

Aprendizado de Máquina para o Auxílio à Localização de Pessoas em Ambientes *Indoor* Monitorados por Câmeras

Luiz C. C. Moitinho¹, Alcides X. Benicasa¹

¹Departamento de Sistemas de Informação – Universidade Federal de Sergipe (UFS)
Caixa Postal – 49.506-036 – Itabaiana – SE – Brasil

Abstract. *Security is a topic of great relevance in society, for this, the development of studies and the use of mechanisms have become essential to assist this social right, such as surveillance camera systems commonly adopted to monitor environments. The use of this technology in conjunction with Artificial Intelligence (AI) was responsible for locating 600 criminals wanted by the government of Bahia, reaching a degree of similarity of 93% when recognizing an individual. In order to contribute to the surveillance scenario, this paper discusses an intelligent model based on microservices architecture and asynchronous communication responsible for performing facial detection and recognition, reaching, respectively, accuracies of 92.73% and 69.23% in images captured from a real environment monitored from microcontrollers with attached cameras, in order to perform the monitoring of consenting people in an environment.*

Resumo. *A segurança é um tema de grande relevância na sociedade, para isto, o desenvolvimento de estudos e o emprego de mecanismos se tornaram essenciais para auxiliar este direito social, como por exemplo, sistemas de câmeras de vigilância comumente adotados para monitorar ambientes. O uso desta tecnologia em conjunto com Inteligência Artificial (IA) foi responsável por localizar 600 criminosos procurados pelo governo da Bahia [Secom 2023], chegando a atingir um grau de similaridade de 93% ao reconhecer um indivíduo. A fim de contribuir com o cenário de vigilância, o presente trabalho aborda um modelo inteligente baseado na arquitetura de microsserviços e comunicação assíncrona responsável por realizar a detecção e reconhecimento facial, alcançando, respectivamente, acurácias de 92,73% e 69,23% em imagens capturadas de um ambiente real monitorado à partir de microcontroladores com câmeras acopladas, com o intuito de realizar o monitoramento de pessoas consentidas em ambiente indoor.*

1. Introdução

Sistemas de vigilância com o objetivo de monitorar determinados locais e prevenir ataques e danos ganharam notoriedade após ataques terroristas presenciados na Europa e América do Norte, como por exemplo, os ataques realizados aos EUA no dia 11 de setembro de 2001. Conforme [CNN 2021], 2.753 pessoas morreram como consequência da colisão dos voos 11 e 175 da *American Airlines* e *United Airlines*, respectivamente, realizados de forma intencional às torres do *World Trade Center* (WTC) em Manhattan.

Tendo como o objetivo fornecer segurança e dispor de recursos facilitadores para o controle do ambiente no qual está inserido, equipamentos de vigilância associados aos

sistemas de monitoramento inteligentes são capazes de aplicar técnicas de processamento de imagens digitais, inteligência artificial e reconhecer e detectar padrões [Souza 2015].

No trabalho de [Farias and Lima 2018], por exemplo, foi realizado um estudo e análise de técnicas para reconhecimento e armazenagem online de faces mediante a imagens de vídeo feitas através da *Webcam*, bem como a realização de comparações com imagens contidas no banco de dados do sistema desenvolvido para identificar pessoas. Em [Moriggi Filho 2018], foi desenvolvido um sistema funcional que seja capaz de coletar dados de sensores possibilitando um controle das condições de temperatura, umidade e de acesso de ambientes por reconhecimento facial em tempo real. Já em [da Silva Nunes 2016], foi realizado um estudo de algoritmos responsáveis pelo reconhecimento automático de faces humanas em uma imagem por meio da aplicação de técnicas de visão computacional utilizando o *AdaBoost* (do inglês, “Adaptive Boosting”) para detectar faces e *Eigenfaces* para reconhecimento facial.

O uso de câmeras para a vigilância e monitoramento de ambientes necessitam ativamente da intervenção humana para identificação e/ou localização de pessoas. Diante disto, a aplicação de modelos inteligentes neste cenário contribui para automação do processo de monitoramento e reconhecimento.

Para tanto, estudos e técnicas vêm sendo realizadas e podem ser empregadas no desenvolvimento de soluções tecnológicas neste contexto. Permitindo, assim, o aumento da taxa de eficiência ao se realizar o processo de monitoramento associado a identificação, reconhecimento e rastreamento de indivíduos mediante o complexo de câmeras que integram um ambiente.

2. Revisão de Literatura

Na presente seção serão apresentados conceitos e métodos para embasamento e desenvolvimento do trabalho, englobando uma visão sobre legislação, processamento de imagens digitais, visão computacional e detecção e reconhecimento de faces.

2.1. Legislação e Sistemas de Reconhecimento Facial

A Lei Geral de Proteção de Dados Pessoais (LGPD) determina o consentimento da coleta de dados sensíveis com base na privacidade, incluindo a imagem de pessoas. Dito isto, para o caso do uso de recursos biométricos é necessário a autorização prévia. Contudo, há casos em que consentimento não se faz necessário, como por exemplo, situações em que o reconhecimento facial é realizado visando unicamente a segurança pública, defesa da pátria e atividades de combate à criminalidade, para estes casos específicos, a LGPD não é aplicável [Comitê Central 2020].

2.2. Processamento de Imagens Digitais

Com a finalidade de otimizar as taxas de sucesso dos modelos de detecção e reconhecimento facial, a aplicação do processamento de imagens digitais demonstrou ser essencial. Segundo [Gonzalez and Woods 2000], o uso de métodos de processamento de imagens digitais decorre de duas áreas principais: melhoria da informação visual para a análise humana e o processamento de dados de cenas para percepção automática através do uso de máquinas.

O processamento de imagens pode ser dividido em três níveis. O primeiro, baixo nível, corresponde ao processamento inicial de formação da imagem, bem como correção de distorções geométricas e remoção de ruído. O segundo, médio nível, visa aprimorar a qualidade da imagem e viabilizar melhorias para identificar objetos presentes na mesma. E o último, alto nível, diz respeito ao reconhecimento e a interpretação da imagem, a fim de gerar uma atribuição de classes aos objetos presentes na imagem. Apesar destas subdivisões não possuem fronteiras definitivas, elas possuem um esquema de trabalho útil para a categorização dos vários processos de um sistema que faz a análise de imagens [Gonzalez and Woods 2000].

2.3. Visão Computacional

Segundo [Ballard and Brown 1982], a visão computacional é a ciência que estuda e desenvolve tecnologias que possibilitam às máquinas enxergarem e extraírem características do meio, através de imagens capturadas por diferentes sensores e dispositivos. Tais informações extraídas permitem que sejam realizados o reconhecimento, manipulação e processamento de dados sobre os objetos que compõem a imagem capturada. Embora os trabalhos de [Ballard and Brown 1982] tenham sido realizados por volta da década de 80, a visão computacional já vinha sendo estudada desde a década de 50.

O reconhecimento de padrões e processamento de imagens são campos que são relacionados a visão computacional. Como mencionado na seção anterior, o processamento de imagem diz respeito às técnicas para manipulação de informações que estão dispostas em forma de imagem, como por exemplo, remoção de ruídos e realces das bordas. Já o reconhecimento de padrões fornece a capacidade de identificar e classificar objetos representados [Barelli 2018].

2.4. Detecção e Reconhecimento de Faces

Uma das tarefas essenciais para o desenvolvimento deste trabalho é a detecção da presença de uma face em uma imagem. Para [Zhao et al. 2003], detectar faces é a capacidade do algoritmo localizar e pôr em evidência, de alguma maneira, um rosto de um indivíduo em uma imagem, o qual pode ser no formato de vídeo ou foto, tendo como base em seus elementos integrantes: olhos, boca e nariz ou a combinação destes, distinguindo uma face humana a partir de outros tipos de padrões.

Conforme [Santana and Rocha 2015], as técnicas mais utilizadas na realização da detecção de faces são: casamento de padrões que se baseiam na detecção mediante a comparações com formas geométricas, modelos estatísticos, baseado em redes neurais, baseados em tons de pele e o proposto no trabalho de [Viola and Jones 2001].

O algoritmo Viola-Jones desenvolvido por [Viola and Jones 2001], posteriormente estendido por [Lienhart and Maydt 2002], é um mecanismo que têm sido comumente aplicado quando se pretende realizar o reconhecimento de objetos e/ou faces. Este algoritmo é capaz de detectar faces com precisão, alta taxa de acerto, baixa taxa de falsos positivos e com baixo custo computacional.

Conforme [Araujo 2010], o algoritmo é composto de três partes. A primeira é a representação da imagem em um espaço de características baseadas em filtros de *Haar*. Tal técnica é utilizada na detecção de padrões de forma a encontrar elementos da imagem que possuam diferença de contraste, esses quadros (contrastes) são qualificados como

características de *Haar*. A segunda característica é a elaboração de um classificador baseado em *Boosting*, capaz de selecionar as características mais importantes. A terceira, presente em versões mais recentes do algoritmo, trata-se de um conjunto estendido de características proposto por [Lienhart and Maydt 2002], responsável por considerar um novo conjunto de características e versões rotacionadas das características utilizadas na abordagem original.

Outra parte essencial considerada aqui é o reconhecimento facial, de acordo com estudos constantes em [Alves 2021] e [Geitgey 2016], um processo contínuo responsável por resolver cada etapa do reconhecimento facial separadamente é composto por quatro passos: obter todas as faces presente na imagem; posicionar e centralizar as faces; gerar as medidas da face; e comparar as medidas da face junto a base de pessoas conhecidas a fim de determinar a nova face.

Uma vez realizada a detecção facial é possível executar o processo de reconhecimento das áreas que contém rostos na imagem. Um método muito comum de realizar a detecção de faces é chamado de Histograma de Gradientes Orientados (do inglês *Histogram of Oriented Gradients*) [Dalal and Triggs 2005]. De acordo com [Geitgey 2016], o método HOG transforma a imagem a ser analisada para a escala de preto e branco, e então analisa cada *pixel* da imagem e os pixels que estão em torno dele, e desenha uma flecha (gradiente) que aponta para direção onde a imagem fica mais escura.

Uma vez obtida a posição das faces na imagem isolada, a segunda etapa consiste em centralizar os olhos e boca. Esse processo é realizado para simplificar o reconhecimento para o caso de faces de uma mesma pessoa que estejam em ângulos diferentes. Nesta etapa, é utilizado um algoritmo denominado *face landmark estimation*, ou, estimativa de marcas do rosto. O algoritmo se baseia na premissa de que existem 68 pontos específicos presentes em todas as faces, então, um algoritmo de *Machine Learning* (ML) é treinado para conseguir encontrar todos esses pontos em uma face [Geitgey 2016].

A terceira etapa do processo é codificar as faces encontradas. Para isso, é necessário extrair uma série de medidas de cada face. Porém, as medidas que podem fazer sentido para um humano, como cor dos olhos ou tamanho do nariz, podem não fazer sentido para um computador, então, pesquisadores descobriram que uma abordagem muito precisa é utilizar *deep learning* e deixar que a máquina decida quais medidas são mais importantes e as colete. Uma rede neural é então treinada para gerar 128 medidas para cada face [Geitgey 2016].

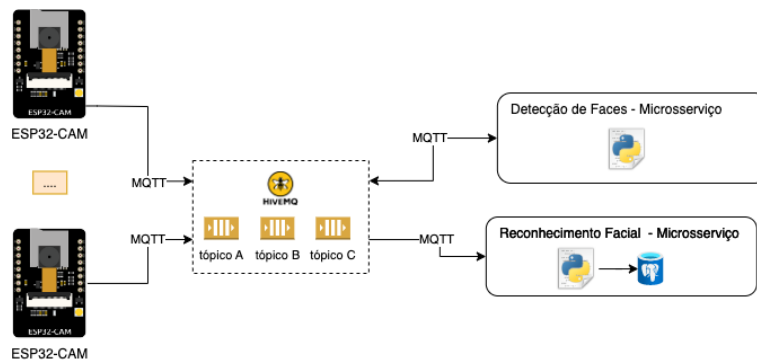
Por fim, segundo [Navlani 2018], para realizar a busca de reconhecimento da face podem ser utilizadas diversas técnicas. Uma técnica de busca de reconhecimento facial bastante difundida e que foi utilizada no desenvolvimento deste trabalho por meio da extensão cube do PostgreSQL é a K-Nearest Neighbors (KNN). No KNN a estrutura do modelo é definida pelo conjunto de dados, no caso, a estrutura do modelo é representada pelas 128 medidas geradas da face. Considera-se K como a quantidade de vizinhos próximos, quando K é igual a 1, o algoritmo só levará em consideração o vizinho mais próximo, obtendo então a face com maior similaridade dentre as existentes na base de dados.

3. Descrição do Modelo

Nesta seção serão apresentados o diagrama geral (Figura 1) do modelo proposto, bem como os seus componentes que integram o presente trabalho, sendo eles: solução embarcada em dispositivos ESP32-CAM responsável pela captura e envio das imagens capturadas de um determinado ambiente *indoor*; microserviço para detecção de faces; e microserviço responsável pelo reconhecimento facial.

A fim de tornar uma solução escalável e resiliente a falhas, foi adotado um *broker* MQTT (do inglês - *Message Queue Telemetry Transport*) de mensagens para estabelecer a comunicação entre os componentes desenvolvidos para o modelo. Sendo assim, é possível realizar o envio de mensagem para um agente intermediário responsável por gerenciar e trafegar as informações essenciais para o processo de detecção e reconhecimento facial, bem como o rastreamento de pessoas consentidas.

Figura 1. Arquitetura do Modelo Proposto.



3.1. Captura e envio de imagens

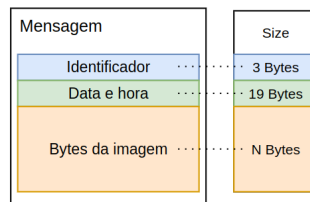
A captura e envio de imagens do ambiente *indoor* são responsabilidades dos microcontroladores ESP32-CAM, sendo assim, o dispositivo deve ser configurado para que consiga registrar fotos do ambiente e as enviar para um tópico no *broker* MQTT no formato de mensagens contendo o identificador do dispositivo, imagem e a data e hora do registro, viabilizando o rastreamento futuro de uma pessoa monitorada. É importante notar que, uma vez que a mensagem esteja publicada no tópico, o microserviço responsável por identificar faces irá realizar a leitura destas mensagens e efetuará as operações necessárias para detecção de faces.

Na Figura 2, pode ser observada a estrutura da mensagem enviada do ESP32-CAM para o tópico, composta aqui por três campos, sendo o primeiro o identificador do dispositivo, possuindo 3 *bytes*, o segundo a data e hora, com 19 *bytes* e, por último, ocupando o restante da mensagem, os *bytes* da própria imagem com resolução de 1280x720 (definida por meio de experimentos e performance do dispositivos).

3.2. Detecção de Faces

Para realizar a detecção de faces foi desenvolvido um microserviço em *Python* nomeado por *ms_identification*, nele estão contidas todas as operações e procedimentos necessários para que seja possível realizar a extração e localização de uma face em uma imagem digital.

Figura 2. Estrutura da mensagem enviada do ESP32-CAM para o *broker* MQTT.



Uma vez que a solução realiza uma leitura de uma nova mensagem no tópico *identificaton* (enviadas pelo ESP32-CAM), é realizado o procedimento para converter os *bytes* dos dados em sua forma original, bem como obter a imagem capturada para que sejam feitas as operações necessárias para a detecção de faces.

Após a obtenção dos *bytes* da imagem, é possível realizar a detecção de faces por meio da aplicação do classificador em cascata (*haarcascade_frontalface_default.xml* presente no openCV) sob a imagem por meio da função *detectMultiScale* do *OpenCV*. Este processo permite que seja retornado uma estrutura de dados possuindo 4 valores para cada face encontrada, sendo o suficiente para formar os pontos de extremidades do recorte da face, permitindo descobrir a sua localização na imagem e por fim obter um recorte da face com base na imagem original.

Uma vez que todas as faces são detectadas as mesmas serão agrupadas com o identificador e data e hora para compor uma nova mensagem a ser publicada no tópico *recognition*, responsável por conter faces a serem reconhecidas.

4. Reconhecimento Facial

Os mecanismos desenvolvidos para realizar o reconhecimento facial estão contidos no microserviço nomeado por *ms_recognition*, o qual utiliza do *broker* MQTT para receber mensagens contendo faces detectadas e que foram publicadas no tópico *recognition*, para que então seja feito o processo de reconhecimento facial. Diante disto, é realizado o processo de leitura e conversão da mensagem para o seu formato original. Obtendo, assim, os valores do identificador do dispositivo, data e hora de captura e as imagens de faces detectadas.

Para cada imagem de face detectada, é utilizado um método que confirma a existência de uma face na foto. Caso não seja detectada uma face na imagem, o processo de reconhecimento para a mesma é descartado. Caso contrário, é realizado o processo de reconhecimento facial, o qual compreende na obtenção das medidas faciais utilizando uma rede neural pré-treinada e a busca por faces de indivíduos consentidos que são similares em banco de dados de medidas faciais, tendo como base um *threshold* de 0,6. Uma vez que esta face seja considerada uma pessoa conhecida, um novo registro na base de dados contendo o identificador do dispositivos, data e hora de captura e referência para a imagem da face em meio físico são persistidos. Tais informações serão responsáveis por compor o trajeto de um indivíduo em um ambiente indoor.

Vale ressaltar que o processo de reconhecimento facial conta com o apoio de um banco de dados para armazenar e buscar características das faces capturadas, para então determinar se o indivíduo é ou não reconhecido.

5. Resultados

5.1. Experimentos Individuais

Para avaliar a solução embarcada presente no ESP32-CAM, foram realizadas baterias de experimentos com duração de 2 minutos em diferentes resoluções, a fim de mensurar o desempenho do dispositivo ao realizar a captura de imagens em relação à qualidade das mesmas, também, foi analisado o tempo de envio das imagens para o tópic no *broker* MQTT, visando avaliar o microcontrolador em termos de valores mínimos e máximos de *Frames Per Second* (FPS) ao enviar uma imagem (com e sem defeito) para o *broker*.

Dentre as 9 resoluções avaliadas, a selecionada foi 1280x720, sendo a mais alta que obteve melhores resultados, totalizando 64 de imagens, 0 imagens defeituosas enviadas e com FPS máximo e mínimo de 1 publicadas no *broker* MQTT. Foi possível obter resoluções maiores, como 1600x1200, contudo a suas taxas de captura e envios não foram satisfatórias para o modelo, totalizando 38 imagens enviadas e com 0 e 1 FPS mínimos e máximos, respectivamente.

Com a finalidade de mensurar a capacidade do modelo em detectar faces, por meio do ESP32-CAM foram capturadas 55 imagens de um indivíduo consentido, o qual manteve uma distância que variou entre 0.8 e 4.2 metros em relação ao microcontrolador. Estas que por sua vez foram avaliadas em etapas para a definição dos limiares utilizados no openCV. Ao final, foram adotados os limiares *minNeighbors* igual 4 para especificar a quantidade de vizinhos necessários para classificar uma região como face e *maxSize* de 300x300 para delimitar o tamanho máximo em *pixels* que uma face deve ter.

Ao final, como pode ser visto na Tabela 1 foram quantificadas 50 Verdadeiros Positivos, onde a face foi detectada corretamente ao menos uma vez, 0 Falsos Positivos, indicando que não houveram imagens em que o modelo detectou somente faces incorretas. 4 Falsos Negativos, devido às imagens em que não foi possível detectar ao menos uma face e 1 Verdadeiro Negativo, referente a imagem em que não havia a presença da pessoa em cena e não foi realizada nenhuma detecção.

O reconhecimento facial no modelo é realizado por meio do cálculo da distância euclidiana entre as medidas das faces, ou seja, para uma face ser reconhecida como uma pessoa monitorada é necessário que o resultado do cálculo esteja abaixo do limiar de dissimilaridade (0.6) adotado no modelo. Para mensurar o reconhecimento facial, foram fornecidos como dados de entrada para o microsserviço um total de 52 faces únicas detectadas pelo microsserviço de detecção facial, já que duas imagens tiveram como resultado duas detecções faciais. Dito isso, assim como pode ser visto na Tabela 2, foi obtido um total de 36 reconhecimentos do indivíduo monitorado, 16 faces não reconhecidas devido a baixa qualidade do recorte facial, uma vez que estas faces tiveram como origem as imagens mais distantes (entre 2.5 e 4.2 metros) e, por fim, 0 reconhecidas incorretamente.

5.2. Experimento Geral

O objeto de estudo baseia-se na adição de uma pessoa consentida em ser monitorada em um ambiente *indoor* no modelo contendo uma imagem com a face aparente (Figura 3), para que sejam computadas as medidas faciais e armazenadas no banco de dados, a fim de obter um padrão para realizar o reconhecimento da face da pessoa monitorada. Ao final deste processo, um novo registro é gerado e o indivíduo se encontra monitorado pelo modelo.

Tabela 1. Matriz de confusão obtida para detecção facial.

	Positivo	Falso
Positivo	50 (VP)	4 (FN)
Falso	0 (FP)	1 (VN)
Total	50	5

Tabela 2. Matriz de confusão obtida para o reconhecimento facial.

	Positivo	Falso
Positivo	36 (VP)	16 (FN)
Falso	0 (FP)	0 (VN)
Total	36	16

Figura 3. Imagem com a face aparente da pessoa monitorada fornecida ao modelo.

Como os experimentos realizados ocorreram em um apartamento, foram criados 3 registros para armazenar os dados dos microcontroladores responsáveis por monitorar um cômodo específico do imóvel, onde os quais realizaram a captura de imagens de uma região e o envio destas por meio de mensagens para o *broker* MQTT.

De forma a tornar o cenário mais próximo de um ambiente real de monitoramento, em que são raros os momentos em que um indivíduo está olhando de forma incisiva para a câmera, durante o experimento geral a pessoa monitorada evitou o contato visual com o microcontrolador, sendo direcionado o olhar para regiões periféricas ou ao redor onde o dispositivo estava, mas não diretamente, removendo assim qualquer viés que pudesse alterar o resultado final do experimento.

O trajeto percorrido pela pessoa monitorada pode ser visualizado na Figura 4, juntamente com as faces detectadas e reconhecidas em cada área vigiada, onde o indivíduo teve o seu primeiro reconhecimento na “Porta de Entrada” e o último na “Cozinha”, ao longo deste período, o mesmo foi reconhecido 29 vezes no ambiente *indoor* enquanto circulava entre os cômodos.

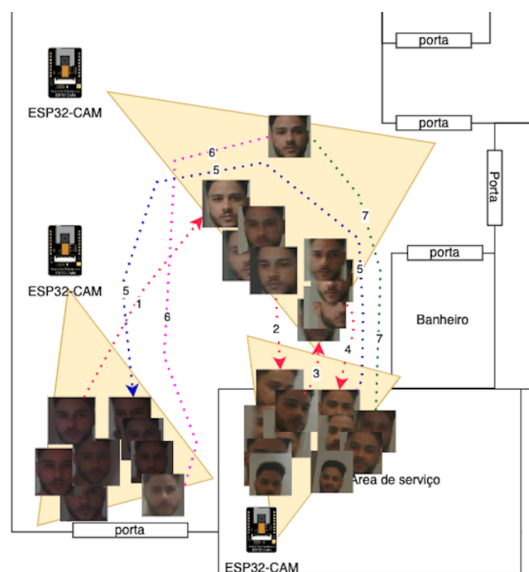
Ao longo de um período de 3 minutos, os dispositivos realizaram capturas de imagens do ambiente com ou sem a presença da pessoa monitorada, onde em sua ausência ou sem o rosto aparente não houve detecção facial. No total, houveram 397 imagens capturadas, enviadas, recebidas e processadas pelo modelo.

6. Considerações Finais

Com a finalidade de que seja realizada a continuação dos estudos, tendo como base o vigente trabalho, há pontos em que é viável o investimento de esforço para o refinamento de técnicas para melhorar os resultados obtidos no modelo.

Dentre eles, está a possibilidade de aplicar o conceito de *threads* no código embarcado no ESP32-CAM, este que por sua vez, permite que sejam criadas um total de 2 *threads*, a fim de dividir responsabilidades de captura de imagens e envio de mensagens.

Figura 4. Rastreamento do monitorado pelo modelo



Para a detecção e reconhecimento facial, vale o estudo e análise mais aprofundado, dos limiares para melhorar a capacidade de detectar e reconhecer padrões, ou realizar experimentos envolvendo outras técnicas, para demonstrar o comparativo e ganhos em relação às técnicas aplicadas no modelo.

Referências

- Alves, T. C. (2021). Proposta de uma sistemática do projeto para manufatura aditiva (dfma) no contexto de desenvolvimento de produtos. Master's thesis, Faculdade Tecnológica. Universidade de Brasília.
- Araujo, G. M. (2010). Algoritmo para reconhecimento de características faciais baseado em filtros de correlação. Master's thesis, Programa de Pós-graduação em Engenharia Elétrica (PPEE) - UFRJ.
- Ballard and Brown (1982). *Computer Vision*.
- Barelli, F. (2018). *Introdução à Visão Computacional: Uma abordagem prática com Python e OpenCV*. Editora Casa do Código.
- CNN (2021). September 11 terror attacks fast facts. <https://edition.cnn.com/2013/07/27/us/september-11-anniversary-fast-facts/index.html/>. acessado em 12 Out. 2021.
- Comitê Central, d. G. d. D. (2020). Guia de boas práticas lei geral de proteção de dados (lgpd). Disponível em: https://www.gov.br/governodigital/pt-br/seguranca-e-protecao-de-dados/guias/guia_lgpd.pdf. acessado em 22 Mar. 2023.
- da Silva Nunes, G. M. (2016). *VISAO COMPUTACIONAL APLICADA A DETECCAO E AO RECONHECIMENTO FACIAL*. PhD thesis, Engenharia Eletrônica e de Computação (Trabalho de Conclusão de Curso). Universidade Federal do Rio de Janeiro.

- Dalal, N. and Triggs, B. (2005). Histograms of oriented gradients for human detection. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, volume 1, pages 886–893. Ieee.
- Farias, L. C. and Lima, M. A. C. (2018). Estudo e análise de técnicas para reconhecimento e armazenagem online de faces.
- Geitgey, A. (2016). Machine learning is fun! part 4: modern face recognition with deep learning. *Medium. Medium Corporation*, 24.
- Gonzalez, R. C. and Woods, R. E. (2000). *Processamento de imagens digitais*. Editora Blucher.
- Lienhart, R. and Maydt, J. (2002). An extended set of haar-like features for rapid object detection. In *Proceedings. international conference on image processing*, volume 1, pages I–I. IEEE.
- Moriggi Filho, E. (2018). Controle de acesso por reconhecimento facial e controle de ambientes.
- Navlani, A. (2018). Knn classification using scikit-learn. Disponível em: <https://www.datacamp.com/community/tutorials/k-nearest-neighbor-classification-scikit-learn>. Acessado em: 16 de Jul. de 2021.
- Santana, L. M. Q. d. and Rocha, F. G. (2015). Processo de detecção facial utilizando viola;jones. *Interfaces Científicas - Exatas e Tecnológicas*, 1(1):35–40.
- Secom (2023). Reconhecimento facial chega a marca de 600 foragidos localizados e presos. Disponível em: www.bahia.ba.gov.br/2023/01/noticias/seguranca/reconhecimento-facial-chega-a-marca-de-600-foragidos-localizados-e-presos/. acessado em 22 Mar. 2023.
- Souza, T. T. L. d. (2015). Auto-calibração de câmeras de vídeo-vigilância por meio de informações da cena. Master's thesis, Instituto de Matemática. Departamento de Mecatrônica.
- Viola, P. and Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. In *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001*, volume 1, pages I–I. Ieee.
- Zhao, W., Chellappa, R., Phillips, P. J., and Rosenfeld, A. (2003). Face recognition: A literature survey. *ACM computing surveys (CSUR)*, 35(4):399–458.