

The Brazilian Data at Risk in the Age of AI?

Raoni F. da S. Teixeira^{1*}, Rafael B. Januzi², Fabio A. Faria^{2*}

¹Faculdade de Engenharia, Universidade Federal de Mato Grosso (FaEng-UFMT)
Cuiabá – MT – Brasil
raoni.teixeira@ufmt.br

²Instituto de Ciência e Tecnologia, Universidade Federal de São Paulo (ICT-UNIFESP)
São José dos Campos – SP – Brasil
rafael.januzi@unifesp.br, ffaria@unifesp.br

Abstract. *Advances in image processing and analysis as well as machine learning techniques have contributed to the use of biometric recognition systems in daily people tasks. These tasks range from simple access to mobile devices to tagging friends in photos shared on social networks and complex financial operations on self-service devices for banking transactions. On July 05th 2021, the Brazilian government announced acquisition of a biometric recognition system to be used nationwide. In the opposite direction to China, Europe and some American cities have already started the discussion about the legality of using biometric systems in public places, even banning this practice in their territory. In order to open a deeper discussion about the risks and legality of using these systems, this work exposes the vulnerabilities of biometric recognition systems, focusing its efforts on the face modality. Furthermore, it shows how it is possible to fool a biometric system through a well-known presentation attack approach in the literature called morphing. Finally, a list of ten concerns was created to start the discussion about the security of citizen data and data privacy law in the Age of Artificial Intelligence (AI).*

Resumo. *Com os avanços das técnicas de processamento e análise de imagens, o uso de sistemas de reconhecimento biométrico em tarefas cotidianas das pessoas já é uma realidade. Dentre essas tarefas estão desde um simples acesso aos dispositivos móveis até a marcação de amigos em fotos compartilhadas em redes sociais e as complexas operações financeiras em equipamentos de autoatendimento para transações bancárias. Em 5 de julho de 2021, o governo brasileiro anunciou a compra de um sistema de reconhecimento biométrico para ser utilizado em todo território nacional. Neste sentido, este artigo propõe a abertura de uma discussão mais aprofundada sobre a adoção de tais sistemas para a identificação dos cidadãos brasileiros e quais os problemas que podem emergir se o sistema não for bem projetado, implantado e gerenciado. Além disso, uma lista de dez questões foi criada para iniciar essa conversa sobre segurança dos dados dos brasileiros na Era da Inteligência Artificial (IA) e o respeito à Lei Geral de Proteção dos Dados (LGPD).*

1. Introdução

Os significativos avanços em algoritmos de processamento e análise de imagens digitais têm contribuído para que os sistemas de reconhecimento facial (RF) rapidamente se tor-

*R. F. da S. Teixeira e F. A. Faria contribuíram igualmente para este trabalho.

nassem um componente importante na rotina de grande parcela da população mundial. Essas tecnologias fazem parte da vida das pessoas desde o simples desbloqueio de seus dispositivos móveis para acesso de suas contas à marcação de amigos em postagens compartilhadas nas redes sociais (e.g., Facebook e Instagram) [Venkatesh et al. 2020].

A adoção generalizada de técnicas de reconhecimento facial tem levantado uma série de preocupações relacionadas com uso dessas tecnologias, principalmente em espaços públicos. Uma das preocupações é que essas tecnologias venham a ser utilizadas para restringir liberdades e monitorar e/ou perseguir grupos minoritários e políticos. Organizações como a Comissão Francesa para Proteção de Dados Pessoais (*Commission Nationale de l'Informatique et des Libertés* – CNIL) [CNIL 2019] e o Instituto AINow da Universidade de Nova York [AI_NOW 2020] têm levantado essas preocupações chamando atenção da opinião pública para a importância do tema e a urgência do debate.

Um efeito prático dessas discussões é que várias cidades americanas tais como Minneapolis¹, São Francisco² e Oakland³ já proibiram o uso de sistema de reconhecimento facial em serviços municipais, incluindo forças policiais. De encontro com essas cidades americanas, em abril de 2021, a Comissão Européia propôs um documento regulatório do uso de tecnologias de inteligência artificial e dentre as aplicações consideradas de alto risco está sistema de reconhecimento biométrico e seu uso em espaços acessíveis ao público para fins de aplicação da lei é proibido a princípio⁴.

Já no caminho oposto do mundo está o Brasil que acaba de anunciar a compra de um sistema nacional de reconhecimento biométrico por meio do consórcio Iafis Brasil e uma empresa fornecedora⁵. Acredita-se que a implantação de um sistema dessa importância precisaria de algum procedimento deliberativo com objetivo de ouvir todas as partes interessadas e abrir uma discussão em profundidade sobre os impactos de seu uso⁷. Essa discussão deve ocorrer observando inclusive o arcabouço legal instituído pela Lei Geral de Proteção de Dados - LGPD⁸ que centraliza regras e estabelece que o cidadão tem controle sobre seus próprios dados.

No Brasil, o documento nacional de identificação civil é emitido pelas Secretarias Estaduais de Segurança Pública (SSPs). Como não há um cadastro centralizado entre esses órgãos, é possível conseguir mais de um documento, cada um deles de um estado diferente, inclusive com números de identificação distintos. Em linhas gerais, as fotos incluídas nos documentos brasileiros de identificação podem ser fornecidas de duas maneiras diferentes dependendo do procedimento adotado pelo estado emissor do documento: 1) a imagem é capturada em uma câmera digital de alta qualidade conectada à

¹<https://techcrunch.com/2021/02/12/minneapolis-facial-recognition-ban/> Acessados em 20/07/2021.

²<https://www.nytimes.com/2019/05/14/us/facial-recognition-ban-san-francisco.html>

³<https://www.sfchronicle.com/bayarea/article/Oakland-bans-use-of-facial-recognition-14101253.php>

⁴<https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>

⁵<https://www.telesintese.com.br/policia-federal-adquire-sistema-de-identificacao-biometrica-com-reconhecimento-facial/>

⁶<https://www.gov.br/pf/pt-br/assuntos/noticias/2021/07/policia-federal-implementa-nova-solucao-automatizada-de-identificacao-biometrica>

⁷<https://www.telesintese.com.br/sistema-de-reconhecimento-facial-da-pf-traz-riscos-e-fere-a-constituicao/>

⁸Lei 13.709 de 14 de agosto de 2018

estação de cadastro oficial; e 2) o cidadão fornece uma foto impressa em papel fotográfico para que seja utilizada na confecção do documento ⁹. Esta segunda maneira é altamente suscetível à falhas por meio de alterações maliciosas de fotografias, causando violações da base de dados do órgão em questão [Venkatesh et al. 2020].

Na literatura já foram relatados muitas formas de alterações fotográficas para mais diversas finalidades. Algumas alterações podem ser não intencionais, introduzidas pelos próprios dispositivos de aquisição ou impressão, os quais podem distorcer a imagem ou modificar sua proporção resultando em alterações da geometria da face do indivíduo. Outras alterações pode ser realizadas de forma intencional por meio do uso de ferramentas de processamento de imagem para tornar uma pessoa mais atraente ou para realização de atos criminosos, como por exemplo, enganar um sistema de reconhecimento automático [Ramachandra and Busch 2017].

Estudos recentes revelaram que os passaportes eletrônicos são particularmente sensíveis ao chamado ataque baseado em metamorfose, no qual a foto do rosto impressa em papel e fornecida pelo cidadão pode ser alterada antes de ser cadastrada no banco de dados [Venkatesh et al. 2020, Seibold et al. 2020]. Esse ataque foi descrito pela primeira vez no artigo [Ferrara et al. 2014] em contexto de verificação de faces em portões de controle automatizado de fronteira (CAF), onde dois sujeitos (ator malicioso e cúmplice) cooperam para produzir uma imagem de rosto adulterada por metamorfose, misturando suas identidades a fim de obter um documento de viagem oficial (passaporte) que pode ser explorado normalmente por ambos os sujeitos. A Figura 1 mostra um esquemático de como a alteração de fotografias pode ser utilizada para burlar um sistema de controle automatizado de fronteira em aeroportos.

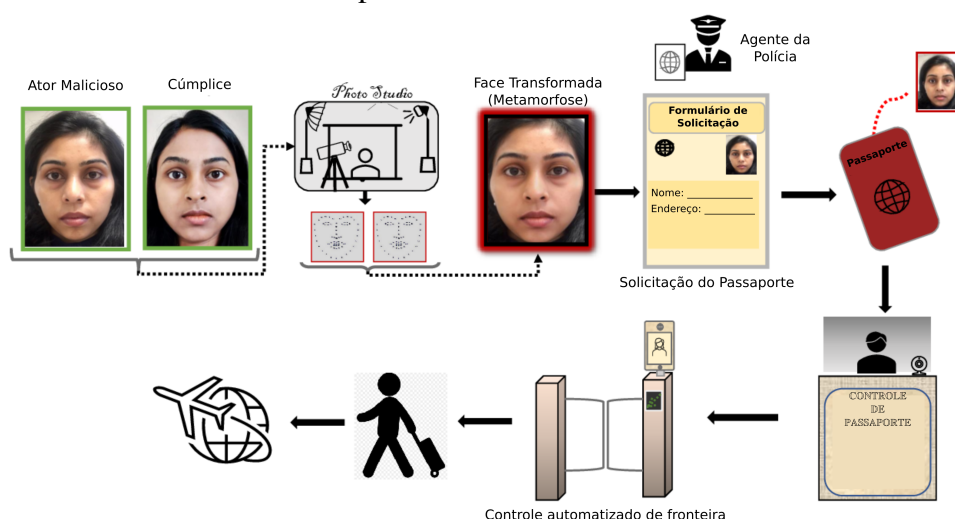


Figure 1. Exemplo de uso de transformação de imagens (metamorfose) para burlar o controle automatizado de fronteiras em aeroportos. Imagem adaptada de [Venkatesh et al. 2020].

Os ataques baseados em metamorfose em sistemas CAF não é apenas um desafio de pesquisa exclusivamente acadêmico. Apenas no ano de 2018, mais de 150 mil pessoas

⁹A Portaria no 2, de 15 de abril de 2019 Instituto de Identificação de Minas Gerais, por exemplo, define o procedimento estadual para envio de fotos.

podem ter entrado ilegalmente na União Europeia segundo estimativas da Comissão Europeia¹⁰. A viabilidade deste ataque foi demonstrada pelo grupo ativista Peng! Kollektiv, que em 2018 solicitou com sucesso um passaporte com foto modificada que combinava as faces de um dos membros do grupo e a política italiana Federica Mogherini (Ex-Alto Representante da União para os Negócios Estrangeiros e a Política de Segurança)¹¹. Um outro caso reportado no *Fifteenth Symposium and Exhibition on the ICAO Traveller Identification Programme*¹² mostra que um cidadão afegão foi detido na fronteira alemã portando um passaporte obtido com uma imagem adulterada. A preocupação vai muito além de documentos de CAF, pois os ataques podem ser estendidos a outros cenários de violação de bases de dados tais como carteira de trabalho, habilitação e até mesmo documento de identificação oficial do país.

As vulnerabilidades de sistemas de reconhecimento facial impulsionaram diversas atividades de pesquisa e desenvolvimento pelo mundo. Vários projetos de pesquisa estão sendo financiados pela União Europeia e Conselhos Nacionais de Pesquisa (e.g., SWAN¹³, ANANAS¹⁴, SOTAMD¹⁵ e iMARS¹⁶) para a proposição de técnicas de detecção de ataques baseados em metamorfose. Uma conferência dedicada foi iniciada pela Frontex (Agência Europeia da Guarda Costeira e de Fronteiras¹⁷). Além disso, o Instituto Nacional de Padrões e Tecnologia dos EUA (*National Institute of Standards and Technology - NIST*) criou bases de dados e uma plataforma para avaliação desse tipo de ataque [Ngan et al. 2020].

Em resposta ao recente anúncio da contratação da Solução Automatizada de Identificação Biométrica (ABIS), o objetivo deste trabalho é abrir uma discussão sobre a confiabilidade e/ou vulnerabilidade de sistemas de reconhecimento biométricos baseados em faces e como as técnicas mais avançadas da área de inteligência artificial podem facilmente corromper esse tipo de sistema se mal projetado, implantado e gerenciado.

Este artigo está organizado da seguinte maneira. A seção 2 descreve a metodologia adotada e alguns conceitos importantes para um melhor entendimento do trabalho. A seção 3 lista algumas questões que precisam ser discutidas antes que um sistema de reconhecimento biométrico possa ser implantado no Brasil. Finalmente, na Seção 4, as conclusões são realizadas.

2. Fundamentação Teórica e Conceitos Importantes

Esta seção apresenta os principais conceitos que foram levados em consideração para suportar a argumentação e contribuição deste trabalho.

2.1. Sistema de Reconhecimento Facial

Reconhecimento facial (RF) é uma expressão genérica usada para descrever um número variado de sistemas que procuram estabelecer algum tipo de correspondência entre duas

¹⁰https://knowledge4policy.ec.europa.eu/dataset/ds00032_en

¹¹<https://pen.gg/campaign/mask-id-2/>

¹²http://biolab.csr.unibo.it/WorkshopPages/Download/4.ferrara_franco.pdf

¹³<https://www.ntnu.edu/iik/swan>

¹⁴<https://www.bmbf.de/en/index.html>

¹⁵<https://www.ntnu.edu/iik/sotamd>

¹⁶<https://cordis.europa.eu/project/id/883356>

¹⁷<https://frontex.europa.eu/future-of-border-control/research-and-innovation/announcements/>

imagens de rostos humanos. Mais precisamente, um sistema de RF é composto por algoritmos que analisam as fotos dos rostos e extraem um conjunto de características distinguíveis. As características correspondem aos atributos físicos da face como, por exemplo, a distâncias entre os olhos e o formato do nariz e, em geral, são codificadas em um objeto denominado *template*. Os *templates* são armazenados em uma base de dados para comparações futuras. A correspondência entre faces é calculada por um algoritmo que verifica a semelhança entre *templates*. O resultado da comparação é um número real chamado de *score* (pontuação). Quanto maior esse número, mais semelhantes são as pessoas. A Figura 2 mostra um diagrama de blocos de uma arquitetura típica de um sistema de RF.

Os blocos representam os quatro principais componentes do sistema:

- **Dispositivo de captura:** Estação de cadastro com câmera e sistema de aquisição;
- **Extração de Características:** Algoritmo de processamento de imagens utilizados para extrair características e criar um *template*;
- **Comparação:** Algoritmo utilizado para calcular a correspondência entre o *template* de referência da base de dados e o *template* da amostra em análise; e
- **Base de dados:** Coleção de dados contendo as informações pré-registradas no sistema (*templates* ou imagens usados como referência).

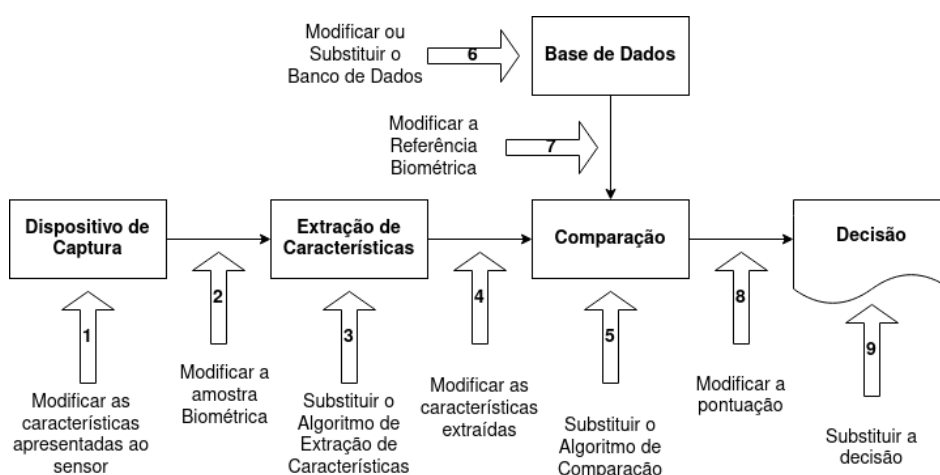


Figure 2. Sistema de reconhecimento facial com pontos de vulnerabilidade. Imagem adaptada de [Scherhag 2020].

As setas apresentam os pontos de vulnerabilidade de um sistema composto por essa arquitetura:

1. **Vulnerabilidade de apresentação:** As características físicas apresentadas ao sensor são modificadas propositalmente utilizando máscaras, maquiagem, modificação física [Yoon et al. 2012] e uso de mídias exibindo fotos e vídeos de pessoas;
2. **Modificação da amostra recém capturada:** A foto coletada é substituída ou modificada deliberadamente. Esse é o exemplo mais simples de ataque de metamorfose em que uma foto adulterada é incluída no sistema;
3. **Substituição do algoritmo de extração de características:** O algoritmo que cria os *templates* é alterado para funcionar de maneira indevida;

4. **Modificação das características extraídas:** As características extraídas pelo algoritmo são manipuladas antes de serem enviadas ao algoritmo de comparação. Técnicas de metamorfose podem ser utilizadas nesse ponto para alterar localmente as características extraídas;
5. **Substituição do algoritmo de comparação de templates:** O algoritmo de comparação é modificado para funcionar de maneira indevida;
6. **Modificação na base de dados:** Os *templates* ou as imagens armazenadas na base de dados são alterados. Inclui vulnerabilidades relacionadas com acesso indevido, roubos de dados e reconstrução dos *templates* de referência [Mai et al. 2019]. Técnicas de metamorfose podem ser utilizadas nesse ponto para alterar a representação do indivíduo na base de dados;
7. **Modificação das características de referência:** As características de referência são manipuladas antes de serem enviadas ao algoritmo de comparação;
8. **Alteração da pontuação:** A pontuação devolvida pelo algoritmo de comparação é modificada a fim de influenciar o resultado da decisão.
9. **Alteração da decisão final do algoritmo:** A decisão final do algoritmo (identificação, autenticação ou verificação) é alterada para um propósito específico.

Os pontos de vulnerabilidades citados anteriormente são definidos na norma ISO/IEC 30107-1, elaborada pela *International Organization for Standardization (ISO)* e pela *International Electrotechnical Commission (IEC)* [Biometrics 2016]. Ataques explorando vulnerabilidades específicas foram reportadas na literatura [Ramachandra and Busch 2017, Anjos et al. 2014, Mai et al. 2019]. Considerando essas vulnerabilidades, um ataque de metamorfose pode ser aplicado explorando os pontos (2), (4), (6) e (7).

2.2. As Avançadas Redes Generativas Adversariais (GANs)

Dentre as técnicas mais avançadas da área de inteligência artificial atualmente está uma família de algoritmos chamados modelos generativos, que tem como objetivo criar/gerar novos dados dada uma distribuição de interesse. Diferente dos modelos generativos tradicionais, aqueles baseados em máxima verossimilhança (em inglês, *maximum likelihood estimation* – MLE), as GANs são capazes de representar melhor dados complexos e aprender distribuições de dados de alta dimensionalidade. Propostas por [Goodfellow et al. 2014], GANs são definidas como modelos generativos otimizados baseados em estratégia de treinamento adversarial.

Nesta abordagem, as GANs são compostas de duas redes neurais: um gerador (G) e um discriminador (D). Essas redes são treinadas usando uma estratégia de otimização chamada MinMax, onde a função objetivo é compartilhada por ambas redes. Por um lado, o gerador tem como objetivo enganar/confundir o discriminador em acreditar que as amostras criadas e as amostras reais pertencem a mesma distribuição. Por outro lado, o discriminador tem o papel de diferenciar as amostras reais das amostras geradas (do inglês, *fake*) [Zhang et al. 2020, Faria and Carneiro 2020].

Na literatura, diversas arquiteturas GANs têm sido propostas para as mais variadas aplicações tais como compressão de dados para vídeo conferências [Wang et al. 2021], reconstrução de imagens de diferentes domínios [Richardson et al. 2021], criação de paisagens [Park et al. 2019] e geração de imagens de faces humanas extremamente realísticas

e de alta resolução [Karras et al. 2019, Karras et al. 2020]. A Figura 3 mostra exemplos de imagens *fake* criadas pela arquitetura StyleGAN2. É importante notar a alta qualidade das imagens *fake* criadas pela arquitetura que podem ser facilmente confundidas com imagens reais.

Dentre as aplicações bem intencionadas desenvolvidas na literatura também estão aquelas com propósitos escusos, como a utilização da StyGAN para criação de novas faces humanas baseada em transformação de metamorfose. Essa transformação têm como objetivo transformar uma imagem de um indivíduo em outro por meio de uma operação de interpolação do espaço latente. A Figura 4 mostra exemplos de imagens resultantes de operação de interpolação de dois indivíduos alvo (a) e (e). É possível observar que as imagens (b), (c) e (d) apresentam uma mistura de diferentes características de ambos indivíduos.



Figure 3. Alguns exemplos das imagens criadas pela StyleGAN2. Imagem adaptada de [Karras et al. 2020].

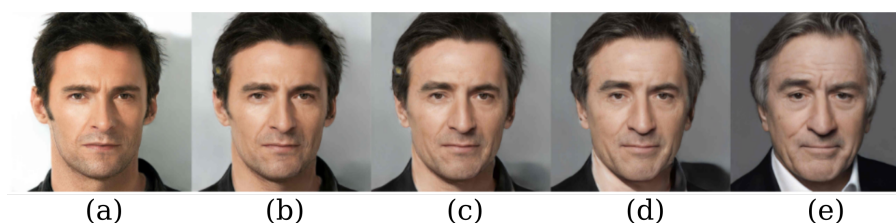


Figure 4. Exemplos de transformações de imagens por metamorfose entre dois indivíduos. Imagem adaptada de [Venkatesh et al. 2020].

3. As Dez Preocupações!

Nesta seção são listadas dez importantes questões que precisam ser levadas em consideração em uma discussão sobre a adoção de um sistema de reconhecimento biométrico com ampla abrangência no Brasil. As ideias por trás dessas questões nasceram da análise dos pontos críticos de sistemas reais (Figura 2) e de casos recentes mostrando mau uso da identificação de pessoas. Seleccionamos 3 (três) desses casos para ilustrar que as questões apresentadas tocam diferentes aspectos do projeto e da implantação do sistema, a seguir:

- **Caso I:** Em 2018, o grupo ativista Peng! Kollektiv obteve um passaporte alemão usando uma foto modificada que combinava as faces de um dos membros do grupo e da política italiana Federica Mogherini¹⁸.

¹⁸<https://pen.gg/campaign/mask-id-2/>

- **Caso II:** Ainda em 2018, uma equipe de pesquisadores do MIT e Stanford encontrou viés de gênero e raça em 3 (três) sistemas comerciais de reconhecimento facial - incluindo sistema utilizado por agências oficiais¹⁹.
- **Caso III:** Em dezembro de 2021, o senhor José Domingos Leitão foi preso indevidamente pela Polícia Civil do Distrito Federal, após ser confundido por um criminoso²⁰. A identificação de Domingos Leitão foi realizada comparando imagens de uma câmera de segurança com as registros oficiais do sistema de identificação.

3.1. O uso de métodos chamados “caixa preta”

As soluções comerciais de RF disponíveis são métodos “caixa preta”, em que a implementação algorítmica não é conhecida (i.e. proprietária) e/ou não pode ser interpretada por seres humanos como é o caso dos algoritmos de aprendizado profundo. Os problemas relacionados ao desconhecimento da implementação do método de reconhecimento biométrico causam uma insegurança em todo o processo de identificação. Os **Casos II e III** mostram que sistemas caixa-preta estão sujeitos ao enviesamento (i.e. preconceito) e, até mesmo, podem levar inocentes para prisão. As respostas para as seguintes perguntas são primordiais: (1) Será que alguém pode ser incriminado por um sistema que tem sua implementação desconhecida? (2) Como proteger as pessoas de soluções pseudo-científicas?

3.2. O Reconhecimento Facial em uma nação extremamente heterogênea

Um país, como o Brasil, que possui uma população altamente heterogênea e, consequentemente, apresenta uma mistura elevada de características físicas (e.g., coloração da pele, olhos e cabelos e formato do rosto), não pode ignorar o viés dos algoritmos biométricos. O **Caso II** nos mostra que o enviesamento é real e que seria de bom tom responder às seguintes perguntas: (1) “Como podemos garantir que o sistema conseguirá, não apenas identificar os indivíduos de maneira confiável, mas também diferenciar quaisquer dois indivíduos?”; (2) “As taxas de acerto dos algoritmos são equivalentes para os diferentes grupos demográficos?”; e (3) “Caso o software seja importado, as taxas de erro na população original podem ser reproduzidas na população brasileira? (i.e. efeito *cross-race* [Phillips et al. 2009])”

3.3. A falta de transparência e/ou padronização no processo de aquisição de imagens

Para países de grandes dimensões, como é o caso do Brasil, onde será necessário implantar uma capacidade de aquisição de dados muito ampla, distribuída adequadamente por todo o território e fazendo uso de uma grande quantidade de profissionais e equipamentos, algumas perguntas precisam ser respondidas: (1) “Como é possível garantir que os processos estejam alinhados de maneira satisfatória em todos os pontos de coleta?”; (2) “Qual o impacto de ruídos operacionais na execução dos processos estabelecidos?”; e (3) “Qual o impacto de divergências técnicas de dispositivos de coleta entre diferentes fabricantes?” Uma vez que a qualidade de sistema de reconhecimento biométrico é altamente relacionada com a qualidade das bases de dados, esses questionamentos sobre

¹⁹<https://news.mit.edu/2018/study-finds-gender-skin-type-bias-artificial-intelligence-systems-0212>

²⁰<https://noticias.r7.com/brasil/programa-da-policia-civil-identifica-homem-errado-e-inocente-e-presos-17122021>

as barreiras a serem enfrentadas na aquisição de imagens se tornam bastante relevantes, podendo colocar o sucesso da implantação do sistema em risco. O **Caso III** é um típico exemplo dos erros que podem ocorrer quando se compara imagens adquiridas em tempos e usando equipamentos distintos.

3.4. A distinguibilidade das características utilizadas na comparação

Mesmo após todo avanço dos últimos anos, ainda faltam estudos estatísticos em escala nacional que demonstrem como as características dos *templates* utilizados pelos algoritmos de RF individualizam as pessoas da base de dados. Em particular, as perguntas fundamentais continuam sem resposta: (1) “Qual é a probabilidade de encontrar duplicações no espaço de características projetado pelo algoritmo para representar o rosto de uma pessoa?”; e (2) “A pessoa identificada é um *doppelgänger* (gêmeo sem relação biológica)”? A resposta para essa pergunta é de extrema importância para evitar acontecimentos iguais ao ocorrido na investigação do *Federal Bureau of Investigation - FBI* contra Steve Talley²¹ e na investigação da Polícia Civil do DF do **Caso III**.

3.5. A integração entre diferentes base de dados já existentes

Um risco da integração das bases oficiais de identificação brasileiras é propagação de erros e violações de bases de dados locais para outras bases em diferentes estados da nação. Iniciativas de integração já começam a ser implementadas a medida que os estados brasileiros iniciam seu processo de transformação digital. Este é o caso, por exemplo, da operação anunciada recentemente para interligar as bases de dados (incluindo as base de identificação de pessoas) nos governos de Mato Grosso e Amapá²². A principal pergunta a ser respondida nesse caso é: “Todas as bases de dados envolvidas são íntegras?” A resposta para essa pergunta passa pela prevenção de ataques e pela garantia de que cada indivíduo foi registrado somente uma vez (i.e. não existe indivíduo se passando por outra pessoa). Até onde sabemos, as bases de dados brasileiras não foram avaliadas sob essa perspectiva.

3.6. A fiscalização por entidades especializadas e independentes

Um fator essencial na discussão é a falta de mecanismos que permitam que entidades independentes examinem os sistemas biométricos brasileiros levando em conta o quadro regulamentar existente e apontem caminhos para melhorar a proteção dos dados pessoais. Mais especificamente, essas entidades poderiam, por exemplo, testar e avaliar o sistema utilizado pela Polícia Civil do DF (**Caso III**) e direcionar esforços para seu aperfeiçoamento.

3.7. A implantação de algoritmos de detecção de ataque em sistemas oficiais

Até onde sabemos, as Secretarias Estaduais de Segurança Pública (SSPs) brasileiras não utilizam algoritmos de detecção automática de ataques de metamorfose para avaliação das imagens de identificação. Desde 2018, o relatório da *Commission to the European Parliament* recomendam a implementação de tal prática para garantir uma maior

²¹<https://theintercept.com/2016/10/13/how-a-facial-recognition-mismatch-can-ruin-your-life/>

²²<https://x-road.global/scaling-interoperability-across-states-for-national-digital-transformation-in-brazil>

robustez no processo de aquisição e armazenamento das imagens em bancos de dados [EUCouncil 2018]. Um dos desafios para a construção desses tipos de algoritmos é a falta de bases públicas para serem utilizadas nos projetos de pesquisas e principalmente ao longo do seu desenvolvimento. Note que de acordo com o artigo 11º da LGPD, a garantia de prevenção à fraudes nos processos de identificação é uma das hipóteses que justificam o tratamento de dados biométricos. O **Caso I** é uma alerta que não pode ser ignorado.

3.8. Definição de protocolos e diretrizes para interpretação dos resultados dos algoritmos

Uma questão crucial a ser enfrentada quando se pretender aplicar algoritmos em problemas de segurança pública é: “Como interpretar o resultado produzido?”. Para evitar análises subjetivas é necessário estabelecer diretrizes claras que consideram o contexto da aplicação (e.g. qualidade e resolução dos dados utilizados e as limitações tecnológicas). O **Caso III** é elucidativo. Até onde sabemos, o laudo elaborado pela Instituto de identificação da Polícia Civil do DF não levou em conta: (1) diferenças entre as condições de captura (e.g. resolução da câmera, controle de iluminação) e (2) mudanças decorrentes da passagem do tempo desde a elaboração do documento de identificação. Vale destacar que esse tipo de problema é independente da falta de conhecimento sobre o sistema (Preocupação 3.1), pois mesmo utilizando um sistema explicável será necessário considerar o contexto e fatores como resolução e passagem do tempo não podem ser desprezados [Best-Rowden and Jain 2018].

3.9. O treinamento, atualização e capacitação de profissionais da área de segurança pública

É importante que todos os agentes envolvidos com o processo de identificação de pessoas recebam treinamento e entendam os pontos de vulnerabilidade dos sistemas. Os ataques evoluem em conjunto com a tecnologia (ao contrário do que se acreditava há alguns anos, sabemos hoje que é possível reconstruir *templates* biométricos [Mai et al. 2019]). A era da IA é um momento histórico em que é possível automatizar tarefas que antes exigiam mais tempo e recursos. Portanto, os agentes envolvidos com o sistema de segurança precisam estar preparados para os novos desafios que aparecerão durante o processo. Os **Casos I, II e III** são alguns dos exemplos das situações para as quais os agentes devem estar preparados.

3.10. A Lei Geral de Proteção dos Dados (LGPD)

A Lei 13.709 de 14 de Agosto de 2018, porém que entrou em vigor em 18 de Setembro de 2020, a LGPD estabelece diretrizes importantes e obrigatórias para a coleta, processamento e armazenamento de dados pessoais do povo brasileiro. Ela foi inspirada na GDPR (*General Data Protection Regulation*), que entrou em vigência em 2018 na União Europeia, trazendo grandes impactos para empresas e consumidores²³. Dentre os principais objetivos da proposição desta lei está “Assegurar o direito à privacidade e à proteção de dados pessoais dos usuários, por meio de práticas transparentes e seguras, garantindo direitos fundamentais.”. Portanto, a utilização generalizada de sistemas biométricos vulneráveis em âmbito nacional para controlar a vida dos brasileiros poderá

²³https://www.sebrae.com.br/sites/PortalSebrae/canais_adicionais/conheca_lgpd

trazer consequências desastrosas para toda sociedade chegando até mesmo a desrespeitar a LGPD, pois não existe garantias que os dados dos brasileiros estarão protegidos de quaisquer ataques que tais sistemas possam sofrer no futuro.

4. Conclusões e Trabalhos Futuros

Este trabalho teve como objetivo iniciar uma discussão sobre o anúncio de contratação de um sistema de reconhecimento biométrico chamado “Solução Automatizada de Identificação Biométrica” (ABIS) por meio do consórcio Iafis Brasil e uma empresa fornecedora da solução. Nós pudemos mostrar que essa discussão prévia é de extrema importância para uma nação e apontamos uma série de 10 (dez) preocupações que podem ajudar na definição de padrões e na criação de políticas públicas direcionadas. As recomendações falam por si mesmo e possuem uma correspondência e emparelhamento com as preocupações e incluem a padronização na aquisição (3.3), definição de protocolos (3.8) e implantação de algoritmos de detecção de ataques (3.7). Esses pontos passam por iniciativas semelhantes ao programa *Face Recognition Vendor Test (FRVT)*²⁴ implementado nos Estados Unidos, em que uma série de avaliações independentes em larga escala são realizadas em sistemas de reconhecimento facial. Para tanto, é necessário base de dados nacionais, estudos e análises distribuídos pelo país. Nesse sentido, entendemos que, além das políticas públicas, serão necessários parcerias envolvendo o setor público, organizações da sociedade civil e universidades. Esperamos que este trabalho possa servir como apoio para as pessoas responsáveis pelos órgãos competentes e que elas possam tomar as melhores decisões no futuro.

Agradecimentos

Os autores agradecem a Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pelos processos #2018/23908-1 e #2017/25908-6, a Fundação de Amparo à Pesquisa do Estado de Mato Grosso pelo processo 0204690/2017, a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela infraestrutura do laboratório GIBIS e apoio financeiro.

Referências

- AI NOW (2020). Regulating biometrics: Global approaches and open questions. Technical report, AI Now Institute, New York, USA.
- Anjos, A., Chakka, M. M., and Marcel, S. (2014). Motion-based counter-measures to photo attacks in face recognition. *IET Biom.*, 3:147–158.
- Best-Rowden, L. and Jain, A. K. (2018). Longitudinal study of automatic face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(1):148–162.
- Biometrics, I. -. (2016). Information technology biometric presentation attack detection - part 1: Framework. Technical report, International Organization for Standardization, Geneva, Switzerland.
- CNIL (2019). Facial recognition: for a debate living up to the challenges. Technical report, CNIL - Commission Nationale de l’Informatique et des Libertés, Paris, France.

²⁴<https://www.nist.gov/programs-projects/face-recognition-vendor-test-frvt>

- EU Council (2018). Report from the commission to the european parliament and the council on the implementation of the action plan to strengthen the eu response to travel document fraud. Technical report, Council of the European Union, Brussels, Belgium.
- Faria, F. A. and Carneiro, G. (2020). Why are generative adversarial networks so fascinating and annoying? In *2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pages 1–8.
- Ferrara, M., Franco, A., and Maltoni, D. (2014). The magic passport. In *IEEE International Joint Conference on Biometrics, Clearwater, IJCB 2014, FL, USA, September 29 - October 2, 2014*, pages 1–7. IEEE.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014). Generative adversarial nets. In Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N. D., and Weinberger, K. Q., editors, *Advances in Neural Information Processing Systems 27*, pages 2672–2680.
- Karras, T., Laine, S., and Aila, T. (2019). A style-based generator architecture for generative adversarial networks. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4401–4410.
- Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., and Aila, T. (2020). Analyzing and improving the image quality of stylegan. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8107–8116.
- Mai, G., Cao, K., Yuen, P. C., and Jain, A. K. (2019). On the reconstruction of face images from deep face templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(5):1188–1202.
- Ngan, M., Grother, P., Hanaoka, K., and Kuo, J. (2020). Face recognition vendor test (frvt) part 4: Morph - performance of automated face morph detection.
- Park, T., Liu, M.-Y., Wang, T.-C., and Zhu, J.-Y. (2019). Semantic image synthesis with spatially-adaptive normalization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Phillips, P., O’Toole, A., Jiang, F., Narvekar, A., and Ayadd, J. (2009). An other-race effect for face recognition algorithms.
- Ramachandra, R. and Busch, C. (2017). Presentation attack detection methods for face recognition systems: A comprehensive survey. *ACM Comput. Surv.*, 50(1).
- Richardson, E., Alaluf, Y., Patashnik, O., Nitzan, Y., Azar, Y., Shapiro, S., and Cohen-Or, D. (2021). Encoding in style: A stylegan encoder for image-to-image translation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2287–2296.
- Scherhag, U. J. (2020). *Face Morphing and Morphing Attack detection*. PhD thesis, Technische Universität Darmstadt.
- Seibold, C., Samek, W., Hilsmann, A., and Eisert, P. (2020). Accurate and robust neural networks for face morphing attack detection. *Journal of Information Security and Applications*, 53:102526.
- Venkatesh, S., Ramachandra, R., Raja, K. B., and Busch, C. (2020). Face morphing attack generation & detection: A comprehensive survey. *CoRR*, abs/2011.02045.
- Wang, T.-C., Mallya, A., and Liu, M.-Y. (2021). One-shot free-view neural talking-head synthesis for video conferencing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 10039–10049.
- Yoon, S., Feng, J., and Jain, A. K. (2012). Altered fingerprints: Analysis and detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(3):451–464.
- Zhang, H., Zhang, Z., Odena, A., and Lee, H. (2020). Consistency regularization for generative adversarial networks. In *International Conference on Learning Representations – to appear*.