



**UNIVERSIDADE FEDERAL DA FRONTEIRA SUL
CAMPUS DE CHAPECÓ
CURSO DE CIÊNCIA DA COMPUTAÇÃO**

RAPHAEL BORGES DOS SANTOS FILHO

**INCLUSÃO DE PESSOAS TRANSGÊNERO NOS ALGORITMOS DE
RECONHECIMENTO AUTOMATIZADO DE GÊNERO PARA RECONHECIMENTO
FACIAL**

**CHAPECÓ
2021**

RAPHAEL BORGES DOS SANTOS FILHO

**INCLUSÃO DE PESSOAS TRANSGÊNERO NOS ALGORITMOS DE
RECONHECIMENTO AUTOMATIZADO DE GÊNERO PARA RECONHECIMENTO
FACIAL**

Trabalho de conclusão de curso apresentado como requisito para obtenção do grau de Bacharel em Ciência da Computação da Universidade Federal da Fronteira Sul.
Orientador: Prof. Me. Adriano Sanick Padilha

**CHAPECÓ
2021**

Filho, Raphael Borges dos Santos

Inclusão de pessoas transgênero nos algoritmos de reconhecimento automatizado de gênero para reconhecimento facial / Raphael Borges dos Santos Filho. – 2021.

35 f.: il.

Orientador: Prof. Me. Adriano Sanick Padilha.

Trabalho de conclusão de curso (graduação) – Universidade Federal da Fronteira Sul, curso de Ciência da Computação, Chapecó, SC, 2021.

1. Visão Computacional. 2. Reconhecimento Automatizado de Gênero. 3. Transgênero. I. Padilha, Prof. Me. Adriano Sanick, orientador. II. Universidade Federal da Fronteira Sul. III. Título.

© 2021

Todos os direitos autorais reservados a Raphael Borges dos Santos Filho. A reprodução de partes ou do todo deste trabalho só poderá ser feita mediante a citação da fonte.

E-mail: oborgesrapha@gmail.com

RAPHAEL BORGES DOS SANTOS FILHO

**INCLUSÃO DE PESSOAS TRANSGÊNERO NOS ALGORITMOS DE
RECONHECIMENTO AUTOMATIZADO DE GÊNERO PARA RECONHECIMENTO
FACIAL**

Trabalho de conclusão de curso apresentado como requisito para obtenção do grau de Bacharel em Ciência da Computação da Universidade Federal da Fronteira Sul.

Orientador: Prof. Me. Adriano Sanick Padilha

Este trabalho de conclusão de curso foi defendido e aprovado pela banca avaliadora em:
26/5/2021.

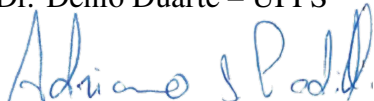
BANCA AVALIADORA



Prof. Me. Adriano Sanick Padilha – UFFS



Prof. Dr. Denio Duarte – UFFS



Prof. Dr. Leandro Bordin – UFFS

AGRADECIMENTOS

Dedico este trabalho a minha comunidade, a comunidade LGBTQIA+, aos nossos que são mortos por ser quem somos. É uma homenagem por toda nossa luta por direitos como cidadãos, luta por respeito por ser quem somos, luta pela não invisibilização da nossa existência... Luta por nossas vidas.

Agradeço a minha mãe Maria Borges (Bia) por todo o apoio e suporte nesses anos de caminhada longe de casa. Aos meus amigos, que sempre caminharam, se aventuraram, deram forças quando eu mesmo já não tinha e batalharam junto comigo, em especial Naomi Fernanda de Mello que, desde que a conheci, esteve comigo nas empreitadas da vida e hoje brilha como uma estrela no céu. Agradeço ao meu orientador Me. Adriano Sanick Padilha por desde o começo acreditar e me incentivar na realização deste trabalho me motivando, acreditando, aconselhando e orientando. A todos os meus professores e pelo conhecimento transmitido. Também aos membros das bancas que muito contribuíram para a execução e aprimoramento deste trabalho: Dr. Leando Bordin e Dr. Denio Duarte.

“I had my ups and downs, but I always find the inner strength to pull myself up. I was served lemons, but I made lemonade.”

(Beyoncé, Freedom)

RESUMO

As técnicas e aplicações de reconhecimento facial estão crescendo, tornando-se presentes cada vez mais no cotidiano. Muitas aplicações estão surgindo sem uma análise crítica do seu funcionamento e dos impactos na sociedade. As aplicações de reconhecimento facial que utilizam o reconhecimento automatizado de gênero (AGR), abrem uma brecha para o *misgender* tecnológico, que afeta diretamente a vida de pessoas transgênero. Este trabalho teve como objetivo analisar os impactos de um *dataset* que contenha pessoas transgênero na sua composição, pois os trabalhos que discutem este tipo de tecnologia não as incluem ou sequer consideram suas existências. Além disso, as principais aplicações existentes no mercado têm um desempenho muito inferior em comparação às pessoas cisgênero. A construção deste *dataset*, foi testada em um algoritmo de AGR já existente, o *Gender Net*, e analisado os efeitos do *dataset* com pessoas transgênero, na intenção de amenizar o *misgender* tecnológico. As métricas de avaliação de modelos de classificação foram utilizadas para compreender os efeitos dessa inclusão. A análise mostrou através dos valores das métricas que o modelo é eficiente para classificações realmente positivas e também é eficiente para classificações relevantes, que se tratando de um *dataset* incluindo pessoas transgênero, significa amenização de problemas de *misgender* tecnológico, pois estas pessoas estão presentes nas classificações e são tidas como positivas e relevantes.

Palavras-chave: Visão Computacional. Reconhecimento Automatizado de Gênero. Transgênero.

ABSTRACT

The techniques and applications of facial recognition are on the rise, becoming even more present in real life. Many of these applications are being created without a critical analysis of their workings and impacts on society. Automatic Gender Recognition (AGR) opens a door for technological misgendering, which affects transgender people's life. Research done on this type of technology do not include this slice of population. Additionally, the AGR delivers a worse performance in comparison with the identification of cisgender individuals, as stated. As such, this piece of research analyzed a dataset, which has transgender people in its composition. The analysis of this dataset, was tested on pre-existent AGR algorithms, the Gender Net and analyzed the effect of this dataset with transgender people to reduce the cases of technological misgendering. Evaluation metrics for classification models were used to understand the effects of this inclusion. The analysis showed through the values of evaluation metrics that the model is efficient for really positives classifications and relevant classifications, and when it comes to transgender people means decrease the technological misgendering problems, because these persons are present in positive and relevant classifications.

Keywords: Computer Vision. Automatic Gender Recognition, Transgender

LISTA DE ILUSTRAÇÕES

Figura 1 – Representação de uma Rede Neural Convolutacional, através das camadas de extração de características (convolução e <i>pooling</i>), a camada de conexão (<i>fully conected</i>) e a camada de classificação, onde a) mostra o fluxo do aprendizado a partir de um objeto e b) mostra o mesmo fluxo através do reconhecimento deste objeto.	22
Figura 2 – Representação da arquitetura da rede neural convolutacional proposta por Gil Levi e Tal Hassner no desenvolvimento do <i>Gender Net</i> e <i>Age Net</i>	24
Figura 3 – Fases e etapas do desenvolvimento do <i>dataset</i>	26
Figura 4 – Fase de aquisição de imagens	27
Figura 5 – Fase de pré-processamento	29
Figura 6 – Gráficos comparativo da evolução da precisão no processo de treinamento de cada grupo.	32

LISTA DE TABELAS

Tabela 1 – Configuração do computador utilizado	31
Tabela 2 – Tabela da rotina de treinamento do Adience benchmark	31
Tabela 3 – Tabela da rotina de treinamento das imagens coletadas	31

LISTA DE EQUAÇÕES

Equação 3.1 - Acurácia	23
Equação 3.2 - <i>Recall</i>	23
Equação 3.3 - Precisão	23
Equação 3.4 - <i>F-Score</i>	24

SUMÁRIO

	Lista de Equações	10
1	INTRODUÇÃO	12
2	OBJETIVOS	14
2.1	OBJETIVO GERAL	14
2.2	OBJETIVOS ESPECÍFICOS	14
2.3	JUSTIFICATIVA	14
3	REVISÃO BIBLIOGRÁFICA	16
3.1	DIVERSIDADE SEXUAL	16
3.1.1	Sexo Biológico	16
3.1.2	Orientação Sexual	16
3.1.3	Identities, Performances e Expressões de Gênero	17
3.1.4	<i>Misgender</i>	18
3.2	TECNOLOGIA	19
3.2.1	Visão computacional	19
3.2.2	Reconhecimento Facial	19
3.2.3	Reconhecimento Automatizado de Gênero	20
3.2.4	Aprendizado de Máquina: <i>Machine Learning</i> e <i>Deep Learning</i>	20
3.2.5	Os problemas de Aprendizado de Máquina: Regressão e Classificação	21
3.2.6	Redes Neurais Convolucionais (CNN)	21
3.2.7	Métricas de Avaliação de Modelos	22
3.3	TRABALHOS RELACIONADOS	24
4	DESENVOLVIMENTO	26
4.1	FASE DE AQUISIÇÃO DE IMAGENS	26
4.2	FASE DE PRÉ-PROCESSAMENTO	27
4.3	TREINAMENTO DA REDE NEURAL	29
4.4	AVALIAÇÃO DE MODELO	30
5	DISCUSSÃO DOS RESULTADOS	31
6	CONCLUSÃO	33
	REFERÊNCIAS	34

1 INTRODUÇÃO

A Inteligência Artificial (IA) tem tido grandes avanços e utilizações nos últimos anos, trazendo diversas ferramentas que expandem as tomadas de decisões e as limitações humanas. O uso desta tecnologia vem causando mudanças no modo de vida da sociedade, como por exemplo as redes sociais e os sistemas de sugestões, a área da saúde com as previsões de diagnósticos, o mercado financeiro e as previsões sobre as especulações de ações, a segurança pública e privada com os sistemas de vigilância, entre outros. Como exemplos deste uso pode-se citar as operações da polícia civil durante o carnaval a partir de 2019, utilizando sistemas de reconhecimento facial para identificação de criminosos entre os foliões (UOL, s.d.), e o próprio investimento federal em tais tecnologias conforme aprovado na portaria Nº 793 de 24 de outubro de 2019 no Diário Oficial da União (JUSTIÇA E SEGURANÇA PÚBLICA, 2019). Entretanto, há a necessidade de se discutir sobre o funcionamento dessas tecnologias e seus impactos sociopolíticos.

No âmbito da comunidade LGBTQIA+ (Lésbicas, Gays, Bissexuais, Travestis e Transgênero, *Queers*¹, Intersexuais, Assexuados e mais), as siglas "T" e "Q" representadas pelas pessoas Travestis, Transgênero e *Queers* têm sido excluídas, ignoradas ou negligenciadas com a criação e evolução dessas tecnologias, principalmente no reconhecimento automatizado de gênero (AGR - *Automatic Gender Recognition*), que utiliza de variadas técnicas (como reconhecimento facial e reconhecimento de voz) para reconhecer o gênero de uma pessoa e tomar decisões a partir disto.

As decisões reconhecimento automatizado de gênero pode ocasionar situações de preconceito de gênero, *misgender*. Este tipo de situação enfraquece a luta por direitos e reconhecimento das identidades destas pessoas.

Este não reconhecimento foi explicitado por Keyes (2018) em seu trabalho que analisou cerca de 58 artigos sobre algoritmos de AGR. Análise a qual mostrou que em nenhum dos artigos foram feitas menções a pessoas transgênero e mais de 90% deles não referenciavam ou diferenciavam os estudos e questionamentos de gênero e expressão de gênero.

Outro trabalho, que também faz uma análise sobre dos efeitos do AGR sob pessoas transgênero, é dos pesquisadores Scheuerman, Paul e Brubaker (2019), que analisa a eficiência das principais aplicações de AGR existentes de grandes empresas no mercado, como Google e Facebook. Esta análise também traz resultados não otimistas com relação a estas pessoas.

A não menção ou discussão dos trabalhos com relação aos estudos de gênero e de expressão de gênero refletem nas tecnologias e aplicações que utilizam do gênero e são voltadas para a sociedade, gerando mais exclusão a uma parcela de pessoas as quais estão em constante luta por reconhecimentos de suas identidades. Desta forma, cria-se um questionamento sobre a responsabilidade e a preocupação social ou até despreparo com as questões de humanidades pelos criadores de tais tecnologias e estudos. Assim como esses avanços podem trazer grandes

¹ Queer (em português, "excêntrico", "insólito") é um termo "guarda-chuva" proveniente do inglês, usado para designar pessoas que não seguem o modelo de heterossexualidade ou do binarismo de gênero (BORTOLETTO, 2019).

frutos à sociedade, revolucionando diversos modos de vida, praticidades e usabilidades, tais avanços sem a menor preocupação social podem se tornar um fator discriminatório na sociedade, gerando mais preconceito, exclusão e desmerecimentos de causas e lutas como as da comunidade LGBTQIA+.

Com isso, abre-se um grande espaço para questionamentos sobre efetividade ao trazer essas questões no desenvolvimento desses estudos e aplicações:

- Tais discussões tornariam essas tecnologias mais eficientes, inclusivas e mais próximas da sociedade?
- A inclusão de pessoas transgênero no *dataset* confundiria a aplicação final, piorando os resultados existentes?
- Seria necessário fazer novos ajustes dos algoritmos ou apenas incluir tais pessoas nos dados de treinamento da IA amenizariam estes casos?
- Como seria a melhor prática de se montar um *dataset* que minimizasse tais problemas?
- Existe a necessidade de um reconhecimento automatizado de gênero? Quais os impactos destes reconhecimentos?
- Quem são as pessoas que criam e desenvolvem essa tecnologia? Para quem estes recursos são pensados? Quais suas finalidades?

Partindo destes questionamentos este trabalho pretende analisar o funcionamento dos algoritmos de AGR em reconhecimento facial trabalhando e incluindo as questões de gênero e expressão de gênero no seu desenvolvimento.

2 OBJETIVOS

2.1 OBJETIVO GERAL

Fazer uma análise de um *dataset* com imagens de pessoas transgênero levando em consideração a construção de gênero de maior reconhecimento na sociedade (a binária: homem e mulher) sob a perspectiva da expressão de gênero, para analisar o comportamento do algoritmo *Gender Net* (LEVI; HASSNER, 2015) de AGR de reconhecimento facial utilizando as métricas de avaliação de modelos com a matriz de confusão.

2.2 OBJETIVOS ESPECÍFICOS

- Avaliar a eficiência e desempenho de um *datasets* que inclui pessoas transgênero para algoritmos de AGR de reconhecimento facial;
- Comparar o comportamento de um algoritmo de AGR para reconhecimento facial, com um *dataset* que possui pessoas transgênero;
- Estabelecer métricas para a criação de *dataset* que amenizem os problemas de *misgender* encontrados nos algoritmos de AGR.

2.3 JUSTIFICATIVA

A falta de preocupação com responsabilidade ética no desenvolvimento de tecnologias como o AGR, pode gerar problemas de *misgender* tecnológico e invalidando a luta por identidade e garantia de direitos das pessoas transgênero e toda comunidade LGBTQIA+.

Citado alguns possíveis cenários, em que a falta de preocupação com a responsabilidade ética do desenvolvimento desta tecnologia possa gerar algum tipo de problema. Começando por uma universidade que faz o cadastramento dos rostos dos estudantes na matrícula, para ser utilizado no reconhecimento facial em diversos serviços inteligentes internos, com alguns ligados ao gênero cadastrado e informado pelo estudante. Caso o(a) estudante durante seu período na universidade se reconheça por outro gênero e mude no seu cadastro, esses sistemas inteligentes com reconhecimento facial que contêm aplicações de AGR. Caso esses sistemas não sejam adaptativos a este estudante, possivelmente ocorrerão problemas de *misgender* tecnológico, todas as vezes que tal estudante for utilizar um dos sistemas que utilize AGR durante seu período de transição, até que chegue em algo "aceitável" para o sistema, indiretamente impondo um padrão expressivo de gênero.

Outro cenário, podemos citar um caixa eletrônico que utilize de um algoritmo de AGR para reforçar a sua segurança antifraudes, fazendo uma comparação com o gênero cadastrado do cliente com o gênero de quem está utilizando o caixa eletrônico. Um reconhecimento errado

pode acionar as políticas antifraudes do banco e trazer transtornos e constrangimentos a uma pessoa transgênero ou *queer*.

Também podemos ter como cenário, um aplicativo de relacionamentos, que utiliza de reconhecimento facial e AGR pra combater possíveis perfis falsos e mal-intencionados, mas com um reconhecimento errado de uma pessoa transgênero ou *queer*, pode acionar suas políticas e impedir com que essa pessoa tenha acesso ao aplicativo, dificultado e limitando as possibilidades dessa pessoa se relacionar com outras e ter relações afetivas.

Por último, em um cenário onde um governo contrate, de alguma empresa, um serviço de reconhecimento facial para monitorar sua população e neste serviço inclui AGR, mas tal governo tem vieses e posicionamentos políticos declaradamente contra pessoas LGBTQIA+, essa tecnologia, sem uma análise crítica em seu desenvolvimento e sem responsabilidade ética, colocaria em risco a vida de uma pessoa transgênero ou *queer*.

Por isso este trabalho, busca a partir desta análise crítica da responsabilidade ética do desenvolvimento do AGR, fazer uma análise do impacto que a inclusão de pessoas transgênero no *dataset* tem para o aprendizado de um algoritmo de AGR, o *Gender Net*.

3 REVISÃO BIBLIOGRÁFICA

3.1 DIVERSIDADE SEXUAL

Vivemos em um planeta com bilhões de pessoas espalhadas pelos quatro cantos da Terra, cada uma com sua história, cultura, individualidades, coletividades e contextos étnicos, sociais, identitários, econômicos e sexuais. Uma vez que os trabalhos de Os Keyes (2018) e Scheuerman, Paul e Brubaker (2019) apontam que as aplicações e os problemas envolvendo AGR entram em conflito direto com a diversidade sexual na sociedade, precisamos entender melhor sobre esses contextos e a diversidade existente neste tópico para podermos compreender esses conflitos.

A diversidade sexual é um tema da sexualidade humana, que é formada por uma combinação de múltiplos fatores que traduzem a individualidade de cada ser. Em um material desenvolvido numa parceria entre o Ministério da Saúde e Ministério da Educação Brasileiro, a explicação sobre a diversidade sexual é dada "[...] a partir de três eixos fundamentais, mas lembrando sempre que, em nossas vidas, esses fatores interagem de maneira dinâmica. São eles: o sexo biológico, a identidade de gênero e a orientação sexual. [...]"(SAÚDE; EDUCAÇÃO, 2020, p. 16).

3.1.1 Sexo Biológico

A composição do sexo biológico "[...] é constituído pelas características fenotípicas (órgãos genitais externos, órgãos reprodutores internos, mamas, barba) e genotípicas (genes masculinos e genes femininos) presentes em nosso corpo. [...]"(SAÚDE; EDUCAÇÃO, 2020, p. 16), ou seja, a transcrição de um conjunto de informações biológicas que distinguem um ser vivo através das informações cromossômicas, capacidades reprodutivas, características fisiológicas e órgãos genitais. Essa distinção biológica nos seres humanos se dá entre "[...] combinação dos cromossomos X e Y, existem somente dois sexos: XY produz um ser chamado de macho e XX, um ser chamado de fêmea [...]"(SAÚDE; EDUCAÇÃO, 2020, p. 16), mas também há o "intersexo", quando há uma mistura entre as características biológicas do "macho" e "fêmea". O que se define como características para classificação de um e outro varia de acordo com a espécie do ser vivo.

3.1.2 Orientação Sexual

A orientação sexual "[...] é uma atração espontânea e não influenciável que só pode ser conhecida plenamente pelo indivíduo que a vivencia [...]"(SAÚDE; EDUCAÇÃO, 2020, p. 17), bem como uma manifestação da afetividade e/ou atração sexual de um indivíduo, ou seja, um apontamento para quem este indivíduo involuntariamente sente desejo, atração sexual. No campo da orientação sexual existe uma pluralidade de orientações. Entretanto, existem quatro

tipos majoritários:

- Heterossexual: Pessoa que se sente atraída por pessoas do gênero oposto ao seu.
- Homossexual: Pessoa que se sente atraída por pessoas do mesmo gênero ao seu, sendo Gay quando se trata de um Homem que se sente atraído por um outro Homem e Lésbica quando se trata de uma Mulher que se sente atraída por outra Mulher.
- Bissexual: Pessoa que se sente atraída por pessoas, sem distinção dos gêneros.
- Assexual: Pessoa que não sente atração sexual, independente do gênero, entretanto essas pessoas podem desenvolver afetividade romântica com outras pessoas e, ainda assim, não sentir desejo e/ou atração sexual.

3.1.3 Identidades, Performances e Expressões de Gênero

Por mais que na dimensão biológica haja uma divisão entre "macho" e "fêmea" nas suas características e comportamentos, na dimensão social estas características e comportamentos são expressas pela cultura da sociedade. Sendo assim, as definições, papéis, performances e formas de expressão dos símbolos/signos relacionados a um gênero são produtos da realidade cultural e não necessariamente da anatomia de seus corpos.

- Cisgeneridade: Indivíduos que se identificam com os padrões sociais pré-estabelecidos, performances, símbolos/signos de um gênero concordantes com o seu sexo biológico. Por exemplo: um indivíduo que nasceu com o sexo feminino e se identifica com os padrões do gênero feminino e suas performances estabelecidas na sua sociedade (como vestimentas e papéis sociais).
- Transgeneridade: Indivíduos que não se identificam com os padrões sociais pré-estabelecidos, performances, símbolos/signos de um gênero concordantes com o seu sexo biológico. Exemplo: um indivíduo que nasceu com o sexo feminino, porém não se identifica com os padrões do gênero feminino e suas performances estabelecidas na sociedade (como vestimentas e papéis sociais) e se identifica e se reconhece com os padrões do gênero masculino.
- *Queers*: este segmento gira em torno de uma teoria de discussões do questionamento identitário das imposições de gênero, performances e sexualidades existentes, como uma luta política a partir da própria existência como explicado por Richard Miskolci:

A nova política de gênero - que também pode ser chamada de queer - se materializa no questionamento das demandas feitas a partir dos sujeitos; em outras palavras, chama a atenção para as normas que os criam. Essa mudança de eixo na luta política se fundamenta em duas concepções distintas com relação à dinâmica das relações de poder: uma que as compreende a partir da visão do poder como algo que opera para repressão, e outra que o concebe como mecanismos sociais

disciplinadores. Na perspectiva do poder opressor, os sujeitos lutam contra o poder por liberdade, enquanto na do poder disciplinar, a luta é por desconstruir as normas e as convenções culturais que nos constituem como sujeitos [...] (MISKOLCI, 2017, cap. 1, p. 27)

Essas discussões de poder, normas e convenções culturais dizem respeito à concordância e identificação com os padrões de gênero. No caso de pessoas não-binárias há discordância e não identificação com ambos os padrões de gênero, porém os reconhecem. No caso de pessoas agênero há a discordância e não identificação com ambos os padrões de gênero sem auto reconhecimento de um gênero, inexistência de um gênero; também há a concordância e identificação parcial, fragmentada e transitória entre os padrões de gênero existentes como o caso de pessoas de gênero fluído.

Todas essas construções em cima de gênero que são externalizadas são chamadas de expressões de gênero, que é a reprodução visual dos símbolos, signos e comportamentos que foram atrelados a um gênero. Por exemplo: o uso de vestidos e saias estão em maior parte atrelados ao gênero feminino; sendo assim, indivíduos (independente do seu sexo) que se utilizam de tais símbolos e signos, muitas vezes são lidos como feminino (como Drag Queen) ou tachados como apropriador do feminino (no caso de homens gays afeminados ou uma pessoa transgênero que possua traços tidos como masculinos).

A expressão de gênero também é utilizada como forma de evidenciar que os gêneros são performances através dos símbolos, signos e reprodução de comportamentos. Há um movimento que traz o questionamento dessa performance de gêneros que é a arte *Drag Queen*, na qual um indivíduo se utiliza da expressão de um gênero que acaba, por muitas vezes, colocando em cheque o gênero de quem performa pelas pessoas que veem. Por exemplo: as diversas ocasiões na mídia, em que artistas Drag Queens como Pablló Vittar são lidas como uma mulher, devido a sua performance de gênero (roupas, peruca, maquiagem, comportamento e etc.), por mais que Pablló Vittar se intitule como "um homem de peruca".

3.1.4 *Misgender*

Conforme livre tradução do dicionário Cambridge de língua inglesa, *misgender* é "o uso de pronome errado ou outra palavra de especificação de gênero ao referir-se ou falar a alguém, especialmente as pessoas transgênero"(MISGENDER. . . , 2020). Além da definição, *misgender* refere-se a situações e experiências que vão além dos pronomes de tratamento. Como explicado pelo pesquisador Kevin A. McLemore, a prática do *misgender* acontece de formas mais sutis, que são vivenciadas, ou seja, experiências que moldam as identidades sociais e a forma como a pessoa se sente. O autor segue especificando que além do tratamento de gênero essas experiências incluem "[...] negar acesso a algum espaço de gênero de preferência de

alguém [...]"(MCLEMORE, 2015), como exemplo negar ou impedir o acesso ao banheiro por questões de gênero.

No caso da tecnologia, o reconhecimento de gênero é um fator determinante para o funcionamento da aplicação, como um sistema de reconhecimento facial de um caixa eletrônico, um sistema de reconhecimento facial anti-fraude do transporte público, o sistema de reconhecimento facial de anti-fraude dos aplicativos de relacionamento e entre outros. Os erros de reconhecimento facial podem causar situações de constrangimento ou invalidação da identidade do indivíduo, que podem ser classificados como *misgender* tecnológico.

3.2 TECNOLOGIA

Da pedra lascada ao fogo, à roda, ao tecido, ao gênero, à foice e o martelo, à máquina a vapor, ao manuseio da eletricidade, ao software e à Inteligência Artificial. Tecnologia é o conceito que envolve ferramentas, estruturas e processos, que são utilizados como forma de revolucionar o modo de vida do ser humano. Há diversas formas de tecnologia, com diversas funções, mas entre as existentes, ao pensar em interpretação do mundo ao nosso redor através dos computadores, temos a visão computacional.

3.2.1 Visão computacional

Uma excelente descrição do que é visão computacional, entre tantas áreas existentes dentro da computação, é a feita por Richard Szeliski em seu livro *Computer Vision: Algorithms and Applications*, em que ele descreve esta área através do que ela faz, "[..] uma forma de descrever o mundo que vemos, através de uma ou mais imagens e reconstrução de suas propriedades através de formas, iluminação e distribuição de cores [...]"(SZELISKI, 2011, p. 5). É através da visão computacional que podemos interpretar o mundo real com computadores, com o auxílio de imagens, vídeos, câmeras e sensores. Além disso, ela é uma área com múltiplos tópicos, dentre eles processamento de imagem, detecção e reconhecimento de objeto e reconhecimento facial.

3.2.2 Reconhecimento Facial

O reconhecimento de objetos para Szelisk é "[...] entre tantas tarefas visuais que mandamos os computadores performar, analisar uma cena e reconhecer os objetos que a constituem talvez seja das mais desafiadoras [...]"(SZELISKI, 2011, p. 657), mas porquê deste desafio? O autor explica que as diferenças de poses, variações de classes de objetos (como diferentes raças de cachorros) tornam essa tarefa mais complexa, necessitando exaustivas comparações num banco de dados. É justamente isso o que torna o reconhecimento facial, entre tantas tarefas de reconhecimento de objetos, algo que Richard chama de "[...] discutivelmente a que tem o maior

sucesso [...]”(SZELISKI, 2011, p. 668), por lidar apenas com rostos, as variações dos parâmetros (poses, classes, cores e etc.) são menores, o que torna essas atividades menos complexas. Mas por que "discutivelmente" a que tem o maior sucesso? Richard não entra em detalhes sobre o "discutivelmente", mas uma interpretação possível é justamente o que engloba este trabalho, ou seja, temos uma técnica muito bem sucedida em execução, porém com aplicações questionáveis, como o reconhecimento automatizado de gênero.

3.2.3 Reconhecimento Automatizado de Gênero

As aplicações de reconhecimento facial utilizam-se das propriedades e características de um rosto para identificar sujeitos. Além disso, também é possível utilizar dessas características para extrair informações como raça, idade e gênero.

Segundo Os Keyes, o reconhecimento do gênero é a proposta dos Algoritmos de Reconhecimento Automatizado de Gênero (*automated gender recognition - AGR*), ou seja, identificar o gênero de uma pessoa através de características, de acordo com as construções e definições de gênero daquele modelo e de como tais características estão expressas e externalizadas, de forma automatizada e computacional utilizando imagens ou vídeos.

É através desta forma automatizada e computacional, que as técnicas de aprendizado de máquina entram como ferramenta que possibilita esse aprendizado computacional e reconhecimento automatizado.

3.2.4 Aprendizado de Máquina: *Machine Learning* e *Deep Learning*

Quando falamos sobre o aprendizado computacional estamos nos referindo às técnicas de aprendizado de máquina onde os autores Moacir A. Ponti e Gabriel B. Paranhos da Costa referem-se ao aprendizado de máquina (do inglês *Machine Learning*, ML) através da literatura como métodos que "[...] buscam diretamente por uma única função que possa, a partir de um conjunto de parâmetros, gerar o resultado desejado.[...]"(PONTI; COSTA, 2017, ch. 3, p. 66). Em outras palavras, uma única função $f(.)$ que define através dos seus parâmetros, os padrões do conjunto dados (*dataset*) onde são possíveis separar por classes ou realizar previsões em cima dos dados. Por outro lado, o aprendizado profundo (do inglês *Deep Learning*, DL) segundo os autores "[...] temos métodos que aprendem $f(.)$ por meio da composições de funções [...]"(PONTI; COSTA, 2017, cap. 3, p. 66), ou seja, cada uma dessas funções serão chamadas de "camadas" e tais camadas têm seus próprios conjuntos de parâmetros que realizam uma transformação dos dados recebidos e geram uma saída de acordo com os parâmetros da próxima camada. Dessa forma os autores transcrevem o aprendizado profundo como $f(x) = f_L(\dots f_2(f_1(x_1, W_1); W_2)\dots), W_L)$, onde x_1 refere-se aos dados de entrada e W_L os parâmetros relacionados a cada função f_L .

As utilizações com *Deep Learning* em Visão Computacional, principalmente, têm tido

resultados muito melhores do que as utilizações com *Machine Learning*. Ainda não se sabe ao certo os reais motivos dessas melhorias dos resultados, mas sua utilização neste segmento é muito crescente. E ao falar sobre *Machine Learning* é preciso entender que há dois tipos de problemas que englobam esta área, que são os problemas de Regressão e Classificação.

3.2.5 Os problemas de Aprendizado de Máquina: Regressão e Classificação

O problemas de regressão, segundo Duarte e Ståhl, são problemas "[...] onde os valores prescritos são contínuos. Prevendo taxas de câmbio de moedas, temperatura e tempo, onde um evento pode vir a ocorrer [...]"(DUARTE; STÅHL, 2019). Portanto os problemas de regressão são problemas onde o intuito é poder identificar um padrão de comportamento dos conjuntos de dados (*dataset*) que permita que valores possam ser previstos e antecipados.

Já os de Classificação, segundo os mesmos autores, "[...] são problemas similares aos problemas de regressão; a diferença é que os descritores são valores discretos"(DUARTE; STÅHL, 2019). Com isso, podemos fazer a representação dos dados através de classes, logo aprendendo seus padrões e separando os dados entre tais classes, como por exemplo o reconhecimento de raças de cachorros, marcas de carros, transações bancárias suspeitas ou até o reconhecimento do gênero de um indivíduo através do reconhecimento facial.

Cada um desses tipos de problemas tem seus próprios conjuntos de algoritmos com suas utilizações para determinadas situações. Em problemas de classificações onde há a necessidade de classificar elementos visuais como imagens e vídeos, os algoritmos mais utilizados têm sido as Redes Neurais.

3.2.6 Redes Neurais Convolucionais (CNN)

As Redes Neurais que são algoritmos de *Deep Learning*. Dentre os modelos de redes mais utilizados, temos as Redes Neurais Convolucionais (*Convolutional Neural Networks - CNN*) "[...] O que caracteriza esse tipo de rede é ser composta basicamente de camadas convolucionais, que processa as entradas considerando campos receptivos locais. Adicionalmente inclui operações conhecidas como *pooling*, responsáveis por reduzir a dimensionalidade espacial das representações [...]"(PONTI; COSTA, 2017, cap. 3, p. 74).

Pode-se observar a estrutura de uma CNN na figura 1, que mostra a combinação de varias camadas de convolução que aplicarão filtros na imagem através de uma combinação linear dos *pixels* vizinhos e essa combinação de várias camadas de convolução vão revelando as características das imagens, o que a literatura chama de mapas de características (*feature map*).

Entre as camadas de convolução temos as camadas de *pooling*, que são responsáveis por fazer a redução das dimensões espaciais das camadas e assim também fazer a redução do tamanho das imagens com o propósito da redução do custo computacional, já que tais operações de convolução tendem a escalonar conforme a profundidade da rede.

Por fim, temos a camada *fully connected* que é a camada que vetoriza o resultado das convoluções e *pooling*, gerando o mapa de características que será utilizado como a entrada de dados pra o algoritmo classificador.

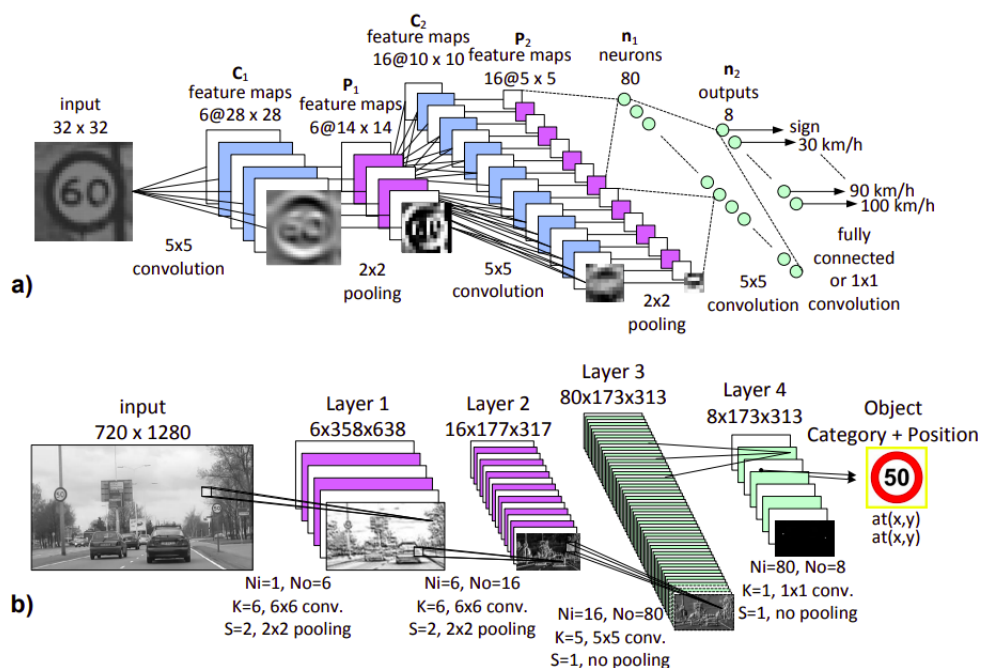


Figura 1 – Representação de uma Rede Neural Convolucional, através das camadas de extração de características (convolução e *pooling*), a camada de conexão (*fully connected*) e a camada de classificação, onde a) mostra o fluxo do aprendizado a partir de um objeto e b) mostra o mesmo fluxo através do reconhecimento deste objeto.

Fonte: Maurice Peemen (PEEMEN et al., 2016)

Porém para saber a eficiência do nosso aprendizado e classificação, onde o resultado final chamaremos de modelo, é preciso a utilização de métricas que irão descrever a eficiência deste modelo.

3.2.7 Métricas de Avaliação de Modelos

Como já mencionado anteriormente, cada tipo de problema de aprendizado de máquina tem seus próprios conjuntos de algoritmos com suas utilizações para determinadas situações. Esses algoritmos também possuem diferentes formas de serem avaliados para ter um entendimento do seu funcionamento e aprendizado, que indicam se ele está fazendo aquilo que foi proposto para fazer.

Para os problemas de classificação, que é o tipo de problema que este trabalho engloba, normalmente se utiliza a matriz de confusão para a obtenção dos dados do funcionamento de um modelo. A Matriz de confusão é uma tabela que permite com que possamos identificar a frequência de classificação através de classes do modelo. Essas classes são:

- Verdadeiro Positivo (*true positive* - TP): Essa classe indica quando um conjunto real é classificado de forma corretamente. Exemplo: quando em uma aplicação de AGR o gênero reconhecido é o gênero o qual o indivíduo se identifica;
- Falso Positivo (*false positive* - FP): Nesta classe é quando o um conjunto real não é classificado de forma correta. Por exemplo: quando em uma aplicação de AGR o gênero reconhecido não é o gênero o qual o indivíduo se identifica (categorizando um *misgender*);
- Verdadeiro Negativo (*true negative* - TN): Os verdadeiros negativos são os conjuntos reais que são representados pelos conjuntos que não estamos buscando prever. Exemplo: em uma aplicação de AGR binária (que classifica apenas entre Homem e Mulher) um indivíduo que não se identifica como Homem e o modelo não o classificou como Homem;
- Falso Negativo (*false negative* - FN): Essa classe indica quando um conjunto real é classificado de forma incorreta. Exemplo: quando em uma aplicação de AGR o gênero reconhecido não é o gênero o qual o indivíduo se identifica.

Através da coleta quantitativa dessas classes, podemos obter as seguintes através de equações, métricas para realizar a avaliação e descrever a eficiência e comportamento do modelo. Tais equações são:

- Acurácia: Esta métrica diz o quão assertivo foi o modelo dentro das previsões possíveis, ou seja, as previsões que são verdadeiras independente de serem verdadeiras positivas ou verdadeiras negativas. Conforme apresentado na equação 3.1, o valor é obtido através do somatório das classes verdadeiro positivo com verdadeiro negativo, dividido pelo somatório de todas as classes.

$$accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (3.1)$$

- *Recall*: O recall descreve o quão bom é o modelo para as proporções positivas. Conforme apresentado na equação 3.2, o valor é obtido através da classe verdadeiro positivo dividida pelo somatório verdadeiro positivo com falso negativo.

$$recall = \frac{TP}{TP + FN} \quad (3.2)$$

- Precisão: Esta métrica descreve o quão bom é um modelo, através da descrição de o quão bom o modelo é para classificações realmente positivas. Conforme apresentado na equação 3.3, o valor é obtido através da classe verdadeiro positivo dividida pelo somatório verdadeiro positivo com falso positivo.

$$precision = \frac{TP}{TP + FP} \quad (3.3)$$

- *f-Score*: É o balanceamento entre a precisão e o *recall*, uma média harmônica, pontuando o desempenho agregado na avaliação do modelo. Por exemplo, em um modelo de AGR, que enquanto a precisão identifica os acertos da identificação de um gênero, o *recall* identifica a relevância desses acertos. Com isso, o *f-Score* traz esse balanço para avaliar este desempenho do modelo. Quanto maiores forem os valores de precisão e *recall*, maiores serão os valores do *f-Score*, demonstrando o quão bom é o modelo nas classificações e nas relevâncias dessas classificações. Conforme apresentado na equação 3.4, o valor é obtido através da multiplicação do valor de precisão pelo valor de recall, dividido pelo somatório do valor de precisão com o valor de recall, com o resultado da divisão, multiplicado por dois.

$$f - score = 2 * \frac{precision * recall}{precision + recall} \quad (3.4)$$

3.3 TRABALHOS RELACIONADOS

Um dos algoritmos de AGR que tem grande difusão e fácil acesso é o *Gender Net*, desenvolvido e publicado por Gil Levi e Tal Hassner (2015) com a proposta de junto ao *Age Net* fazer um reconhecimento de gênero e idade através de uma rede neural convolucional utilizando o *Audience Benchmark* (HASSNER et al., 2015).

Com uma arquitetura simples, como observado na figura 2, o *Gender Net* (LEVI; HASSNER, 2015) possui três camadas de convoluções, seguindo por mais duas camadas de *Fully Connected*.

Em sua execução, utilizando o algoritmo SVM como classificador, o *Gender Net*(LEVI; HASSNER, 2015) obteve bons resultados de classificação, sendo aproximadamente 84,6% de acurácia para a classificação de gênero. Com isso esse estudo e os modelos gerados foram publicados e disponibilizados de forma gratuita e de código aberto para qualquer pessoa que quisesse utilizar.

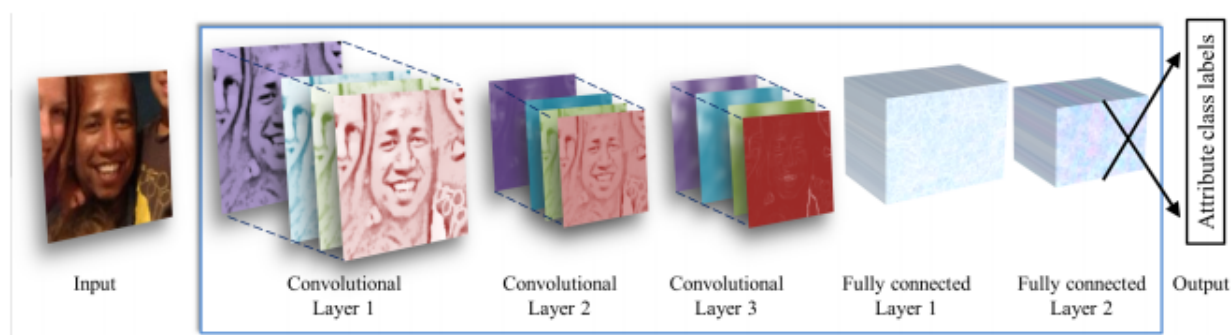


Figura 2 – Representação da arquitetura da rede neural convolucional proposta por Gil Levi e Tal Hassner no desenvolvimento do *Gender Net* e *Age Net*.

Fonte: Levi e Tal Hassner (2015)

Em 2018 o pesquisador Os Keyes publicou um trabalho intitulado *The Misgendering Machines: Trans/HCI Implications of Automatic Gender Recognition*, onde o autor faz uma análise de diversos estudos que tratam do reconhecimento automatizado de gênero. Para entender como tais estudos trabalham com as questões de gênero, Keyes identificou que nenhum dos trabalhos, se quer cita pessoas transgênero e tratam o gênero apenas como algo binário (homem ou mulher). A grande maioria trata como algo imutável e mais da metade, como sendo algo descritivo por características fisiológicas. Tais resultados mostram o descaso do desenvolvimento científico em relação às pessoas transgênero, desmerecendo suas identidades e invisibilizando suas existências.

No ano seguinte, um grupo de pesquisadores inspirados pelo trabalho de Keyes (2018), publicaram um trabalho intitulado *How Computers See Gender: An Evaluation of Gender Classification in Commercial Facial Analysis Services* (2019). Este trabalho teve como objetivo utilizar os principais serviços de AGR disponíveis no mercado por grandes empresas como Google, Facebook, Amazon, IBM, Microsoft, Clarifai, Face++ e entre outras. Para testar e analisar esses serviços com um *dataset* com pessoas transgênero e *queers*, através de uma coleta de imagens com as *tags*: *woman, man, transwoman, transman, agender, genderqueer e nonbinary*, definidas através de uma pesquisa. Com esse teste, Scheuerman, Paul e Brubaker (2019) perceberam uma diferença bem delimitada de precisão entre pessoas cisgênero e transgênero. Tais resultados reforçam a não preocupação com as questões de gênero apontadas por Keyes (2018), não só na academia como também no mercado.

Este descaso com as pessoas transgênero tanto por parte da academia quanto do mercado, deixam evidente que a tecnologia não é um campo neutro, mas sim um campo com vieses muito bem estabelecidos por aqueles que os criam. Neste caso são vieses que invisibilizam e marginalizam ainda mais um grupo social que vem lutando muito contra a invisibilização e marginalização. Segundo dados da organização Transgender Europe, através do projeto *Trans-respect versus Transphobia Worldwide* (TVT; TGEU, s.d.), foram reportados assassinatos de 3664 pessoas transgêneros e *queers* entre primeiro de janeiro de 2008 a 30 setembro de 2020, sendo 1520 dos caso ocorridos no Brasil.

Através dessa reflexão, quais seriam os impactos de uma pequena mudança que através de uma análise crítica que contraponha os vieses deste tipo de tecnologia (como incluir imagens de pessoas transgênero no *dataset* para o processo de aprendizado) teria em uma aplicação de AGR já conhecida e estabelecida como o *Gender Net* (LEVI; HASSNER, 2015)?

4 DESENVOLVIMENTO

Este trabalho consiste em realizar uma análise na inclusão de pessoas transgênero no *dataset* para algoritmos de AGR em reconhecimento facial, com o objetivo de amenizar os problemas de *misgender* acarretados por estes algoritmos. Para isso foi utilizado o algoritmo de AGR *Gender Net* (LEVI; HASSNER, 2015), por ser um dos mais conhecidos, de grande difusão e por ser de código aberto.

A construção deste trabalho foi dividida em quatro fases, contendo etapas que formam aquela fase. Cada etapa possui um objetivo, que é complementar e pré-requisito ao desenvolvimento da etapa seguinte, definindo assim um passo a passo estrutural para realização e reprodução deste trabalho como pode ser visto na figura 3, que ilustra todas as fases e etapas realizadas neste trabalho e como cada etapa depende uma da outra para acontecer.

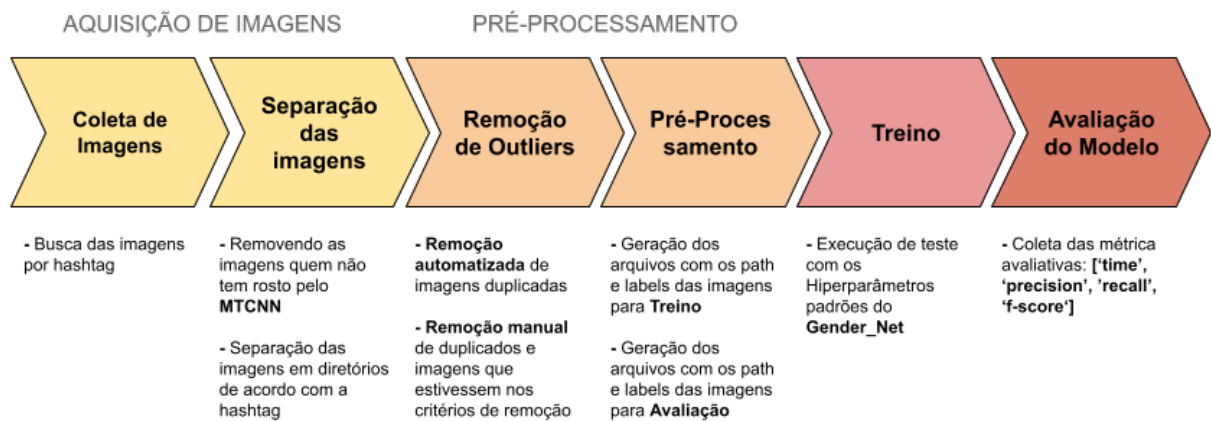


Figura 3 – Fases e etapas do desenvolvimento do *dataset*

Fonte: própria

4.1 FASE DE AQUISIÇÃO DE IMAGENS

Nesta fase o objetivo principal é fazer a coleta das imagens, que serão utilizadas no processo de treinamento e teste da rede neural, e separá-las de acordo com a rotulagem para a classificação, bem como o processo ilustrado pela figura 4. Para isso, foram definidas duas etapas para esta fase: de coleta de imagens e separação de imagens.

A primeira etapa é a de coleta de imagens, onde foi feita a pesquisa e o *download* das imagens inspirado pelo trabalho de Scheuerman, Paul e Brubaker (2019) utilizando as *hashtags*: *woman*, *man*, *transwoman*, *transman*. As imagens foram agrupadas em um diretório central, onde os nomes das imagens serviram como metadado para registrar a data e horário de coleta e a *hashtag* a qual pertencia.

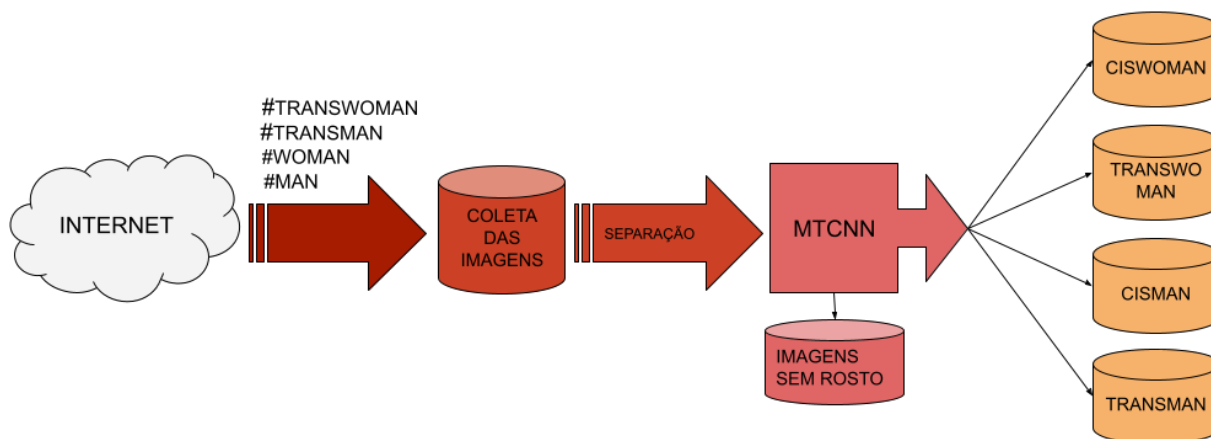


Figura 4 – Fase de aquisição de imagens

Fonte: própria

A etapa final da fase de aquisição de imagens é a separação das imagens, onde o objetivo é separar em diretórios as imagens de acordo com a *hashtag* a qual a imagem foi coletada. Antes da imagem ser separada, ela passa por um filtro, um detector facial para remoção de imagens as quais não possuem algum rosto, que será utilizado para a classificação. Para esse filtro, foi utilizado o detector facial MTCNN, que é uma implementação de detecção facial baseada no artigo de Zhang (ZHANG et al., 2016) feita por Iván de Paz Centeno. Este algoritmo possui uma boa precisão no seu processo de detecção facial, é *open source* e de fácil integração ao projeto.

Feita a coleta e separação das imagens, o objetivo da fase de aquisição de imagens é concluído, onde pode-se observar a representação na figura 4 que apresenta visualmente todos os processos envolvidos na aquisição de imagens.

4.2 FASE DE PRÉ-PROCESSAMENTO

Na fase de pré-processamento, o objetivo é fazer o preparo das imagens coletadas e separadas, para que estejam prontas para utilização pela rede neural, como ilustrado pela figura 5. Esta preparação consiste na seleção de imagens apropriadas para a classificação e alinhadas com o objetivo do aprendizado, que, neste caso, são fotos de pessoas que estejam com seus rostos à mostra. Depois é feita a separação dessas imagens entre as que serão utilizadas para o treinamento e as que serão utilizadas para validação. Por isso, essa fase foi dividida em duas etapas: a de remoção de *outliers* e a de pré-processamento.

A etapa de remoção de *outliers* tem como objetivo vasculhar as imagens coletadas e remover aquelas, que de alguma forma, possam atrapalhar o aprendizado da rede neural. O que incluem como critério para remoção:

- Imagens repetidas;

- Imagens com baixa qualidade;
- Imagens em que o rosto da pessoa esteja muito afastado;
- Imagens que possuam muitas pessoas na mesma imagem, principalmente casais;
- Imagens em que há cortes, como imagens com o rosto pela metade;
- Imagens em que há adereços que cubram o rosto, como o caso de máscaras;
- Imagens que passaram pelo detector facial da fase anterior, e mesmo assim, não havia algum rosto;
- Imagens que estejam usando filtros de redes sociais, que deformam ou adicionem elementos ao rosto, como por exemplo: nariz e orelha de animais;
- *Prints* de postagem de redes sociais;
- Ilustrações.

O primeiro passo dessa etapa foi fazer a remoção automatizada das imagens repetidas. Para isso, foi feito um algoritmo que para cada imagem coletada, calculava-se a média aritmética entre dos pixels da imagem, passando pelos canais de cores (RGB - *Red, Blue, Green*), pois este valor é único para cada imagem que representa suas cores, tonalidades e dimensões. Assim, comparando essa média com a média de outra imagem, ao encontrar uma média com o mesmo valor, temos uma imagem repetida.

Entretanto fazer essa comparação de cada imagem com todas as demais é um processo computacionalmente muito custoso, sendo na ordem de complexidade de $O(N^{N-1})$. Para otimizar esse processo, as imagens foram renomeadas com o valor do somatório da médias aritméticas dos canais de cores, pois quando uma imagem possuía o mesmo nome, ou seja, o mesmo valor de média, o sistema operacional automaticamente removia as imagens repetidas sem haver a necessidade de comparação, fazendo com que a complexidade caísse para $O(N)$.

No segundo passo dessa etapa de remoção de *outliers*, foram realizadas as remoções manuais das imagens, levando em consideração os critérios de remoção citados anteriormente e a remoção de imagens repetidas que passaram despercebidas pelo processo de remoção automatizada, por haver diferenças mesmo que mínimas, de tonalidade, contraste, iluminação, saturação, elemento como marca d'água, mudanças mínimas de ângulo e/ou cortes, pois a mínima mudança que seja, alterava o valor de média das imagens.

Com a conclusão desta etapa, o total coletado foi de 8.816 imagens, sendo: 2.453 imagens de homens cisgênero, 2.213 imagens de mulheres cisgênero, 1.651 imagens de homens transgênero e 2.499 imagens de mulheres transgênero. Em comparação com o *Adience benchmark* (HASSNER et al., 2015) *dataset* utilizado pelo *Gender Net* (LEVI; HASSNER, 2015) originalmente, que possui um total de 12.232 imagens, a quantidade coletada e as proporções de imagens foram consideradas uma boa quantidade e de proporções equilibradas, tendo em vista

que não há uma valor de proporção de pessoas transgênero e cisgênero, pois nenhum país incluiu em seus senso algo que pudesse obter esse valor e, atualmente, os únicos dados quantitativos relacionados às pessoas transgêneros são dados ligados à mortalidade e consumo de pornografia associados a essas pessoas.

Com o conjuntos de imagens já preparado após a remoção de *outliers* para ser utilizado no aprendizado, a última etapa da fase de pré-processamento é fazer a separação dessas imagens em dois conjuntos: as imagens que serão utilizadas para o treinamento e aprendizado da rede neural e as imagens que serão utilizadas para fazer a avaliação e teste da eficiência deste aprendizado.

Para isso, foi utilizado a função de *Train Test Split* do *Scikit Learn*, uma biblioteca de *Machine Learning* de código aberto que faz a separação de forma randômica de todo o conjunto de dados para o conjunto de treino e o conjuntos de teste, nas proporções padrões de 75/25: 75% do conjunto para treino e 25% do conjunto para teste. Na sequência, foi gerado um arquivo de texto para cada um dos conjuntos (treino e teste), contendo por linha o caminho da imagem e a rotulagem da imagem para aprendizado, sendo 0 ou 1: 0 para homens cisgênero e transgênero e 1 para mulheres cisgênero e transgênero.

Desta forma, como visto na figura 5 que ilustra todas as etapas envolvidas na fase de pré-processamento, todo o conjunto de imagens está preparado para ser utilizado no aprendizado no algoritmo de AGR, o *Gender Net* (LEVI; HASSNER, 2015), nas fases de treinamento e avaliação de modelo.

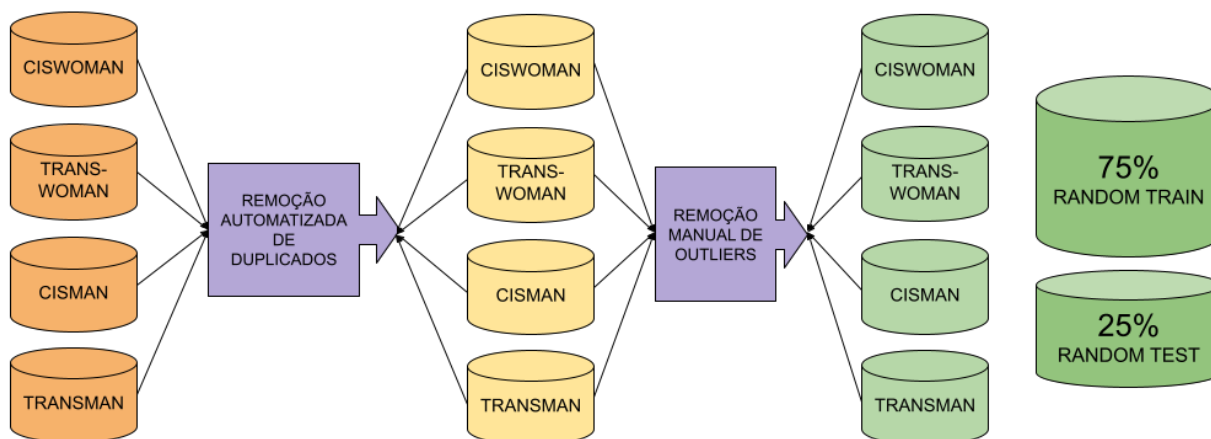


Figura 5 – Fase de pré-processamento

Fonte: própria

4.3 TREINAMENTO DA REDE NEURAL

A fase de treinamento é uma fase de etapa única, onde foi baixado os resultados do *Gender Net* disponibilizado pelos autores (LEVI; HASSNER, 2015); baixado a implementação do trabalho por Daniel Pressel utilizando *Tensorflow*, a biblioteca de código aberto de *machine learning* do Google na versão um, que possibilita a reprodução do *Gender Net* com um conjunto

de imagens próprio; baixado o *Adience benchmark* (HASSNER et al., 2015) *dataset* original do *Gender Net* (LEVI; HASSNER, 2015).

Com todos esses arquivos e dados, a rotina de treinamento foi montada da seguinte forma:

- Foram utilizados os hiperparâmetros padrões da implementação por Daniel Pressel: tamanho da imagens sendo de 227x227 pixels; taxa de aprendizado (*learning rate*): 0.01; épocas por decaída (*steps per decay*) de 10.000 épocas (número de épocas antes do *learning rate* decair); tamanho do *batch* por 128x128 pixels; otimizador (*optimizer*) sendo *Momentum*; função de perda (*loss function*) sendo entropia cruzada (*cross entropy*) utilizando *sparse softmax* com *logits*;
- Foram feitos treinamentos com os grupos de imagens do *Adience benchmark* (HASSNER et al., 2015) e o grupo de imagens coletadas;
- Para visualizar a evolução do aprendizado, foi feito um treinamento com cada grupo de imagens (*Adience benchmark* (HASSNER et al., 2015) e imagens coletadas, nas seguintes quantidades de épocas de aprendizado: 1.000 épocas, 2.000 épocas, 4.000 épocas, 8.000 épocas, 16.000 épocas e 32.000 épocas.

Com a execução dessa rotina de treinamento, foi gerado um modelo de aprendizado para cada um dos grupos de imagens em cada uma das quantidades de épocas de aprendizado. Estes modelos foram utilizados para realização da fase seguinte, a fase de avaliação de modelo.

4.4 AVALIAÇÃO DE MODELO

A fase de avaliação de modelo também é uma fase de etapa única, onde foi executado o processo de avaliação dos modelos gerados na fase anterior com o *Gender Net* (LEVI; HASSNER, 2015), para cada grupo de imagem e quantidade de épocas de aprendizagem, extraindo as métricas de Precisão, *Recall* e *F-Score* através do processo de avaliação da implementação do *Gender Net* (LEVI; HASSNER, 2015) por Daniel Pressel.

5 DISCUSSÃO DOS RESULTADOS

Todo o processo, desde a aquisição das imagens aos treinamentos, foi feito em um computador com as configurações presentes na tabela 1.

Sistema Operacional	Elementary OS 5.1.7 Hera (Linux 4.15.0-139-generic)
Processador	Quad Core Intel® Core™ i7-8550U CPU @ 1.80GHz
Chip Gráfico	Intel Corporation UHD Graphics 620 (rev 07)
Memória RAM	8,0 GB memória
Armazenamento Ocupado	19,4 GB

Tabela 1 – Configuração do computador utilizado

Após a rotina de treinamento esquematizada e executada, foram obtidas as métricas de avaliação de modelos: precisão, *recall* e *f-score* de cada um dos treinamentos.

O primeiro grupo de treinamento foi o grupo do *Adience benchmark* (HASSNER et al., 2015) por ser o *dataset* usado originalmente pelo *Gender Net* (LEVI; HASSNER, 2015) e pode-se observar na tabela 2 os resultados das métricas obtidas deste treinamento.

Adience benchmark						
	1K ÉPOCAS	2K ÉPOCAS	4K ÉPOCAS	8K ÉPOCAS	16K ÉPOCAS	32K ÉPOCAS
TEMPO DE EXECUÇÃO (min)	62	175	270	544	985	2160
PRECISION	0,553	0,553	0,604	0,797	0,866	0,881
RECALL	1,000	1,000	1,000	1,000	1,000	1,000
F-SCORE	0,712	0,712	0,753	0,887	0,928	0,937

Tabela 2 – Tabela da rotina de treinamento do Adience benchmark

De maior destaque dos valores de métricas obtidos, pode-se observar a evolução crescente da precisão e sua estabilização, já que o aprendizado de 16.000 épocas e 32.000 épocas possuem uma grande quantidade de épocas, e uma pequena evolução das precisões e *f-score* demonstrando estabilização do aprendizado na faixa dos 88% de precisão.

A estabilização da evolução do aprendizado também está presente no treinamento do grupo de imagens coletadas e com os valores das métricas menores, sendo 77,9% observável na tabela 3.

Imagens coletadas						
	1K ÉPOCAS	2K ÉPOCAS	4K ÉPOCAS	8K ÉPOCAS	16K ÉPOCAS	32K ÉPOCAS
TEMPO DE EXECUÇÃO (min)	67	145	244	493	985	2145
PRECISION	0,464	0,464	0,566	0,713	0,766	0,779
RECALL	1,000	1,000	1,000	1,000	1,000	1,000
F-SCORE	0,634	0,634	0,722	0,833	0,868	0,876

Tabela 3 – Tabela da rotina de treinamento das imagens coletadas

Pode-se observar essa estabilização na figura 6, que é um gráfico onde é possível observar a evolução da precisão em todas as épocas de aprendizado do experimento para cada um dos grupos de treinamento. Através desta observação é possível compreender que o aprendizado

ficará nesta faixa de precisão obtida nas 32.000 épocas, ou seja, os valores de precisão não devem aumentar significativamente para além dos valores obtidos.

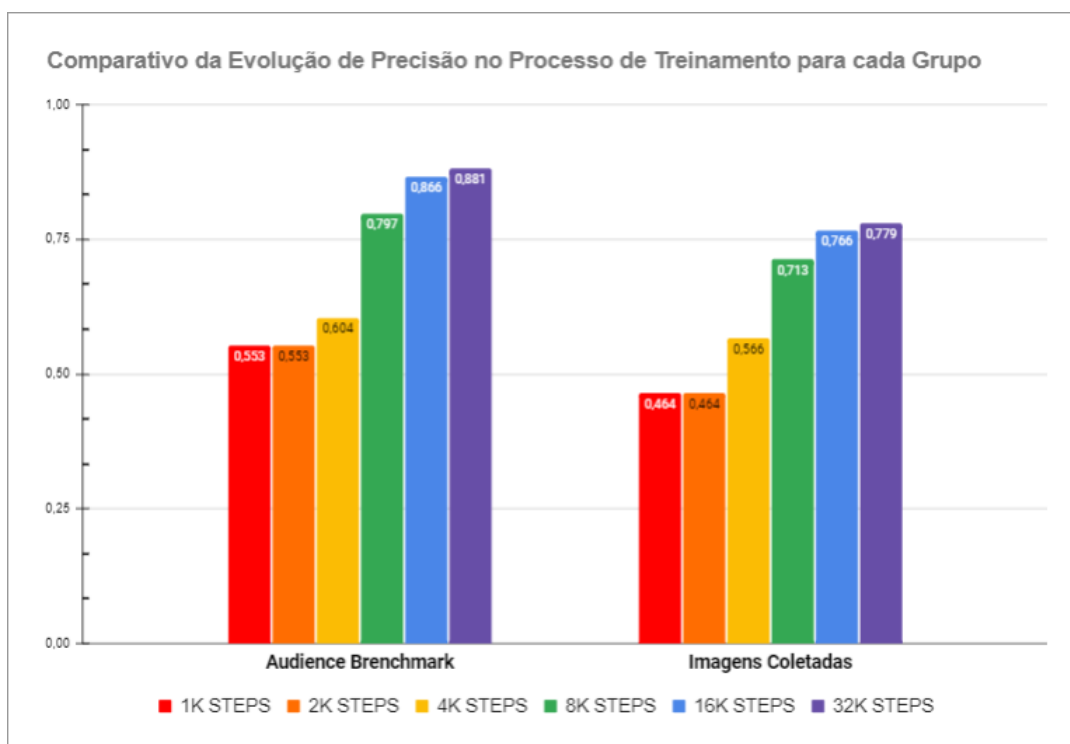


Figura 6 – Gráficos comparativo da evolução da precisão no processo de treinamento de cada grupo.

Fonte: própria

Estes valores de precisão são altos para um modelo, pois a precisão descreve um modelo através da eficiência para classificações realmente positivas e, nesta análise, o *Audience Benchmark* obteve 88,1% de precisão. Comparado com o grupo de imagens coletadas, houve pouca diferença de precisão sendo 77,9% para o grupo de imagens coletadas.

Com relação ao *recall*, que é a métrica que demonstra a eficiência do modelo para as proporções positivas, que em todos os casos foram de 100%, demonstrou grande eficiência para as proporções positivas.

O *f-score* que descreve a eficiência do modelo para as classificações que são relevantes e nesta análise obteve crescimento gradual, semelhante à precisão. Como o *f-score* é uma relação entre o *recall* e a precisão, e pelo fato do *recall* ter sido de 100% em todos os casos, mostra que o *f-score* está diretamente ligado à precisão.

Com a análise dos comparativos dos valores das métricas, observa-se que conforme o modelo se torna mais eficiente para classificações realmente positivas, ele também se torna mais eficiente para classificações relevantes. Quando se trata de um *dataset* incluindo pessoas transgênero na sua composição, significa que problemas na classificação envolvendo as pessoas transgênero são amenizados, pois este grupo está presente nas classificações tidas como positivas e relevantes.

6 CONCLUSÃO

O desenvolvimento deste trabalho possibilitou analisar o impacto em um algoritmos de reconhecimento facial para reconhecimento automatizado de gênero, o *Gender Net*, utilizando um *dataset* diverso, com um conjunto de imagens que contenha pessoas transgênero em sua composição buscando compreender se há mudanças negativas ou positivas com essa inclusão e se há a possibilidade de amenizar problemas de *misgender* tecnológico.

Tais problemas acarretados pelo desenvolvimento deste tipo de tecnologia comprovam que seus desenvolvedores não se preocupam em estarem próximos das realidades da diversidade de gênero e sexualidade, ressaltando ainda mais a não neutralidade do desenvolvimento tecnológico.

O modo como uma tecnologia é pensada e construída faz toda a diferença na experiência de uso e aplicação, principalmente para grupos minoritários de pessoas, como as pessoas transgênero, fazendo com que tal experiência possibilite situações de constrangimento, invisibilização de suas identidades e exposição a situações de risco a sua segurança.

Através dos resultados obtidos, mesmo sem alteração alguma do algoritmo analisado, o *Gender Net*, em uma comparação do treinamento do *Audience Benchmark* com os grupos de imagens coletadas, houve uma diferença de precisão menor para os grupos de imagens coletadas, porém uma diferença baixa de 10,2%, mas que por conseguir englobar um escopo maior de pessoas, ou seja, por englobar as pessoas transgênero, acaba por melhorar sua experiência de uso e aplicação.

Esta análise demonstra que, mesmo partindo de uma perspectiva de interpretação do gênero como algo binário (reconhecendo apenas Homem e Mulher), uma ação simples, que é a inclusão de pessoas transgênero no *dataset*, ajuda a amenizar problemas de *misgender* tecnológico, o que expõe ainda mais a necessidade de se discutir sobre questões de diversidade de gênero e sexualidade dentro do campo da tecnologia.

Neste contexto é importante ressaltar que nenhuma tecnologia é neutra, é preciso responsabilidade ética, política e social, fazendo necessário trazer pautas como lgbtqia+fobia, racismo, capacitismo, acessibilidade e desigualdade social no seu desenvolvimento. Dada a importância ao assunto, como trazer discussões sobre e fazer dessa não neutralidade tecnológica algo inclusivo e eficiente para alcançar mais camadas da sociedade também possibilita questionar a necessidade de uma tecnologia. Como o questionamento da necessidade do conhecimento automatizado de gênero, em que sua proposta é determinar do gênero de uma pessoa de forma automatizada, levando em consideração apenas uma interpretação de gênero e de forma fisiológica. Por si só, essas considerações são excludentes, falhas, preconceituosas e limitantes para algo que é de caráter extremamente pessoal, autoexpressivo, político, existencial e diverso. Com tal responsabilidade, alternativas mais eficientes e menos problemáticas poderiam e podem ser desenvolvidas.

REFERÊNCIAS

- BORTOLETTO, GUILHERME ENGELMAN BORTOLETTO. LGBTQIA+: identidade e alteridade na comunidade. CELAAC USP, São Paulo, SP, Brasil, 2019. Disponível em: <http://celacc.eca.usp.br/sites/default/files/media/tcc/guilherme_engelman_bortoletto.pdf>.
- DUARTE, Denio; STÄHL, Niclas. Machine Learning: A Concise Overview. In: [s.l.: s.n.], set. 2019. p. 27–58.
- HASSNER, Tal et al. Effective Face Frontalization in Unconstrained Images. In: IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). [S.l.: s.n.], jun. 2015. Disponível em: <<https://osnathassner.github.io/talhasner/projects/frontalize>>.
- JUSTIÇA E SEGURANÇA PÚBLICA, Ministério da. **PORTARIA Nº 793, DE 24 DE OUTUBRO DE 2019**. 208. ed. [S.l.]: Diário da União, 25 out. 2019. Disponível em: <<http://www.in.gov.br/en/web/dou/-/portaria-n-793-de-24-de-outubro-de-2019-223853575>>.
- KEYES, Os. The Misgendering Machines: Trans/HCI Implications of Automatic Gender Recognition. **Proc. ACM Hum.-Comput. Interact.**, ACM, New York, NY, USA, v. 2, CSCW, 88:1–88:22, nov. 2018. ISSN 2573-0142. Disponível em: <<http://doi.acm.org/10.1145/3274357>>.
- LEVI, Gil; HASSNER, Tal. Age and Gender Classification Using Convolutional Neural Networks. In: IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) workshops. [S.l.: s.n.], jun. 2015. Disponível em: <https://osnathassner.github.io/talhasner/projects/cnn_agegender>.
- MCLEMORE, Kevin A. Experiences with Misgendering: Identity Misclassification of Transgender Spectrum Individuals. **Self and Identity**, Routledge, v. 14, n. 1, p. 51–74, 2015. Disponível em: <<https://doi.org/10.1080/15298868.2014.950691>>.
- MISGENDER. [S.l.]: Cambridge Dictionary, 27 out. 2020. Disponível em: <<https://dictionary.cambridge.org/pt/dicionario/ingles/misgender>>.
- MISKOLCI, Richard. **Teoria Queer: um aprendizado pelas diferenças**. [S.l.]: Autêntica, 2017.
- PEEMEN, M. et al. The neuro vector engine: Flexibility to improve convolutional net efficiency for wearable vision. In: 2016 Design, Automation Test in Europe Conference Exhibition (DATE). [S.l.: s.n.], 2016. p. 1604–1609.
- PONTI, Moacir; COSTA, Gabriel B. Paranhos. **Como funciona o Deep Learning**. Uberlândia, MG, Brasil: Sociedade Brasileira de Computação – SBC, out. 2017.

SAÚDE, MINISTÉRIO DA; EDUCAÇÃO, MINISTÉRIO DA. **Adolescentes e Jovens para a Educação entre Pares: Diversidade Sexual**. 69. ed. Brasília, DF, Brasil, 2 dez. 2020.

Disponível em: <http://www.unfpa.org.br/Arquivos/guia_diversidades.pdf>.

SCHEUERMAN, Morgan Klaus; PAUL, Jacob M.; BRUBAKER, Jed R. How Computers See Gender: An Evaluation of Gender Classification in Commercial Facial Analysis Services.

Proc. ACM Hum.-Comput. Interact., Association for Computing Machinery, New York, NY, USA, v. 3, CSCW, nov. 2019. Disponível em: <<https://doi.org/10.1145/3359246>>.

SZELISKI, Richard. **Computer vision algorithms and applications**. 2011. Disponível em: <<http://dx.doi.org/10.1007/978-1-84882-935-0>>.

TVT; TGEU. **Trans day of remembrance 2020**. Acesso em: 5 dez. 2020. Disponível em:

<https://transrespect.org/wp-content/uploads/2020/11/TvT_TMM_TDoR2020_Tables.pdf>.

UOL. **Câmeras de reconhecimento facial acham criminosos no Carnaval de Salvador**.

Acesso em: 25 jan. 2020. Disponível em: <<https://noticias.uol.com.br/cotidiano/ultimas-noticias/2019/03/05/cameras-de-reconhecimento-facial-acham-criminoso-no-carnaval-de-salvador.html>>.

<<https://noticias.uol.com.br/cotidiano/ultimas-noticias/2019/03/05/cameras-de-reconhecimento-facial-acham-criminoso-no-carnaval-de-salvador.html>>.

ZHANG, K. et al. Joint Face Detection and Alignment Using Multitask Cascaded

Convolutional Networks. **IEEE Signal Processing Letters**, v. 23, n. 10, p. 1499–1503, out. 2016.