Universidade de Brasília – UnB
Faculdade UnB Gama – FGA
Engenharia Eletrônica

Proposta e Avaliação de Métodos Ativos para Detecção de Falsificação Facial Eletrônica

Autor: Lucas Pereira Pires
Orientador: Dr. Diogo Caetano Garcia

Brasília, DF
2023

Lucas Pereira Pires

Proposta e Avaliação de Métodos Ativos para Detecção de Falsificação Facial Eletrônica

Monografia submetida ao curso de graduação em Engenharia Eletrônica da Universidade de Brasília, como requisito parcial para obtenção do Título de Bacharel em Engenharia Eletrônica.

Universidade de Brasília – UnB

Faculdade UnB Gama – FGA

Orientador: Dr. Diogo Caetano Garcia

Brasília, DF

2023

Lucas Pereira Pires
Proposta e Avaliação de Métodos Ativos para
Detecção de Falsificação Facial Eletrônica/ Lucas Pereira Pires. – Brasília, DF, 2023-
38 p. : il. (algumas color.) ; 30 cm.

Orientador: Dr. Diogo Caetano Garcia

Trabalho de Conclusão de Curso – Universidade de Brasília – UnB
Faculdade UnB Gama – FGA, 2023.

1. Reconhecimento facial. 2. Falsificação facial eletrônica. I. Dr. Diogo Caetano Garcia. II. Universidade de Brasília. III. Faculdade UnB Gama. IV. Proposta e Avaliação de Métodos Ativos para
Detecção de Falsificação Facial Eletrônica.

CDU

Lucas Pereira Pires

Proposta e Avaliação de Métodos Ativos para Detecção de Falsificação Facial Eletrônica

Monografia submetida ao curso de graduação em Engenharia Eletrônica da Universidade de Brasília, como requisito parcial para obtenção do Título de Bacharel em Engenharia Eletrônica.

Trabalho aprovado. Brasília, DF, 20 de fevereiro de 2023 – Data da aprovação do Trabalho de Conclusão de Curso:

Dr. Diogo Caetano Garcia
Orientador

Dr. Cristiano Jacques Miosso
Convidado 1

Dr. Daniel M. Muñoz Arboleda
Convidado 2

Brasília, DF
2023

Resumo

Em nossa sociedade, a aparência de uma pessoa é a forma mais direta de tentar reconhecê-la. Baseado neste comportamento humano, o reconhecimento facial é uma importante área de estudo que busca a identificação de uma pessoa por suas características faciais. Entretanto, ameaças a sistemas do tipo podem ser facilmente reproduzidas, utilizando imagens individuais obtidas em redes sociais, gerando resultados falso positivos para impostores. Assim, associada à pesquisa de técnicas de reconhecimento facial, está a busca de métodos de prevenção a essas ameaças.

A falsificação facial indireta é uma conhecida ameaça que levanta maior interesse na prevenção de ataques a sistemas de reconhecimento. Trata-se da reprodução da imagem de uma pessoa, por foto impressa ou reprodução digital, que é amplamente conhecida pela sua simplicidade de execução. Em grande parte, pesquisas relacionadas à prevenção desse tipo de ameaça são baseadas no reconhecimento de padrões por processamento de imagens 2D. Esta pesquisa propõe e avalia técnicas ativas de detecção de falsificação facial por apresentação de imagem em aparelhos eletrônicos com tela digital, que considere informações além da imagem 2D capturada, utilizando da reflexão da tela.

Primeiramente é feita uma revisão teórica das ferramentas de processamento de imagens utilizadas e dos termos detecção, reconhecimento e falsificação faciais, assim como das técnicas atuais disponíveis em métodos ativos de detecção de falsificação eletrônica. Após isso, são apresentados os métodos propostos, baseados inicialmente na simplicidade de aplicação. O primeiro método consiste em um dispositivo com emissores e receptores na faixa do infravermelho para captação automática da reflexão. O segundo busca a detecção da reflexão de uma fonte luminosa, posicionada próxima à câmera, por ferramentas de processamento de imagens.

O primeiro método foi analisado a partir da leitura de tensão no fototransistor receptor, por meio de um conversor analógico digital de 10 bits. Foram realizados testes iniciais em dois aparelhos eletrônicos, em distâncias na faixa de 4 a 15 centímetros. Os dispositivos captaram uma maior intensidade luminosa, de reflexão, em distâncias menores. Na distância máxima averiguada, de 15 centímetros, a leitura para um dos aparelhos foi de 941 e 878, para dois dispositivos desenvolvidos, em comparação com o valor de 953, lido na ausência dos aparelhos de apresentação.

O segundo método foi considerado mais viável para aplicação e consequentemente mais experimentado. Os experimentos foram realizados alternando condições de iluminação, cor de fundo e distância, com o ajuste de duas variáveis, de limiarização da imagem e de decisão, obtendo um valor de acurácia máximo foi de 89,6% e de F1-score de 88,9%. O sistema foi capaz de diferenciar a foto real da reprodução de imagem na maior parte dos casos e a reflexão luminosa impossibilitou a detecção de face, em alguns casos, na tentativa de fraude em distâncias menores.

Palavras-chave: Reconhecimento facial. Falsificação facial eletrônica. Processamento de imagens. Reflexão de fontes luminosas.

Abstract

In our society, a person's appearance is the most direct way of trying to recognize them. Based on this human behavior, facial recognition is an important field of study that seeks to identify a person by facial features. However, threats to such systems can be easily reproduced using individual images obtained on social media, generating false positive results for impostors. Thus, associated with the search for facial recognition techniques, is the search for methods to prevent these threats.

Indirect face-spoofing is a known threat that raises increased interest in prevention of attacks on recognition systems. This is the reproduction of the image of a person, by printed photo or digital presentation, which is widely known for its simplicity of execution. For the most part, research related to prevent this type of threat are based on pattern recognition by 2D image processing. This research proposes active techniques for face spoofing detection by image presentation on electronic devices with digital screen, that consider information in addition to the captured 2D image, using the reflection of the screen.

Firstly, a theoretical review is made of the image processing tools used and of the terms face detection, recognition and spoofing, as well as the current techniques available in active methods for electronic face-spoofing detection. Next, the proposed methods are presented, initially based on the simplicity of application. The first method consists of a device with emitters and receivers in the infrared range for automatic capture of reflection. The second searches the detection of the reflection of a light source, positioned close to the camera, by image processing tools.

The first method was analyzed from the voltage reading on the receiver phototransistor, via a 10-bit analog-to-digital converter. Initial tests were carried out on two electronic screens, at distances in the range of 4 to 15 centimeters. The devices captured a greater luminous intensity, of reflection, in smaller distances. At the maximum distance investigated, of 15 centimeters, the reading for one of the screens was 941 and 878, for two developed devices, in comparison with the value of 953, read in absence of presentation devices.

The second method was considered more viable for application and consequently more experienced. The experiments were carried out alternating lighting conditions, background color and distance, with the adjustment of two variables, image thresholding and decision making, obtaining a maximum accuracy value of 89.6% and an F1-score of 88.9%. The system was able to differentiate the real photo from the image reproduction in most part of the cases and the light reflection made it impossible to detect the face, in some cases, in the fraud attempt at smaller distances.

Key-words: Facial recognition. Electronic Face-spoofing. Image processing. Reflection of luminous sources.

Lista de ilustrações

Sumário

1 Introdução

Aplicações do reconhecimento facial podem ser vistas em diferentes áreas, como em investigações forenses, vigilância inteligente, monitoramento de segurança, indústria de transporte aéreo, entre outras. O contexto internacional de crise sanitária provocada pela pandemia de Covid-19 impulsionou a procura pela biometria facial como uma alternativa sem contato à identificação de identidade, gerando também um aumento no interesse social pela alternativa (NORSTROM; CONSULTING, 2021).

A busca pelo reconhecimento ocorre também, possivelmente de forma mais intensa, por órgãos governamentais para rastreamento de foragidos da justiça e desaparecidos. Por exemplo, a prefeitura da cidade de São José do Campos investiu em óculos que contêm uma câmera, com uma de suas atribuições sendo o reconhecimento facial (PESSOA, 2022), o que resultou na localização de uma pessoa desaparecida, com problemas psiquiátricos (BAND VALE, 2022).

Dada a importância e relevância do tema tratado, o termo reconhecimento facial é de grande interesse da comunidade acadêmica, na busca de um sistema capaz de identificar indivíduos por fotos e vídeos. Este interesse pode ser evidenciado com os avanços observados nos últimos anos, como a taxa de erro de reconhecimento que passou de 4,1% em 2014 para 0,08% em abril de 2020 (CRUMPLER, 2020). Além disso, o espaço ocupado pelo reconhecimento facial no mercado global foi de 3,83 bilhões de dólares em 2020, e é esperado que chegue em um total de 16,74 bilhões de dólares em 2030 (YOGENDRA et al., 2022).

Em geral sendo aplicado no contexto de segurança, parte do interesse em desenvolvimento é dado na integridade de tais sistemas, quando alvejados por tentativas de fraude. Gent (2022) apresentou a iniciativa de uma conferência *hacking* de estimular o desenvolvimento de ataques ao reconhecimento facial, por meio de um desafio premiado, em busca de alimentar a pesquisa de contra-medidas que mitigariam fraudes milionárias referentes a sistemas de reconhecimento.

O contexto atual da tecnologia proporciona uma ameaça aos sistemas de reconhecimento facial, de modo que os impostores utilizam, nos atos de falsificação facial, dos dados e imagens disponíveis. Com o amplo uso das redes sociais, a aquisição de dados de biometria facial se tornou algo de simples realização. E com isso em evidência, a prevenção dos ditos ataques de apresentação, ou falsificação facial indireta, leva a um interesse maior na busca de prevenção de ataques, por se tratar de um processo de simples reprodução.

Os métodos atualmente desenvolvidos, que utilizam apenas das distorções presentes da captura de uma apresentação, são capazes de determinar a tentativa de fraude, mas tendem a sofrer com o contínuo avanço da tecnologia de telas, com cada vez mais qualidade de resolução e reprodução de cores. Desta forma, a busca por novas tecnologias é essencial na prevenção de ataques. Os métodos ativos são uma forma de agregar informação ao sistema de reconhecimento, com a integração de um dispositivo físico, na busca de melhorias no processo de detecção de fraudes.

1.1 Objetivos

1.1.1 Objetivo Geral

O presente projeto tem como objetivo geral o desenvolvimento e avaliação de métodos ativos de detecção de falsificação facial indireta, por apresentação de imagens em dispositivos eletrônicos, que utilizem da imagem capturada pelo sistema em conjunto com uma informação agregada a um dispositivo físico externo para a tomada de decisão.

1.1.2 Objetivos Específicos

- Investigar métodos ativos na solução do problema de falsificação facial eletrônica;
- Verificar o perfil de reflexão de diferentes aparelhos com tela digital;
- Implementar, por métodos de processamento de imagem, a detecção da reflexão de pontos luminosos em telas digitais;
- Avaliar a viabilidade de detecção de fraude em duas abordagens ativas distintas (uma com base em LEDs infravermelhos, outra com base em *ring light*);
- Desenvolver e avaliar o desempenho de um detector de fraude no reconhecimento facial, com base na abordagem considerada mais viável pela diferenciação de reflexões entre imagens de faces reais e imagens de telas de reprodução.

2 Revisão Bibliográfica

No desenvolvimento dos objetivos e das atividades presentes neste documento, faz-se necessária a utilização de conhecimentos previamente estabelecidos. A construção deste capítulo tem por finalidade expor conceitos básicos, de forma resumida, relacionados a processamento digital de imagens, apresentado as ferramentas principais utilizadas, além de definir detecção facial, reconhecimento facial e falsificação facial.

2.1 Processamento Digital de Imagens

A visão é o sentido do ser humano que capta estímulos visuais do ambiente na forma de luz, transmitindo-os para o cérebro para interpretação (GOLDSTEIN, 2009). Trata-se de um processo fundamental na percepção humana do mundo. A imagem é uma forma de armazenar essa informação visual, por meio de formas, cores e texto. De acordo com a teoria da comunicação visual (MARRIOTT; MEYER, 1998), a imagem é uma ferramenta de comunicação utilizada amplamente, desde a pré-história até os dias atuais, capaz de captar a atenção e transmitir a informação de forma objetiva, sendo utilizada em diversos contextos, como na arte e na ciência.

Uma imagem pode ser definida por uma função $f(x, y)$, em que o valor de f representa a intensidade de cinza e x e y a posição do elemento de intensidade f na matriz. Se os valores de f, x e y são quantidades finitas e discretas, então essa pode ser dita como uma imagem digital (GONZALEZ; WOODS, 2009). Por este ponto de vista, pode-se realizar diversas operações com a imagem ou entre imagens. Estas operações podem ter como resultado outra imagem, de mesmas dimensões ou dimensões diferentes da original, ou ainda resultar em valores vetoriais ou escalares. Tais resultados são buscados como uma forma de realizar modificações visuais na imagem original ou extrair informações, na forma de uma imagem modificada a ser analisada, ou valores que possam representar características da imagem.

Na análise de imagens digitais, de acordo com a definição de imagem como uma função, um *pixel* é o menor elemento da imagem, que pode ser representado por sua posição (x, y). A própria denominação vem do inglês *picture element*, que significa um elemento da imagem. Um *pixel* tem uma intensidade relacionada, quando se trata de uma imagem em escalas de cinza, ou ainda um valor vetorial que representa a intensidade de múltiplos canais, como em uma imagem colorida com canais vermelho, verde e azul.

O processamento digital de imagens é uma área da computação que engloba aquisição, análise e modificação de imagens (GONZALEZ; WOODS, 2009). As aplicações do processamento de imagens são diversificadas em várias áreas, com diferentes fontes de aquisição e objetivos de análise ou modificação. Normalmente, imagens digitais são relacionadas com a captação de radiação eletromagnética, de forma que o tipo de sensor e a banda específica de sensibilidade definem o tipo de imagem. Por exemplo, no campo da medicina, são utilizadas para diagnóstico imagens geradas por raios X e por ressonância magnética, que são baseadas em outra faixa do espectro utilizando ondas de rádio.

Outros sinais matriciais, não somente os provenientes de sensores eletromagnéticos, podem ser interpretados como imagens, utilizando das técnicas do processamento de imagens para análise da informação contida. Sendo assim, a fonte de uma imagem visual pode ser outra, como por exemplo imagens de ultrassom utilizadas para mapeamento de território utilizado na geologia, ou ainda imagens totalmente artificiais geradas por computador.

Uma grande área que se sustenta nas ferramentas do processamento de imagens é a visão computacional, em conjunto com outras grandes áreas como o aprendizado de máquina e a inteligência artificial. Ela consiste no desenvolvimento de algoritmos com capacidade de interpretar imagens e vídeos digitais. Com aplicações em robótica, em seus sistemas de navegação, a produção acadêmica em visão computacional contribui ativamente no desenvolvimentos de soluções comerciais, *e.g.* veículos autônomos (DAILY et al., 2017).

Outro importante ramo é a detecção e o reconhecimento de objetos, que utiliza principalmente do processamento de imagens em conjunto com o aprendizado de máquina para a construção de classificadores. Os classificadores em si são utilizados para detectar e rotular objetos ou características específicas em uma imagem (DUDA; HART; STORK, 2000). Eles são constituídos de três partes principais: características a serem extraídas, algoritmo de aprendizagem e um conjunto de dados de treinamento.

Em um tipo de aplicação, e com respeito à detecção de objetos ou segmentação, detecção de faces e, em um processo ainda mais complexo, pode-se ressaltar o reconhecimento facial. Como tópicos de notada importância para o problema atacado neste trabalho, serão tratados de forma mais detalhada posteriormente neste capítulo. A seguir, serão apresentados alguns métodos específicos de processamento de imagens utilizados durante o desenvolvimento dos métodos propostos.

2.1.1 Suavização

A suavização de uma imagem, aplicada por meio dos filtros de suavização, consiste no processo de "borramento" da imagem. Também referenciados como filtros passa-baixas, os filtros de suavização realizam a atenuação de ruídos de altas frequências, removendo pequenos detalhes e descontinuidades (GONZALEZ; WOODS, 2009). São comumente utilizados como etapa de pré-processamento, para a remoção de ruídos antes da realização de outros processos como os de detecção e reconhecimento, extraindo da imagem original a informação relevante na detecção e extração de características.

A implementação desta ferramenta pode ser feita de duas formas principais: filtros lineares e não lineares. A filtragem linear é realizada pela convolução com uma resposta impulsional no domínio bidimensional, sendo um caso em particular o filtro de média móvel, geralmente correspondente à média móvel ponderada dos *pixels* da vizinhança. Dos filtros não-lineares aplicados, o mais comum é o filtro de mediana, que possui um conceito geral parecido com os filtros de média (a saída é dada pela mediana da vizinhança), porém com uma complexidade de implementação maior e resultados que mantêm pequenos detalhes da imagem, com um aspecto de "borramento" menor, possivelmente mantendo informações relevantes, de acordo com a aplicação (GONZALEZ; WOODS, 2009).

A forma mais comum de implementação é pelo método linear. Trata-se de um filtro de média, de cálculo rápido, que pode ainda ser dividido em filtros de média simples ou ponderada. De forma geral, para o cálculo da média é definida previamente uma regra de vizinhança, que utiliza a distância entre *pixels* para determinar quais são utilizados no cálculo. Para a média ponderada, são atribuídos diferentes coeficientes para os *pixels* prévios à soma, de acordo com uma função peso que normalmente também depende desta distância, um exemplo é o filtro gaussiano, que utiliza uma função gaussiana para determinar os coeficientes (GONZALEZ; WOODS, 2009).

2.1.2 Limiarização

Considere uma imagem que contém um objeto iluminado em destaque sobre um fundo escuro, de modo que os valores de intensidade de *pixels* do objeto e do fundo correspondem a intervalos distintos (GONZALEZ; WOODS, 2009). Seria possível determinar a região de interesse, como a área da imagem que contém o objeto, pela simples seleção de *pixels* de intensidade presentes no intervalo do objeto.

Com base nessa ideia, a limiarização de uma imagem (do inglês *thresholding*) consiste na manutenção da imagem de acordo com o valor de intensidade e sua diferença entre regiões. Dados os valores de f como a imagem de entrada, g a imagem de saída e x e y a posição do *pixel* na matriz, o processo de limiarização com limiar T pode ser definida pela equação 2.1.

$$g(x,y) = \begin{cases} 255, & f(x,y) \geq T \\ 0, & f(x,y) < T \end{cases} \tag{2.1}$$

Suas aplicações são dadas normalmente na área de segmentação de imagens, determinando áreas de interesse. Um exemplo específico diz respeito ao aprimoramento de imagens de texto (documentos, etc), afetados pela diferença de iluminação em diferentes pontos, utilizando-se de métodos de limiarização local.

A implementação da limiarização depende de seu contexto de aplicação e pode ser determinada de maneira global ou local. A aplicação global utiliza de um único valor, previamente determinado pelo projetista ou processado de acordo com a imagem na busca de um valor ótimo ou de referência. Mesmo na busca de um valor ótimo, o valor limite global pode não considerar diferenças de iluminação durante a imagem. A busca do limite ótimo pode ainda ser feita em regiões, ou depender da sua vizinhança, de forma local, para um resultado mais específico de acordo com a aplicação.

2.1.3 Abertura

A abertura de uma imagem é o processo de remoção de pequenas regiões dessa imagem. É uma técnica composta de duas operações básicas de morfologia matemática: erosão e dilatação (GONZALEZ; WOODS, 2009). O objetivo da utilização desta técnica é remover pequenas regiões de ruído, normalmente de uma imagem previamente segmentada, para a seleção de regiões de interesse com uma área considerada relevante para a análise.

Em resumo, a operação de erosão realiza uma varredura de um elemento estruturante pelas regiões da imagem, de forma a manter os *pixels* em que, quando centrados no elemento estruturante, este esteja contido na região de interesse. O resultado desta operação em uma região é possivelmente a sua diminuição e uma suavização dos contornos, bem como a separação de regiões ligadas por linhas finas. Em contraste, a operação de dilatação faz a varredura do elemento estruturante pela região da imagem, em todos os seus *pixels* expandindo a região na forma do elemento pelo contorno. A dilatação também suaviza o contorno da região, enquanto a expande, possivelmente conectando regiões próximas.

Realizadas nessa sequência, a erosão elimina os pontos de área desprezível, reduzindo a área da região, enquanto a dilatação expande novamente a região para manter a área total da região de interesse, concluindo o processo de abertura na imagem. Paralelamente, pode-se também definir o processo de fechamento como a realização inversa da sequência, com o objetivo de remover pequenos furos internos à região.

2.2 Detecção Facial

Uma importante etapa no estudo e na análise facial por imagens, em especial o reconhecimento, é a detecção facial. Anterior ao reconhecimento, a detecção facial determina a presença de face na imagem e identifica a sua região, e a partir desta área podem ser feitos o processamento e as comparações necessárias à análise desejada. A detecção é feita pela classificação de objetos, identificando a existência de faces na imagem e determinando as coordenadas e a área total da face encontrada. O classificador é dado por uma série de descritores de características, que procuram por padrões na imagem, de acordo com o objetivo.

Uma dificuldade na classificação de objetos em imagens é a necessidade de realizar diversos cálculos em toda a matriz de *pixels*. Uma solução reconhecida pela comunidade científica é o algoritmo de Viola e Jones (2001), que além de propor um método eficiente de classificação, ainda apresenta uma forma de simplificar os cálculos utilizando a chamada imagem integral, sendo empregados em aplicações atuais pela simplicidade e eficácia.

A imagem integral, assim como a imagem digital original, pode ser definida como uma função $F(x, y)$ que representa o valor acumulado dos *pixels* acima e a esquerda da posição (x, y), calculado como na equação 2.2. Este valor acumulado permite o rápido cálculo do valor médio de intensidade em uma área retangular. A princípio, este retângulo é definido apenas com um de seus vértices fixado na origem do plano $(x, y) = (0, 0)$, mas com poucas operações de soma e subtração é possível definir o valor acumulado de qualquer retângulo na imagem.

$$F(x, y) = \sum_{i=0}^{x} \sum_{j=0}^{y} f(x, y) \tag{2.2}$$

Os valores médios de *pixels* de áreas retangulares adjacentes podem ser comparados em configurações específicas como forma de detecção de características na imagem. Algumas destas características, por sua vez, podem ser acumuladas para a decisão de um classificador na detecção de objetos. Este é o procedimento proposto por Viola e Jones (2001), baseado em classificadores *Haar-like* para detectar características da face humana. Alguns classificadores são disponibilizados para uso em projetos *open-source* pela biblioteca OpenCV[1]. Um destes detectores, o detector de face frontal, foi utilizado neste projeto.

[1] <https://github.com/opencv/opencv/tree/master/data/haarcascades>

2.3 Reconhecimento Facial

A importância do reconhecimento facial começa pela interação social, que utiliza da aparência de uma pessoa para o seu reconhecimento. Além de sua importância social, o reconhecimento facial levanta principalmente um interesse tecnológico, de forma que a identificação automática possibilita aplicações de cunho comercial e institucional.

No âmbito comercial, aplicações de controle de acesso são beneficiadas pela evolução do tema tratado. Haja vista aplicações como a entrada de edifícios, onde se faz necessária a coleta da identidade dos ingressantes para posterior verificação, ou ainda no controle de áreas restritas que necessitam de uma decisão para a liberação de acesso de acordo com o reconhecimento da identidade.

A aplicação institucional do reconhecimento pode ser exemplificada pela identificação de uma pessoa pelo sistema de vigilância governamental, podendo rastrear pessoas procuradas pela justiça ou desaparecidas (PESSOA, 2022). A utilização em sistemas de vigilância governamental fomenta a pesquisa na área, apesar de o rastreamento levantar discussões éticas em sua utilização, sobre o rastreamento em geral da população e a violação da privacidade individual.

Apesar da discussão ética na aplicação em larga escala do reconhecimento, a biometria facial aplicada em contextos restritos possibilita especialmente um fator de segurança, no sentido de determinação da identidade em aplicações de vigilância e controle de acesso, de maneira menos invasiva em comparação com as demais formas de biometria. O interesse na automatização da identificação minimamente invasiva potencializa a pesquisa na área da tecnologia em questão.

De forma geral, o processo de reconhecimento facial segue o fluxo apresentado na Figura 1, podendo ser dividido em algumas etapas principais: a detecção da face, a normalização da imagem (quanto a resolução, cor, entre outros aspectos) e a identificação da identidade pela comparação das características extraídas da imagem com a base de dados previamente estabelecida. O reconhecimento é feito, por exemplo, por meio do mapeamento da face em um espaço de características, sendo o vetor de características obtido comparado com a posição neste espaço de pessoas cadastradas para a tomada de decisão. Tolba, El-Baz e El-Harby (2008) fizeram uma revisão dos principais métodos explorados pela literatura para o problema do reconhecimento.

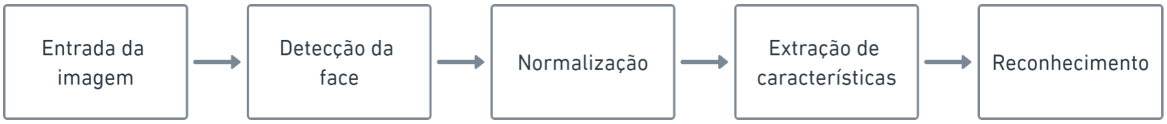


Figura 1 – Fluxograma do processo de reconhecimento facial.

Tratando-se de uma aplicação em segurança, ameaças aparecem explorando as vulnerabilidades do sistema. Rusia e Singh (2022) fizeram uma análise extensiva do material disponível em reconhecimento facial e suas ameaças. O avanço da tecnologia, com a popularização da internet e sua utilização na socialização, proporcionou o compartilhamento voluntário de dados pessoais através das redes sociais. Esta cultura das redes sociais coloca sistemas de reconhecimento em risco, com a facilidade de aquisição de dados de usuários, em especial dados de biometria como fotos, vídeos e gravações de voz. A utilização de dados de um usuário como forma de burlar um sistema de reconhecimento facial é conhecido como falsificação facial, e será melhor discutido a seguir.

2.4 Falsificação Facial

O termo falsificação facial (do inglês *face-spoofing*) é utilizado para caracterizar um ataque intencional a um sistema de reconhecimento facial. Trata-se de uma tentativa de burlar o sistema utilizando-se de técnicas para favorecer a detecção de características presentes na identidade de uma pessoa autorizada, para acesso indevido de um impostor.

A falsificação facial pode ser dividida em ataques diretos e indiretos (KUMAR; SINGH; KUMAR, 2017). O ataque direto, ou ataque 3D, é dado pela alteração visual do impostor, utilizando uma máscara ou semelhante. Ataques deste tipo demandam uma maior sofisticação do impostor na reprodução de características, de forma que a maior parte dos ataques é de forma indireta (PEREIRA et al., 2013).

O ataque indireto, também chamado de ataque 2D ou ataque de apresentação, é feito pela reprodução de uma imagem da pessoa autorizada, por foto ou vídeo. Este modo de ataque levanta um maior interesse e preocupação pela sua simplicidade, ainda mais no contexto atual e na disponibilidade de dados biométricos através das redes sociais. A falsificação facial eletrônica, se refere ao ataque indireto pela apresentação da imagem por um dispositivo com uma tela digital, e.g. computadores, *smartphones*, etc.

Diversos sistemas de reconhecimento se baseiam em imagens 2D e processamento de imagens para determinar o resultado. Em especial, o ataque indireto oferece grande ameaça aos sistemas automáticos, que possuem apenas as informações de imagem coletadas para determinar se o que é apresentado num dado momento é a imagem da pessoa ou de uma foto dela.

Algumas técnicas de reconhecimento e prevenção de violação estão disponíveis na literatura. A técnica de redução de dimensionalidade reduz o espaço de características, consequentemente reduzindo a quantidade de variáveis no cálculo de reconhecimento (MEYTLIS; SIROVICH, 2007). A extração de características é uma importante etapa na detecção e reconhecimento facial, para a representação da imagem por um vetor de variáveis (WANG; HU; DENG, 2018).

Em destaque no objetivo deste documento, faz-se necessário a revisão de técnicas disponíveis na prevenção de ataques indiretos. Técnicas de classificação são utilizadas para determinar a presença de ataques de falsificação, por exemplo o algoritmo proposto por Garcia e Queiroz (2015), que busca pelo padrão *moiré*, uma distorção característica de capturas de reprodução por telas digitais. A maior parte do material disponível na prevenção destes ataques busca por diversas distorções presentes, sendo técnicas robustas que em geral apresentam bons resultados. Entretanto, investigar novas metodologias é essencial na busca de manter a robustez dos sistemas quanto às novas tecnologias, como o avanço na resolução das telas e na qualidade de reprodução.

Metodologias ativas, como as propostas neste documento, utilizam informações complementares para além da imagem única capturada no sistema. Mhou, Haar e Leung (2017) apresentam um método que utiliza duas câmeras, como fator de segurança, sendo uma câmera auxiliar infravermelha. A análise foi baseada no fato de diferentes materiais apresentarem perfis de reflexão distintos, indicando se os resultados foram positivos em um ambiente controlado com baixa luminosidade.

Esse perfil de reflexão infravermelho é utilizado também pela tecnologia comercial Face-ID da empresa Apple (2022). Por se tratar de uma solução comercial, informações oficiais sobre seu funcionamento são limitadas. A empresa promete com sua câmera *TrueDepth*, através da reflexão infravermelho, realizar um mapeamento espacial do rosto em um modelo 3D, que é então utilizado para a caracterização da face em um espaço de características. Assim, a análise de profundidade e de textura, providas pelas leituras desta câmera, são fatores que contribuem na prevenção de ataques de falsificação.

3 Métodos propostos

A partir dos objetivos definidos e dos conceitos apresentados, pretende-se desenvolver métodos ativos de detecção da falsificação, utilizando da imagem capturada pelo sistema, bem como de uma informação agregada por um dispositivo externo ao sistema de reconhecimento.

A proposta do trabalho é averiguar a aplicabilidade de alternativas, comparativamente as soluções ativas apresentadas na seção 2.4, por exemplo na detecção de reflexão de um LED (Diodo Emissor de Luz, do inglês *Light emitter diode*) no caso da falsificação por reprodução eletrônica. A Figura 2 exemplifica este objetivo de detecção, apesar de não ser adequada para a análise do método, uma vez que as imagens não foram capturadas em condições similares de iluminação e equipamento. Na Figura 2b, em que ocorre a captura da reprodução da imagem frente a um LED aceso, é possível perceber uma concentração da reflexão da iluminação.

Dessa forma, o desenvolvimento dos métodos se baseia na característica de reflexão de fontes luminosas comparativamente entre a captura real de uma pessoa e a captura de uma reprodução da foto da pessoa, potencialmente obtidas por exemplo em redes sociais, com o intuito de falsificar a identidade. Serão então propostos dois métodos de detecção, por sensoriamento analógico e por sensoriamento digital, apresentados a seguir.

(a) (b)

Figura 2 – Fotografia para análise de reflexão do LED na tela digital. (a) Imagem original com a face de uma pessoa, retirada de uma base de dados aberta (WEYRAUCH et al., 2004); (b) Reprodução eletrônica da imagem capturada por uma câmera com o *flash* ativado.

3.1 Detecção por Sensoriamento Analógico

O primeiro método abordado, em concordância com os métodos disponíveis, utiliza-se da reflexão de ondas no espectro do infravermelho. O uso de dispositivos neste espectro é atrativo por não ser perceptível ao olho nu, de forma que a experiência final do usuário do sistema não é afetada. Apesar disto, além de não ser perceptível ao olho da pessoa fotografada, a percepção das ondas infravermelho por uma câmera comum existe, mas é baixa e na reflexão da tela também não é perceptível.

A detecção por sensoriamento analógico proposta é composta por conjuntos de emissores e receptores na faixa do infravermelho para a detecção automática da falsificação por um dispositivo físico, de maneira independente à câmera. A arquitetura geral é então composta pelo dispositivo de detecção e pela câmera, ambos em comunicação com o sistema de reconhecimento, de forma que o dispositivo de detecção possa invalidar o reconhecimento, ou ainda bloquear o fluxo geral do sistema enquanto detectada a reflexão do sinal infravermelho. A expectativa nesse método é captar através do receptor a intensidade luminosa e é esperado que essa seja maior no caso da presença de uma tela digital, em comparação com a presença de uma pessoa, como exemplificado na Figura 2. Esse aspecto esperado é avaliado nos experimentos conduzidos nesse trabalho.

Tratando isoladamente do dispositivo de detecção, os materiais utilizados foram emissores IR204-A[1], receptores PT204-6C[2] e uma placa Arduino Nano[3], que é uma plataforma construída em torno do microcontrolador ATmega328. Ele é utilizado para realizar a conversão analógica-digital do sinal dos receptores para posterior análise.

Para a aplicação do método, foram construídos dois dispositivos distintos, o primeiro com apenas um par de emissor e receptor, próximos e quase paralelos, representando o caso mais simples (Figura 3), sendo o receptor à esquerda lido pela porta analógica do Arduíno. Já o segundo dispositivo foi realizado com 4 receptores na área central, com 4 emissores espaçados do centro igualmente e ocupando as posições a 90° entre si em um círculo (Figura 4). O incremento de múltiplos emissores e receptores adiciona redundância ao sistema, com o objetivo de aumentar a robustez quanto às variações de distância e angulação da tela.

Os dispositivos desenvolvidos foram inicialmente testados de forma isolada, com o objetivo de identificar primeiramente a possibilidade de detecção da reflexão, pela intensidade luminosa percebida pelo receptor, em telas de dispositivos eletrônicos comerciais disponíveis, para depois serem integrados ao sistema com a câmera e o processador do sistema de reconhecimento. Os resultados serão apresentados no Capítulo seguinte.

[1] <https://en.everlight.com/wp-content/plugins/ItemRelationship/product_files/pdf/IR204-A.pdf>
[2] <https://en.everlight.com/wp-content/plugins/ItemRelationship/product_files/pdf/PT204-6C.pdf>
[3] <https://store-usa.arduino.cc/products/arduino-nano>

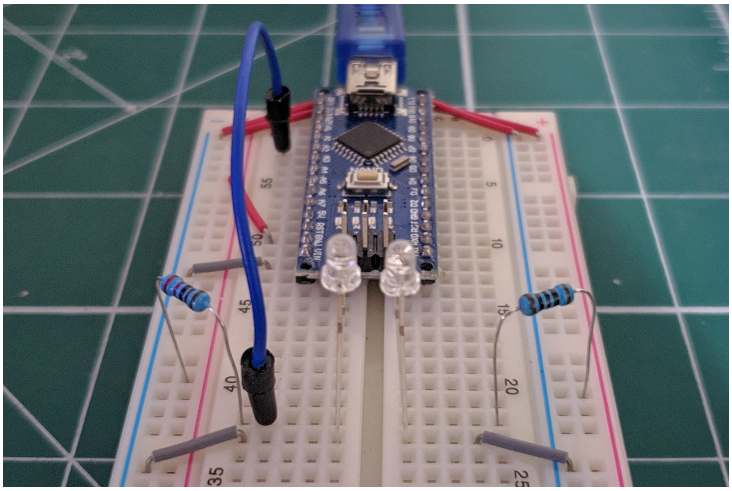


Figura 3 – Dispositivo sensor de reflexão proposto com um par emissor/receptor.

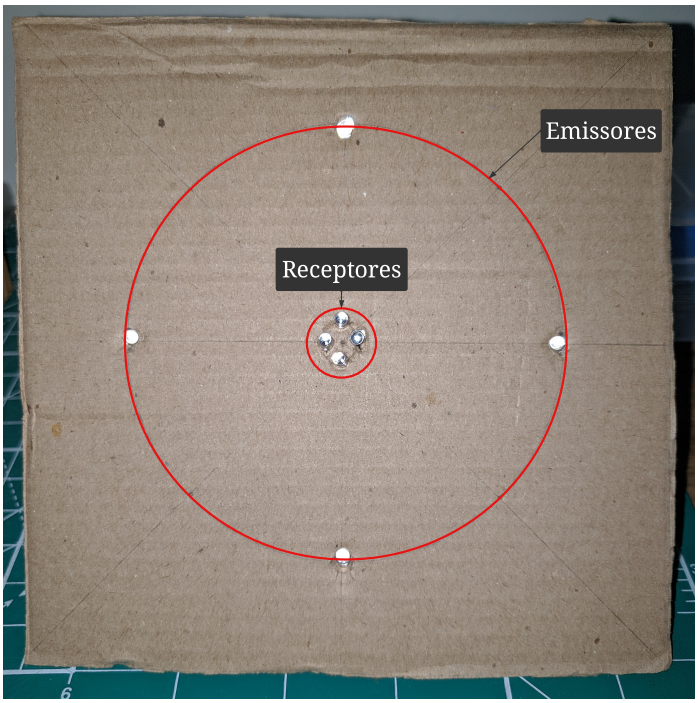


Figura 4 – Dispositivo sensor de reflexão proposto com quatro pares emissor/receptor, com os receptores concentrados no centro do círculo de emissores.

3.2 Detecção por Sensoriamento Digital

A solução por sensoriamento digital consiste em detectar a concentração da reflexão por técnicas de processamento de imagens. O fluxograma do algoritmo desenvolvido é apresentado na Figura 5, que consiste em buscar regiões de pontos luminosos na imagem, baseado no projeto de Rosebrock (2016). Posteriormente ao feito no projeto está a agregação da informação da imagem final pela média dos *pixels* para a tomada de decisão sobre a detecção de falsificação.

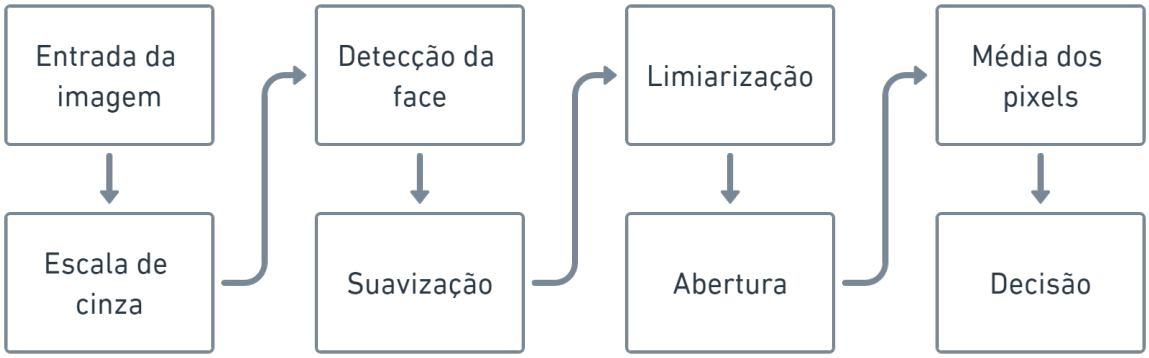


Figura 5 – Fluxograma do processo proposto de detecção de falsificação por sensoriamento digital.

As etapas de imagem do processo apresentado são exemplificadas pela Figura 6, de uma forma sequencial, como no fluxograma. Todas as etapas do fluxogramas serão melhor detalhadas e justificadas. A partir da entrada da imagem na Figura 6a, a imagem foi convertida em escala de cinza na Figura 6b, para simplificar o processo computacional, que agora segue trabalhando com apenas um canal de intensidade por *pixel* e para uma normalização da implementação.

A Figura 6b passa então pelo processo de detecção da face, que é recortada na Figura 6c, de forma a concentrar a região de interesse da aplicação e reduzir a influência do fundo da imagem no restante do processo. Esta imagem recortada é então suavizada para a Figura 6d, com o objetivo de eliminar ruídos de altas frequências na imagem, na forma de *pixels* isolados com valor incompatível com a sua vizinhança.

Partindo da imagem suavizada da Figura 6d, ocorre a etapa de limiarização da imagem (Figura 6e), que transforma a imagem em binária a partir de uma valor limite de intensidade, reduzindo todos os valores abaixo para o mínimo de intensidade, na cor preta, e os valores acima no máximo de intensidade, na cor branca. Esta é a principal etapa na detecção dos pontos luminosos, de forma que o limiar definido deve separar as fontes de luz e suas reflexões dos demais objetos ou pessoas da foto. Por fim a Figura 6e passa pela abertura, para remover regiões de área desprezível para a aplicação com resultado na Figura 6f.

Figura 6 – Exemplo de aplicação do método proposto para a detecção por sensoriamento digital, dividido em etapas sequenciais. (a) Imagem de entrada, contendo a região da face de uma pessoa (WEYRAUCH et al., 2004); (b) Imagem em escala de cinza; (c) Recorte da detecção da face; (d) Imagem suavizada; (e) Imagem binária após a limiarização; (f) Abertura da imagem binária. Nas imagens (a)-(d), as regiões brancas na região da face foram adicionadas somente neste texto para a não-identificação do indivíduo, não sendo parte do algoritmo de detecção.

A etapa de limiarização na Figura 6e foi ajustada para fins de exemplificação do processo, com um valor de limiar baixo. Idealmente esse limiar deve ser tal que proporcione a separação apenas de regiões correspondentes aos pontos luminosos. O método aplicado pode ser visto na Figura 7 que foi retirado da base utilizada nos testes que serão apresentados no próximo Capítulo, capturadas em configurações semelhantes, com uma pessoa e uma reprodução. A partir da imagem final, binarizada e aberta, a decisão pode ser feita a partir da área total de regiões com valor alto, para isto está a etapa do cálculo da média dos *pixels* apresentado na Figura 5.

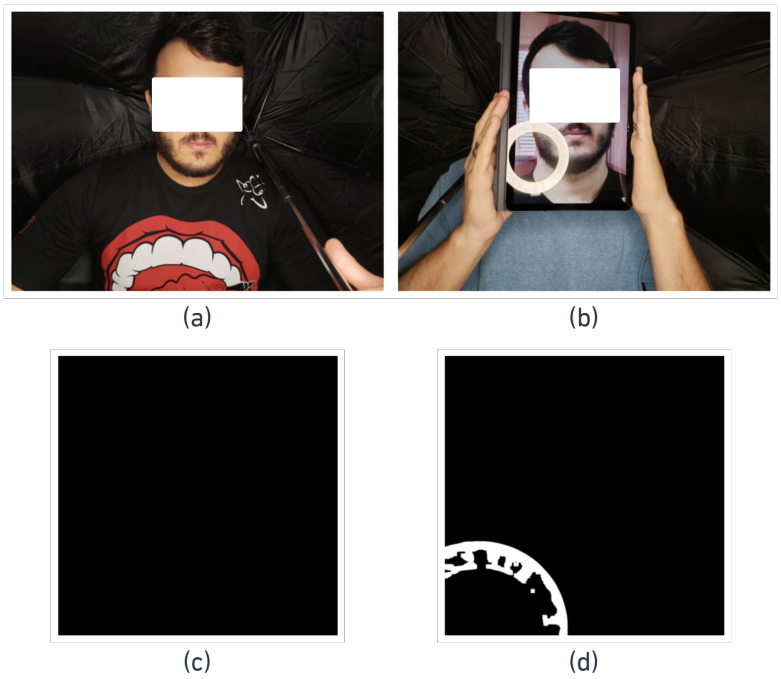


Figura 7 – Método aplicado nas situações da imagem (a) de uma pessoa e (b) da reprodução de uma foto. Os resultados são (c) e (d), respectivamente. Nas imagens de entrada, as regiões brancas na região da face foram adicionadas somente neste texto para a não-identificação do indivíduo, não sendo parte do algoritmo de detecção.

4 Testes e Resultados

4.1 Detecção por Sensoriamento Analógico

Os testes foram realizados com os dispositivos, de acordo com a configuração das Figuras 3 e 4, e os resultados serão apresentados pela leitura do conversor analógico-digital do Arduíno, do fototransistor referente ao dispositivo testado. No caso do dispositivo com 4 receptores, o valor apresentado é referente ao mínimo dentre os 4 valores lidos. Os testes foram realizados de forma a verificar a reflexão de dois aparelhos, o *smartphone* Samsung Galaxy S9+ e o *tablet* Samsung Galaxy Tab S6 Lite, além de um caso ideal que se realizou usando um espelho.

A apresentação do valor de saída do conversor permite uma análise da variação na leitura, de forma a determinar a viabilidade da construção de um algoritmo de detecção automático a partir do dispositivo construído. Para a interpretação dos dados, a tensão máxima em 5V lida do fototransistor, convertida para o valor digital de 1023, representa uma corrente nula no ramo e por consequência baixa captação de fótons. Já a leitura de valores baixos representa uma baixa resistência por parte do fototransistor e uma consequente alta captação dos fótons. Assim, nos resultados, valores mais altos indicam não-reflexão, enquanto valores baixos indicam reflexão.

As Figuras 8 e 9 apresentam os resultados obtidos para os dispositivos apresentados anteriormente, dividido em imagens (a) e (b), correspondentes as Figuras 3 e 4, respectivamente. É possível notar uma melhora na leitura entre os dois dispositivos, comparando os resultados das imagens em uma mesma Figura, tanto para o teste de distância, quanto no de angulação. Entretanto, a faixa de distância analisada foi escolhida para observar essa variação relevante na leitura dentro de um mesmo gráfico, mas se tratam de valores muito baixos, fora do escopo de aplicação. Considerando ainda que, na ausência dos aparelhos reflexivos analisados, situação tomada como valor máximo lido mesmo que na ausência de uma face para comparação dos valores, a leitura para ambos os dispositivos foi de 953, que se trata de valores próximos aos da curva do *tablet*, que obteve uma leitura de 941 e 878 para os testes na distância maior de 15 centímetros. Além disto, ocorreu também uma diferença entre as curvas dos dois dispositivos analisados, o que dificultaria uma aplicação geral do método pela inconsistência entre diferentes aparelhos.

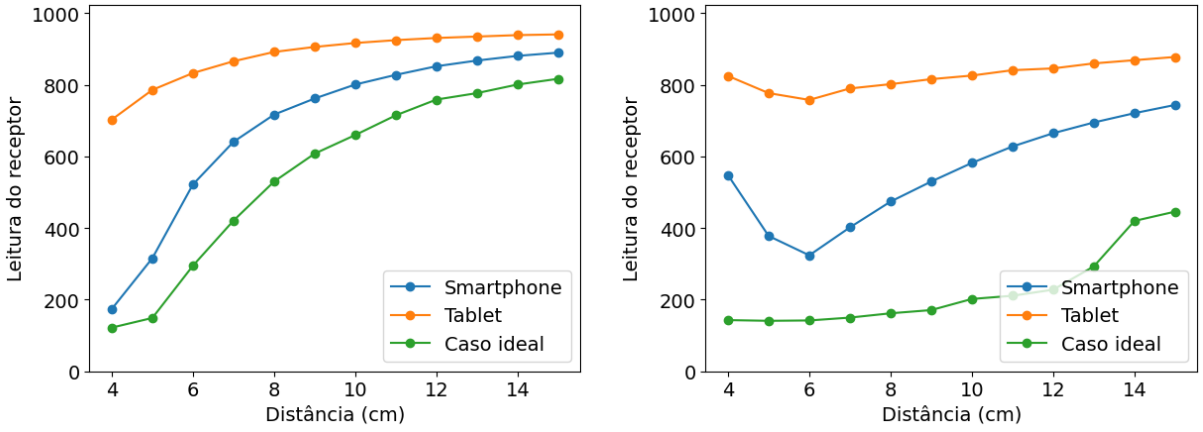


Figura 8 – Resultados para o teste de variação da distância na detecção por sensoriamento analógico (a) para o primeiro dispositivo e (b) para o segundo dispositivo, ambos mantidos paralelos à tela. Os valores do eixo y são dados pela leitura de tensão do receptor infravermelho pelo conversor analógico-digital de 10 bits, resultando na escala de 0 a 1023, de forma que o valor lido é inversamente proporcional à corrente no fototransistor ou a intensidade captada pela reflexão.

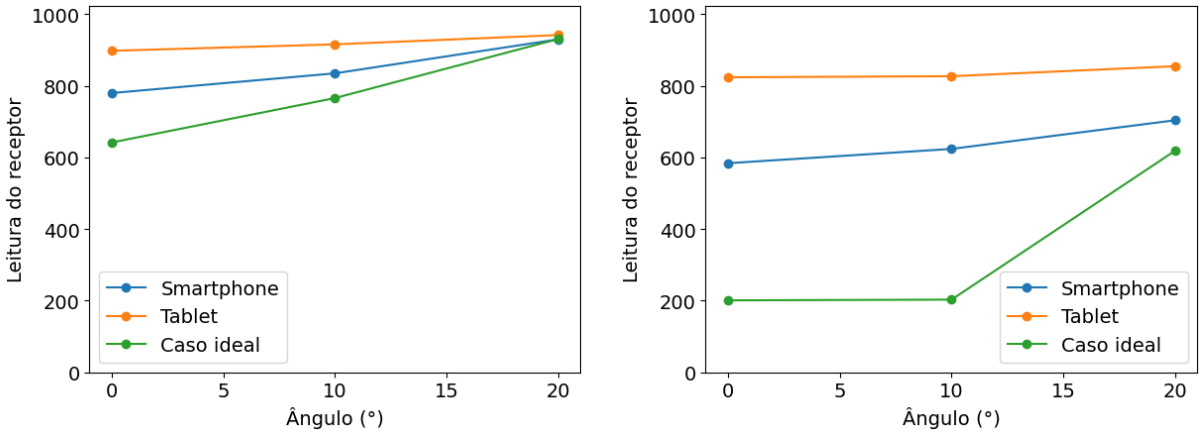


Figura 9 – Resultados para o teste de variação da angulação entre o dispositivo e a tela de reprodução na detecção por sensoriamento analógico (a) para o primeiro dispositivo e (b) para o segundo dispositivo, ambos a uma distância fixa de 10cm. Os valores do eixo y são dados pela leitura de tensão do receptor infravermelho pelo conversor analógico-digital de 10 bits, resultando na escala de 0 a 1023, de forma que o valor lido é inversamente proporcional à corrente no fototransistor ou a intensidade captada pela reflexão.

4.2 Detecção por Sensoriamento Digital

Toda a implementação[1] do roteiro de execução do fluxo proposto e dos testes realizados para o método de detecção por sensoriamento digital foi realizada utilizando a linguagem de programação Python, com os processos de processamento de imagens por meio da biblioteca OpenCV (BRADSKI, 2000). Além dos códigos-fonte de implementação, o repositório do método oferece os dados de entrada da solução, bem como de saída para o valor de limiar ótimo encontrado, além de um relatório geral da execução. Os dados de saída totais acumulam mais de 2500 imagens, portanto não foram adicionados em sua totalidade. Também foram adicionados os arquivos de análise, contendo o roteiro de geração das Figuras de resultado deste trabalho. Estes arquivos estão no formato Python *notebook* na pasta raiz do repositório.

Para validação do método de detecção por sensoriamento digital foram tiradas fotos na construção de um banco próprio a este trabalho. A não-utilização de banco público padronizado foi escolhida para manter a isonomia das condições de captura das imagens de pessoas e de fotos reproduzidas. Para a estruturação deste banco foram considerados alguns parâmetros:

- Cor do fundo: branco ou preto;
- Iluminação ambiente: claro ou escuro;
- Distância da câmera: 40cm, 60cm ou 80cm.

Estes parâmetros intercalados resultam em um total de 12 imagens. Além disto, foram capturadas fotos de duas pessoas, em presença real e com a reprodução de uma foto. Dadas todas as variáveis de ambiente, foram consideradas, no total, 48 imagens de entrada do sistema de detecção, sendo metade em tentativa de falsificação.

As imagens foram capturadas em um ambiente interno com iluminação controlada, com a variação do fundo por meio de um tecido preto. Como fonte luminosa, para a tentativa de captar a reflexão, foi utilizado um dispositivo *ring light* genérico[2], que provê iluminação em um formato circular. Para a captura se utilizou a câmera traseira do *smartphone* Samsung Galaxy S9+, posicionada no centro e atrás do *ring light*. A utilização desta iluminação, independente da iluminação ambiente, circulando a câmera, tem por objetivo adicionar robustez às variações de ângulo da tela de reprodução, sendo uma área de fonte luminosa maior para facilitar a detecção. A reprodução das imagens foi realizada pela tela digital do *tablet* Samsung Galaxy Tab S6 Lite.

[1] <https://github.com/l-pires/TCC>
[2] <https://www.amazon.com.br/Iluminador-Ring-Light-Polegadas-Misto/dp/B07TSFLCZR>.

Além das variáveis citadas na geração dos dados de análise, existem na implementação do sistema duas variáveis principais para ajuste: o valor de limiar de binarização da imagem e o valor de decisão para comparação com o média de *pixels* da imagem final. Estas foram então as variáveis ajustadas para a visualização dos resultados, mais especificamente foram testados os valores de limiar no intervalo $[200, 255]$ em valores inteiros e para o valor de decisão o intervalo $[0, 100]$. Valores além deste intervalo foram experimentados inicialmente, mas valores de limiar menores e de decisão maiores tendem a reduzir o desempenho geral do método, sendo apresentado o intervalo relevante para a análise feita.

Foram salvas as 48 imagens com resolução de 4032x3024 *pixels*, para entrada do algoritmo, nas condições de testes definidas. Para cada valor de limiar o algoritmo desenvolvido gerou uma imagem de saída por imagem de entrada, contendo nesta uma réplica da imagem original, imagem em escala de cinza, recorte da face e a imagem final. Considerando o intervalo para o limiar em valores inteiros e a quantidade de imagens de entrada, foram geradas 2688 imagens de saída. Além das imagens de saída, foi salvo um relatório com os valores de média da imagem final para cada caso, para análise posterior e construção dos resultados. O algoritmo foi executado em um processador Intel Core i7-7500U, com 16GB de RAM (Memória de acesso aleatório, do inglês *Random Access Memory*) em frequência de 2133MHz e uma placa gráfica AMD Radeon R7 M445 com 4GB de memória dedicada no sistema operacional Windows 10. O processo de geração das imagens de saída e do relatório, para os 56 valores distintos de limiar, durou em torno de 3 horas de execução.

Para cada par de valores de limiar e de decisão, determinou-se a quantidade dos resultados para a detecção, sendo verdadeiro positivo V_p a detecção correta em uma imagem de reprodução, verdadeiro negativo V_n a não-detecção em uma imagem da pessoa, falso positivo F_p a detecção em uma imagem da pessoa e falso negativo F_n a não-detecção em uma imagem de reprodução. A partir destes valores, foram calculadas métricas de avaliação de acurácia e F1-score (RODRIGUES, 2019) para determinação de um valor das variáveis em que se encontra o máximo destas métricas, nas condições analisadas:

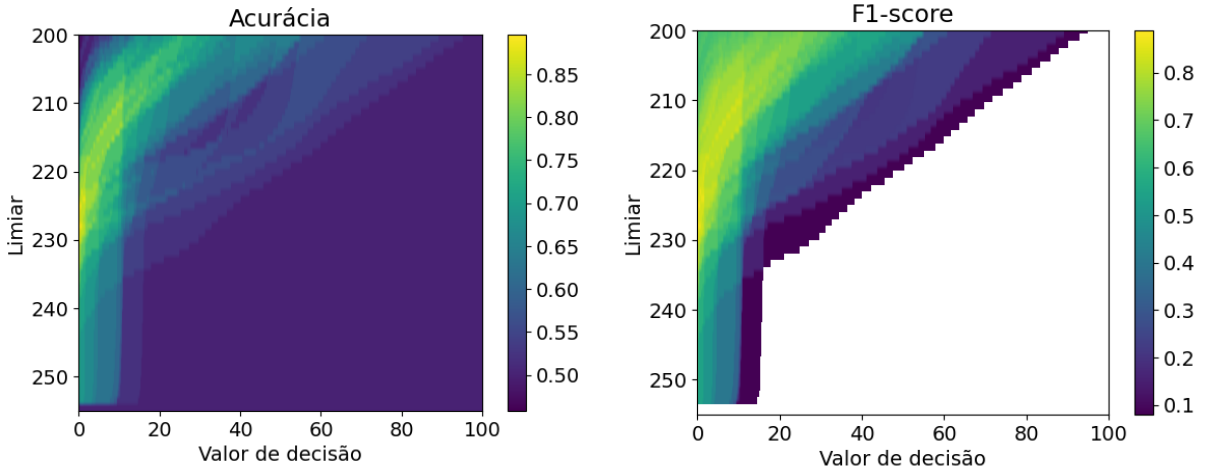
$$\texttt{Acurácia} = \frac{V_p + V_n}{V_p + V_n + F_p + F_n} \tag{4.1}$$

$$\texttt{F1-score} = \frac{1}{1 + \dfrac{F_p + F_n}{2V_p}} \tag{4.2}$$

Os resultados para o método de detecção por sensoriamento digital são apresentados na Figura 10, com as métricas citadas anteriormente. Nestes gráficos, cada ponto representa uma relação entre os valores de limiar e de decisão, sendo então a métrica

calculada pela aplicação destes valores nas 48 imagens. Para a métrica F1-score, com resultados na Figura 10b, o seu valor depende da ocorrência de resultados positivos, entretanto, para combinações de valor de limiar e de decisão maiores estes resultados não aparecem, gerando uma indeterminação no cálculo e a região branca presente na Figura. O valor máximo de ambas as métricas é dado pela mesma posição, correspondente ao limiar de 225 e um valor de decisão de 0,4008. O valor de acurácia máximo foi de 89,6% e de F1-score de 88,9%.

Apesar da tentativa adicionar robustez quanto à angulação da tela com a utilização do *ring light*, em distâncias maiores a tela de reprodução precisa estar paralelamente alinhada a câmera para ocorrer a captura da reflexão. Apesar disto, as regiões de valor alto após a limiarização são mais comuns nas capturas de reprodução, possivelmente pela própria iluminação da reprodução pela tela digital. Outro fenômeno foi observado em distâncias menores, em que a chance de captura da reflexão é maior, a depender da posição da reflexão na imagem reproduzida ocorreram casos em que a reflexão impossibilitou a detecção da face pelo classificador, fato este que bloquearia também o fluxo do reconhecimento.



(a) (b)

Figura 10 – Resultados para o método de detecção por sensoriamento digital, em mapas de cores. (a) Acurácia; (b) F1-score.

5 Conclusão

A partir do objetivo proposto, duas abordagens foram exploradas na busca de uma solução ao problema. No método de detecção por sensoriamento analógico, nenhum critério objetivo foi posto para análise, entretanto os dados aferidos apontam que um dispositivo deste tipo, com os materiais utilizados neste trabalho, teria um intervalo de distância de aplicação curto e para valores baixos, em comparação com valores típicos em um sistema de reconhecimento, uma vez que no intervalo testado de 4 a 15 centímetros o dispositivo apresentou uma variação em toda a faixa de leitura. Ademais, tal dispositivo ainda produziu efeitos de reflexão distintos entre os aparelhos testados, em que um dos aparelhos foram lidos os valores 941 e 878, para dois dispositivos desenvolvidos na maior distância, em comparação com o valor de 953, lido na ausência dos aparelhos de apresentação. Pela distância curta e inconsistência de resultado entre aparelhos, o método foi considerado menos viável dentre os propostos.

Já o método de detecção por sensoriamento digital pode ser analisado de acordo com suas métricas de avaliação. Com uma acurácia máxima de 89,6%, os valores ótimos de aplicação foram um limiar de 225 de intensidade de *pixel* e um valor de decisão para a média de *pixels* da imagem final de 0,4008. As métricas obtidas indicam que houve casos em que a reflexão impossibilitou a detecção de face, o que impossibilitaria também o reconhecimento, mas que resultou em uma imagem final sem região de luminosidade, reduzindo a acurácia aferida.

Em trabalhos futuros, para o método de sensoriamento analógico, é possível fazer a análise alterando os componentes, principalmente na busca de ajuste de intensidade, testando novamente a captação de intensidade em diferentes aparelhos de reprodução e adicionando uma comparação real de captura com o caso de presença da face. Além disso, pode ser feitos experimentos com outras quantidades de emissores e receptores, em configurações diferentes deste texto, em busca de uma condição de aplicação viável. Para o método digital, pode-se estudar melhor a relação da detecção por limiarização da imagem, considerando também o brilho da tela de reprodução, controlando as condições definidas, principalmente de iluminação ambiente, e analisar a influência de cada variável de condição de captura nas métricas de avaliação.

Referências

APPLE. *About Face ID advanced technology*. United States: Apple Inc., 2022. <https://support.apple.com/en-us/HT208108>. Acesso em 30 de janeiro de 2023. Citado na página 22.

BAND VALE. *Desaparecido em São José é localizado com tecnologia de reconhecimento facial*. 2022. <https://www.band.uol.com.br/band-vale/noticias/desaparecido-em-sao-jose-e-localizado-com-tecnologia-de-reconhecimento-facial-16533629>. Acesso em 12 de setembro de 2022. Citado na página 13.

BRADSKI, G. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000. Citado na página 31.

CRUMPLER, W. *How Accurate are Facial Recognition Systems – and Why Does It Matter?* 2020. <https://www.csis.org/blogs/technology-policy-blog/how-accurate-are-facial-recognition-systems-%E2%80%93-and-why-does-it-matter>. Acesso em 12 de setembro de 2022. Citado na página 13.

DAILY, M. et al. Self-driving cars. *Computer*, v. 50, n. 12, p. 18–23, 2017. Citado na página 16.

DUDA, R. O.; HART, P. E.; STORK, D. G. *Pattern Classification*. 2nd. ed. New Jersey, United States: Wiley Interscience, 2000. ISBN 978-0-471-05669-0. Citado na página 16.

GARCIA, D. C.; QUEIROZ, R. L. D. Face-spoofing 2D-detection based on moiré-pattern analysis. *IEEE Transactions on Information Forensics and Security*, Institute of Electrical and Electronics Engineers Inc., v. 10, p. 778–786, 4 2015. ISSN 15566013. Citado na página 22.

GENT, E. *Hackers Compete To Confound Facial Recognition*. 2022. <https://spectrum.ieee.org/facial-recognition>. Acesso em 5 de setembro de 2022. Citado na página 13.

GOLDSTEIN, E. *Sensation and Perception*. Boston, United States: Cengage Learning, 2009. ISBN 9780495601494. Citado na página 15.

GONZALEZ, R. C.; WOODS, R. E. *Processamento Digital De Imagens*. São Paulo: Pearson, 2009. ISBN 9788576054016. Citado 4 vezes nas páginas 15, 16, 17 e 18.

KUMAR, S.; SINGH, S.; KUMAR, J. A comparative study on face spoofing attacks. *Proceeding - IEEE International Conference on Computing, Communication and Automation, ICCCA 2017*, Institute of Electrical and Electronics Engineers Inc., v. 2017-January, p. 1104–1108, 12 2017. Citado na página 21.

MARRIOTT, K.; MEYER, B. *Visual Language Theory*. New York, United States: Springer, 1998. ISBN 978-1-4612-7240-3. Citado na página 15.

MEYTLIS, M.; SIROVICH, L. On the dimensionality of face space. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 29, p. 1262–1267, 7 2007. ISSN 01628828. Citado na página 21.

MHOU, K.; HAAR, D. van der; LEUNG, W. S. Face spoof detection using light reflection in moderate to low lighting. In: *2017 2nd Asia-Pacific Conference on Intelligent Robot Systems (ACIRS)*. [S.l.: s.n.], 2017. p. 47–52. Citado na página 22.

NORSTROM, P.; CONSULTING, A. *Has Covid increased public faith in facial recognition?* 2021. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8612370/>. Acesso em 12 de setembro de 2022. Citado na página 13.

PEREIRA, T. de F. et al. Can face anti-spoofing countermeasures work in a real world scenario? In: *2013 International Conference on Biometrics (ICB)*. [S.l.: s.n.], 2013. p. 1–8. Citado na página 21.

PESSOA, P. *Tecnologia amplia a segurança no aniversário de São José*. Secretaria de Proteção ao Cidadão: [s.n.], 2022. <https://www.sjc.sp.gov.br/noticias/2022/julho/27/tecnologia-amplia-a-seguranca-no-aniversario-de-sao-jose/>. Acesso em 12 de setembro de 2022. Citado 2 vezes nas páginas 13 e 20.

RODRIGUES, V. *Métricas de Avaliação: acurácia, precisão, recall... quais as diferenças?* 2019. <https://vitorborbarodrigues.medium.com/m%C3%A9tricas-de-avalia%C3%A7%C3%A3o-acur%C3%A1cia-precis%C3%A3o-recall-quais-as-diferen%C3%A7as-c8f05e0a513c>. Acesso em 10 de janeiro de 2023. Citado na página 32.

ROSEBROCK, A. *Detecting multiple bright spots in an image with Python and OpenCV*. 2016. <https://pyimagesearch.com/2016/10/31/detecting-multiple-bright-spots-in-an-image-with-python-and-opencv/#download-the-code>. Acesso em 6 de dezembro de 2022. Citado na página 26.

RUSIA, M. K.; SINGH, D. K. A comprehensive survey on techniques to handle face identity threats: challenges and opportunities. *Multimedia Tools and Applications*, Springer Science and Business Media LLC, 6 2022. ISSN 1380-7501. Citado na página 21.

TOLBA, A. S.; EL-BAZ, A.; EL-HARBY, A. Face recognition: A literature review. *International Journal of Computer and Information Engineering*, v. 2, p. 2556–2571, 7 2008. Citado na página 20.

VIOLA, P.; JONES, M. Rapid object detection using a boosted cascade of simple features. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2001. ISSN 10636919. Citado na página 19.

WANG, H.; HU, J.; DENG, W. Face feature extraction: A complete review. *IEEE Access*, Institute of Electrical and Electronics Engineers Inc., v. 6, p. 6001–6039, 2018. ISSN 21693536. Citado na página 21.

WEYRAUCH, B. et al. Component-based face recognition with 3d morphable models. In: *2004 Conference on Computer Vision and Pattern Recognition Workshop*. [S.l.: s.n.], 2004. p. 85–85. Citado 2 vezes nas páginas 23 e 27.

YOGENDRA, B. et al. *Facial Recognition Market*. 2022. <https://www.alliedmarketresearch.com/facial-recognition-market>. Acesso em 12 de setembro de 2022. Citado na página 13.