar postagens por tipo de mídia, além de facilitar a coleta de informações do Instagram, que não possui API oficial para essa finalidade, apenas APIs de desenvolvimento.⁶

Os dados coletados englobam as postagens com fotos realizadas no mês de Dezembro de 2020 e que incluem, na descrição textual, alguma das seguintes *hashtags*: #boy, #handsomeboy, #man, #handsomeman, #girl, #cutegirl, #woman, #beautifulwoman. As *hashtags* escolhidas abrangem diferentes faixas etárias (boy/man e girl/woman), acrescentando ou não adjetivos que denotem beleza (*handsome* para homens e *cute/beatiful* para mulheres). Com essas *hashtags*, será possível observar o tipo de conteúdo presente em diferentes faixas etárias dos dois gêneros, além de permitir análises da percepção de nudez e beleza quando são adicionados os adjetivos escolhidos. A Tabela 1 apresenta o total de postagens coletadas para cada uma das *hashtags* anteriormente listadas.

Para as análises apresentadas, consideramos, para cada *hashtag*, as 100 postagens com o maior índice de engajamento (soma de curtidas e comentários). Todas as postagens presentes neste ranqueamento possuem *links* disponíveis para *download* de suas imagens. A Figura 1 apresenta alguns exemplos de postagens com as *hashtags* coletadas.

Extração de Características. Para quantificar os índices de nudez e beleza associados às postagens coletadas, utilizamos duas ferramentas. A primeira, chamada NudeNet,⁷ é uma biblioteca para linguagem Python, onde o percentual de nudez é mensurado a partir de dois valores: *safe*, que indica que a imagem não é sexualmente explícita e *unsafe*, que indica que a imagem é sexualmente explícita, sendo a soma dos dois fatores igual a 1. A ferramenta, treinada com extenso conjunto de dados de imagens sexualmente explícitas, atribui as probabilidades de uma imagem ser sexualmente explícita ou não aos valores *safe* e *unsafe*. Para quantificar o nível de nudez em nossas postagens, consideramos o valor atribuído à escala *unsafe*.

A segunda ferramenta, para extração do nível de beleza associada à postagem em análise, é a API Face++,⁸ que detecta rostos em uma determinada imagem, inferindo informações como idade, gênero, emoções, intensidade de sorriso, grau de beleza, entre outros, para cada face, com acurácia superior a 90% [Bakhshi et al. 2014]. Nos resultados apresentados, utilizamos a média dos valores da característica *beauty*, que quantifica a

⁶<https://www.instagram.com/developer/>

⁷<https://github.com/notAI-tech/NudeNet>

⁸<https://www.faceplusplus.com/>

		#boy	#handsomeboy	#man	#handsomeman	#girl	#cutegirl	#woman	#beautifulwoman
Curtidas	Média	49117.03	4954.07	19670.11	7607.12	73271.82	7516.92	42258.71	2979.95
	Mediana	43099.5	4263.0	13270.0	6115.5	61912.5	5437.0	35737.0	1761.0
Comentários	Média	259.03	94.53	16320.77	79.35	957.26	158.02	1172.82	62.19
	Mediana	178.5	46.5	188.0	57.5	620.0	79.0	558.5	32.5

Tabela 2. Métricas de engajamento de cada hashtag coletada

beleza do rosto da pessoa presente na imagem e que possui dois campos, *male_score* (nível de beleza dado ao rosto quando avaliado por um homem) e *female_score* (nível de beleza dado ao rosto quando avaliado por uma mulher). Ambos os valores variam entre 0 e 1, com valores mais altos indicando que o rosto detectado é mais bonito.

Limitações das Ferramentas. Apesar de robustas, as ferramentas selecionadas não fornecem informações sobre os dados de treinamento, por exemplo se foram utilizadas imagens de diferentes etnias. No caso da biblioteca NudeNet, é possível acessar todas as imagens utilizadas no treino, que compreendem mais de 19 Gb de dados⁹. A API Face++, por sua vez, não divulga bases de dados ou possui código de seus modelos abertos, levando a críticas sobre vieses em suas ferramentas no passado.¹⁰

3. Análises e Discussões

Na Tabela 2, são apresentadas média e mediana para a quantidade de curtidas e comentários das top-100 postagens coletadas, divididas por sua *hashtag* de origem.

Métricas de Engajamento. Todas as *hashtags* apresentam queda de média de curtidas quando os adjetivos *cute*, *handsome* ou *beautiful* são adicionados, com redução de até 92.94%, no caso do par *#woman* e *#beautifulwoman*. Um comportamento similar é observado na média de comentários, com redução de até 99.51% no caso do par *#man* e *#handsomeman*. De forma análoga, as medianas, tanto para curtidas quanto comentários, são reduzidas quando os adjetivos são adicionados.

Comparando os dois gêneros, todas as medianas para curtidas e comentários nas *hashtags* do gênero feminino apresentam valores maiores que suas contrapartidas masculinas. Comportamento similar ocorre na média de curtidas, exceto no par *#handsomeman* e *#beautifulwoman*, em que há diferença de 60.82%.

A média de comentários apresenta um comportamento distinto, em que as *hashtags* que descrevem homens mais jovens possuem média de comentários menores que suas contrapartidas femininas. Entretanto, comparando as *hashtags* que utilizam as palavras *man* e *woman*, o gênero masculino recebe mais comentários, apresentando médias maiores, com diferença de até 92.81% no par *#man* e *#woman*.

Beleza e Nudez. A Tabela 3 apresenta a média e mediana para os valores de beleza e nudez das postagens coletadas, divididas por sua *hashtag* de origem. A seguir, analisamos a relação de beleza e nudez considerando os gêneros masculino e feminino.

Gênero Masculino. As *hashtags* com a palavra *boy* possuem média de nudez maior que as com a palavra *man*, indicando mais conteúdo explícito em fotos de pessoas mais jovens, com o par *#boy* e *#handsomeboy* apresentando aumento em 600% de sua média de nudez. Entretanto, as médias de beleza são mais elevadas para as *hashtags* com a palavra *man*,

⁹https://archive.org/details/NudeNet_classifier_dataset_v1

¹⁰<https://www.technologyreview.com/2019/01/29/137676/making-face-recognition-less-biased-doesnt-make-it-less-scary/>

		#boy	#handsomeboy	#man	#handsomeman	#girl	#cutegirl	#woman	#beautifulwoman
Nudez	Média	0.05	0.35	0.24	0.23	0.39	0.58	0.39	0.37
	Mediana	0.01	0.12	0.02	0.03	0.17	0.71	0.2	0.16
Beleza	Média	0.61	0.69	0.68	0.71	0.71	0.76	0.7	0.66
	Mediana	0.61	0.69	0.68	0.74	0.73	0.77	0.72	0.66

Tabela 3. Média e mediana para os valores de nudez e beleza de cada *hashtag*

com *#handsomeman* possuindo a maior média dentre todas, com valor 0.71. No gênero masculino, portanto, há indícios de que a percepção de beleza vai na direção contrária da percepção de nudez para fotos de homens mais velhos.

O par de *hashtags* com o adjetivo *handsome* apresenta um comportamento distinto, com queda de 75% da mediana quando utilizamos a palavra *man* ao invés de *boy*. Para a mediana de beleza, essa tendência é inversa, com aumento de 7.24% da mediana de beleza quando a palavra *man* é utilizada. Apesar da queda da média e da mediana das métricas de engajamento, os dados mostram o aumento da média e mediana de nudez e beleza para o par *#boy* e *#handsomeboy*.

Por fim, as fotos de *#cuteboy* apresentam fotos de homens jovens (diversos aparentam ser adolescentes), muitas vezes com roupas íntimas e de banho. Por outro lado, *#handsomeman* apresenta homens de diversas idades, na sua maioria completamente vestidos e vários com roupas sociais, camisas de botões e ternos.

Gênero Feminino. As *hashtags* *girl* e *woman* possuem a mesma média de nudez. Entretanto, quando adicionados os adjetivos *cute* e *beautiful*, *#cutegirl* apresenta aumento de 48.71% na média de nudez, com *#beautifulwoman* apresentando média 36.2% menor que sua contrapartida mais jovem.

O par *#girl* e *#cutegirl* apresenta aumento de média e mediana para nudez e beleza, enquanto *#woman* e *#beautifulwoman* apresenta queda para os os mesmo valores, apresentando indícios de que mulheres mais jovens são percebidas como mais bonitas, além de publicarem mais fotos com conteúdo sexualmente explícito.

Fotos da *hashtag* *#cutegirl*, de forma geral, apresentam mulheres jovens, algumas aparentando ser adolescentes. Várias fotos possuem mulheres de roupas íntimas, academia ou de banho. Um comportamento similar é apresentado nas fotos de *#beautifulwoman*, porém, com mulheres mais velhas.

Comparação entre os gêneros. Todas as *hashtags* que descrevem mulheres possuem médias e medianas de nudez iguais ou maiores que suas contrapartidas masculinas. Um comportamento similar ocorre para as médias e mediana de beleza, exceto no par *#handsomeman* e *#beautifulwoman*. Este processo apresenta indícios de que a fotos de mulheres no Instagram apresentam mais conteúdo sexualmente explícito no conjunto de dados coletado, assim como maiores valores de beleza que postagens de *hashtags* masculinas.

4. Conclusão

Esse trabalho apresentou uma primeira análise sobre nudez em fotos publicadas no Instagram. Com o acréscimo dos adjetivos que indicam beleza, as medições de nudez apresentam crescimento em todas as *hashtags* coletadas, exceto no caso do par *#woman* e *#beautifulwoman*. As *hashtags* que descrevem pessoas jovens possuem crescimento da média de nudez significativo com o acréscimo de adjetivos associados a beleza. O par

#boy e *#handsomeboy* apresenta aumento em 600% de sua média de nudez. Para o par *#girl* e *#cutegirl*, o aumento é de aproximadamente 48.71%. Porém, todas as *hashtags* que descrevem mulheres possuem médias de nudez maiores que suas contrapartidas masculinas. Além disso, fotos de mulheres, de forma geral, apresentam mais curtidas em média, enquanto fotos de homens mais velhos apresentam mais comentários que suas contrapartidas femininas.

Nossos resultados indicam que fotos de mulheres no conjunto de dados coletados do Instagram apresentam níveis maiores de nudez. Além disso, apresentamos indícios de que a nudez influencia no nível de beleza percebida para fotos de pessoas jovens, com médias de beleza aumentando conforme as médias de nudez também aumentam.

Agradecimentos. Este trabalho foi realizado com apoio financeiro do CNPq e FAPEMIG.

Referências

- Araújo, C. S., Meira, W., and Almeida, V. (2016). Identifying stereotypes in the online perception of physical attractiveness. In *International Conference on Social Informatics*, pages 419–437. Springer.
- Bakhshi, S., Shamma, D. A., and Gilbert, E. (2014). Faces engage us: Photos with faces attract more likes and comments on instagram. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 965–974.
- Coetzee, V., Greeff, J. M., Stephen, I. D., and Perrett, D. I. (2014). Cross-cultural agreement in facial attractiveness preferences: The role of ethnicity and gender. *PLoS one*, 9(7):e99629.
- Eisenthal, Y., Dror, G., and Ruppin, E. (2006). Facial attractiveness: Beauty and the machine. *Neural Computation*, 18(1):119–142.
- Fink, B., Grammer, K., and Matts, P. J. (2006). Visible skin color distribution plays a role in the perception of age, attractiveness, and health in female faces. *Evolution and Human Behavior*, 27(6):433–442.
- Grammer, K., Fink, B., Møller, A. P., and Thornhill, R. (2003). Darwinian aesthetics: sexual selection and the biology of beauty. *Biological reviews*, 78(3):385–407.
- Kay, M., Matuszek, C., and Munson, S. A. (2015). Unequal representation and gender stereotypes in image search results for occupations. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 3819–3828.
- Noble, S. U. (2012). Missed connections: What search engines say about women. *Bitch Magazine*, 54:36–41.
- Noble, S. U. (2013). Google search: Hyper-visibility as a means of rendering black women and girls invisible. *InVisible Culture*, (19).
- Sweeney, L. (2013). Discrimination in online ad delivery. *Communications of the ACM*, 56(5):44–54.
- Zehlike, M., Bonchi, F., Castillo, C., Hajian, S., Megahed, M., and Baeza-Yates, R. (2017). Fa* ir: A fair top-k ranking algorithm. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, pages 1569–1578.