

Comparison of Face Detection Methods Under the Influence of Lighting Variation

Renan Sevilha Sakamoto¹, Benjamin Grando Moreira¹, Pablo Andretta Jaskowiak¹

¹ Centro Tecnológico de Joinville (CTJ)
Universidade Federal de Santa Catarina (UFSC)

renan.sakamoto@grad.ufsc.br {benjamin.grando,pablo.andretta}@ufsc.br

Abstract. This article compares facial detection methods under varying lighting conditions, highlighting the impact of illumination on the accuracy of facial recognition algorithms. Different algorithms were analyzed, including Haar-Cascade-based methods, Artificial Neural Network-based methods, and also Histogram of Oriented Gradients. The results indicate that while some methods perform well under varied lighting conditions, others show significant accuracy drops in low-light environments. The research contributes to understanding the limitations and capabilities of facial detection methods under different lighting conditions, making it relevant for the development of more robust systems.

Resumo. Este artigo compara métodos de detecção facial sob variações de iluminação, destacando o impacto da iluminação na precisão dos algoritmos de reconhecimento facial. Foram analisados diferentes algoritmos, incluindo métodos baseados em Haar-Cascaded, baseados em Redes Neurais Artificiais e também Histograma de Gradientes Orientados. Os resultados indicam que, embora alguns métodos apresentem bom desempenho em condições de iluminação variadas, outros mostram quedas significativas de precisão em ambientes com pouca luz. A pesquisa contribui para o entendimento das limitações e capacidades dos métodos de detecção facial em diferentes condições de iluminação, sendo relevante para o desenvolvimento de sistemas mais robustos.

1. Introdução

A identificação biométrica é uma ferramenta para o reconhecimento de pessoas com base em suas características físicas e/ou comportamentais únicas e distintivas. Neste contexto, as características podem incluir impressões digitais, padrões faciais ou de retina, geometria da íris, geometria da mão, voz, termograma facial e assinatura manuscrita [Jain et al. 2000]. O uso de tecnologias biométricas permite verificar ou identificar pessoas por meio destas características, oferecendo elevada precisão e segurança [Pankanti et al. 2000].

Dentre as possíveis abordagens biométricas destaca-se o reconhecimento facial, que utiliza a visão computacional a fim de identificar e validar a identidade de pessoas com base nos traços e características faciais. Uma das grandes vantagens destes sistemas reside na ausência de contato físico, reduzindo, por exemplo a possibilidade de propagação de doenças [Rana et al. 2022]. Para alguns métodos desta abordagem, uma etapa essencial comprehende a detecção e localização de faces, que pode ser afetada por diversos fatores, como, mudança na pose do indivíduo; alteração do ângulo em que é

capturado pela câmera; variação de idade; oclusão, principalmente da porção superior da face e; iluminação, para a qual a reflexão pode dificultar o processamento [Abate 2007].

A variação de iluminação observada em imagens possui impacto direto na precisão e robustez dos sistemas de detecção de faces, e por consequência, no reconhecimento facial. A iluminação é uma variável complexa e multifacetada, suscetível a fatores como diferentes fontes de luz (tanto naturais, quanto artificiais), condições climáticas adversas e ambientes físicos complexos [Gamage e Seneviratne 2014]. Tais variações podem distorcer a aparência de rostos, prejudicando a capacidade dos métodos de reconhecimento em identificar indivíduos com precisão. Portanto, compreender e mitigar os efeitos da variação na iluminação é essencial para o desenvolvimento de sistemas de reconhecimento facial confiáveis e eficientes [Gamage e Seneviratne 2014, Rana et al. 2022].

O presente trabalho explora a interação entre a variação da iluminação e o desempenho de diferentes métodos (pré-treinados) de detecção facial. Foram selecionados para avaliação métodos com diferentes abordagens e complexidades, prontamente disponíveis para uso em bibliotecas de software (em particular, na linguagem Python, utilizada no desenvolvimento deste trabalho). Em específico, foram comparados o algoritmo Haar-Cascaded [Viola e Jones 2001], utilizando diferentes versões, por meio da biblioteca OpenCV [Bradski 2000]; Histogram of Oriented Gradients (HOG) [Dalal e Triggs 2005], por meio da biblioteca Dlib [King 2009]; YOLO v8 (You Only Look Once) [Redmon et al. 2016]; Multi-task Cascaded Convolutional Networks [Zhang et al. 2016] e; o modelo de rede neural profunda disponível no OpenCV para detecção de faces (módulo *dnn* - *deep neural networks*), pré-treinada com a base de dados Caffe [Jia et al. 2014].

Para a avaliação foi empregada a base de dados expandida de faces de Yale [Georghiades et al. 2001], uma vez que é adequada aos objetivos deste estudo. A base de dados de faces de Yale oferece uma ampla gama de imagens faciais capturadas sob várias condições de iluminação e ângulos, além de possuir imagens somente em tons de cinza, com tamanhos padronizados. Tal diversidade de dados permite uma análise mais abrangente e precisa da influência da variação da iluminação no reconhecimento facial.

O restante do trabalho está organizado como segue. Na Seção 2 são discutidos trabalhos relacionados, considerando o contexto de detecção facial sob diferentes condições de iluminação. Na Seção 3 são discutidas as principais etapas da avaliação, passando pela escolha e descrição da base de dados, determinação das condições de iluminação, escolha dos métodos e métricas de avaliação empregadas. Na Seção 4 são apresentados resultados e discussões

2. Trabalhos relacionados

Investigações no campo do reconhecimento facial têm buscado compreender os desafios impostos pela variação da luminosidade e seus efeitos no desempenho dos sistemas de detecção de faces e reconhecimento facial. Trabalhos neste contexto não apenas aprofundaram a compreensão das nuances da interação entre iluminação e características faciais, mas também apresentam abordagens inovadoras para mitigar seus efeitos negativos.

[Lee et al. 2018] realizaram estudos considerando um dispositivo de captura de imagens de baixo custo (uma webcam Logitech C310) e diferentes condições de iluminação. Embora os autores estivessem preocupados com a influência da iluminação

no processo de detecção facial, a avaliação foi realizada de forma ad-hoc, considerando apenas um método da biblioteca OpenCV e poucas imagens, sem uma avaliação sistemática.

O trabalho de [Behling 2019] se concentra na identificação de diferentes emoções em faces a partir de vídeos. Neste contexto, o autor discute que muito embora Redes Neurais Profundas apresentem melhores resultados na detecção das faces, as imagens devem ser submetidas à um pré-processamento a fim de amenizar os efeitos observados em diferentes condições de iluminação, visto que podem afetar o desempenho das técnicas.

[Böhm 2021] concentrou-se em explorar a aplicação de redes neurais convolucionais (CNNs) para aprender representações robustas em sistemas embarcados. Abordagens como a de Böhm não apenas oferecem pontos de vista sobre como as redes neurais podem ser treinadas para discernir características consistentes apesar das mudanças de iluminação, mas também sugerem soluções promissoras para o desafio contínuo de melhorar a confiabilidade dos sistemas de reconhecimento facial.

De acordo com [Kaur e Sharma 2023], um estudo comparativo entre abordagens de detecção de rosto utilizando Haar-Cascaded, Histogram of Oriented Gradients (HOG) com SVM e Multi-Task Cascaded Neural Network, constatou-se que o método Haar, apesar de simples, é robusto e eficiente para detecção de rostos posicionados de maneira frontal com 95% de precisão. Este mesmo método, entretanto, obteve piores resultados considerando rostos de perfil, com óculos e/ou oclusos. Já o método HOG atingiu 98.6% de precisão para rostos posicionados de forma frontal, apresentando ainda uma boa taxa de detecção para rostos com óculos. Por fim, a abordagem baseada em MTCNN obteve o melhor resultado nesse estudo, com 99.7% de precisão, sendo capaz de identificar rostos de forma satisfatória em todos os casos. Os resultados obtidos sugerem que a abordagem baseada em MTCNN é menos propensa a erros mesmo com condições adversas.

Ao examinar os trabalhos relacionados sobre a influência da variação da iluminação no reconhecimento facial, fica claro que essa é uma área de pesquisa em evolução constante. De fato, este tema tem sido abordado em revisões específicas da literatura, apresentadas em artigos de periódicos especializados. Neste sentido, sugerimos ao leitor interessado no tema os artigos de [Zafeiriou et al. 2015] e [Kumar et al. 2019] que apresentam revisões mais amplas no contexto de detecção facial (abordando aspectos de iluminação) e o trabalho de [Ochoa-Villegas et al. 2015] que apresenta uma revisão específica acerca de detecção facial sob diferentes condições de iluminação e seus desafios.

3. Materiais e Métodos

A metodologia deste trabalho começa com a definição, separação e preparação da base de dados, seguida pela escolha dos métodos a serem empregados. A Figura 1 ilustra a sequência de passos seguida neste estudo, abrangendo desde a escolha da base de dados até as detecções dos resultados obtidos através dos diferentes métodos de detecção facial. A análise da variação da iluminação foi conduzida avaliando os diferentes modelos e métodos com o conjunto de dados escolhido, com base nas suas detecções. Posteriormente, os resultados obtidos foram avaliados com métricas de qualidade apropriadas.

É importante realizar uma discussão em relação a base de dados para este trabalho. Inicialmente, foram consideradas para utilização as bases de dados: *Labeled Faces in*

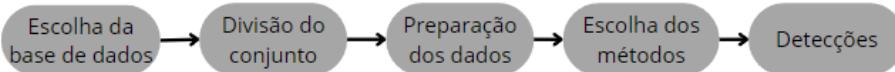


Figura 1. Principais etapas adotadas durante a realização do estudo.

*the Wild*¹, IMDB-Wiki², *MUCT face database*³, *CMU Multi-PIE Database*⁴. As duas primeiras bases de dados supracitadas foram descartadas em virtude da pouca diferença na iluminação apresentada em suas imagens. A base de dados *MUCT face database*, por sua vez, possui variação na iluminação e ângulo das imagens, porém com um número relativamente baixo de imagens. A base de dados *CMU Multi-PIE Database* possui uma base de dados extensa, mas com poucas condições de variação de iluminação.

Em virtude dos pontos anteriormente discutidos, a base de dados estabelecida para o estudo foi a base de dados expandida de faces de Yale [Georghiades et al. 2001]. A base de dados de faces de Yale oferece uma ampla gama de imagens faciais capturadas sob várias condições de iluminação e ângulos, além de seu tamanho ser padronizado e possuir apenas tons de cinza. Tal diversidade de dados permite uma análise mais abrangente e precisa da influência da variação da iluminação no reconhecimento facial.

A base de dados expandida B de Yale (utilizada no estudo) contém 16.380 imagens de 28 indivíduos em 9 poses diferentes e sob 65 condições de iluminação distintas. Essas condições de iluminação foram divididas em cinco subgrupos, conforme o ângulo entre o eixo da câmera e a direção da fonte de luz. Referente a variação da iluminação, foram estabelecidas 5 categorias diferentes, com relação ao ângulo da luz, separadas em: Muito Claro, com ângulos de 0° a 12°, Claro, de 12° a 25°, Médio, com 26° a 50°, Escuro, com ângulos entre 51° e 70° e Muito Escuro com 71° a 130°. A Figura 2 apresenta todos os 28 indivíduos compreendidos na base de dados de Yale, que foi utilizada neste trabalho. As faces, tal qual apresentadas na figura, fazem parte do primeiro grupo de iluminação, com o ângulo entre o eixo da câmera e a direção da fonte da luz entre 0° e 12° graus.

É importante notar que a base de dados não contempla efeitos adicionais, como oclusão, isto é, não há obstrução da visão da face. Além disso, as faces são sempre apresentadas sem a presença acessórios como gorros, bonés, máscaras ou óculos. Tendo em vista que o objetivo deste trabalho é avaliar somente aspectos pertinentes à influência da iluminação na detecção facial, acreditamos que estas características tornam a base de dados apropriada, isolando ou removendo possíveis fatores confundidores.

Como o objetivo é avaliar apenas a questão da iluminação a divisão do conjunto de dados foi feita com base nos ângulos de iluminação, resultando em 6 grupos de imagens conforme descrito acima. O último grupo, referido como ambiente, é utilizado exclusivamente para a classificação de verdadeiros negativos, onde os algoritmos não devem

¹Disponível em: <http://vis-www.cs.umass.edu/lfw/>

²Disponível em: <https://data.vision.ee.ethz.ch/cvl/rrothe/imdb-wiki/>

³Disponível em: <http://www.milbo.org/muct/The-MUCT-Landmarked-Face-Database.pdf>

⁴Disponível em: <http://www.cs.cmu.edu/afs/cs/project/PIE/MultiPie/Multi-Pie/Home.html>



Figura 2. Exemplo de faces contidas na base de dados adotada no estudo. Imagens obtidas do trabalho de [Georghiades et al. 2001].

detectar faces, tendo em vista que não há iluminação incidente na face. A Figura 3 ilustra a separação dos 6 grupos diferentes de iluminação para este trabalho. A partir das imagens é possível ter uma noção das diferentes condições de iluminação consideradas.

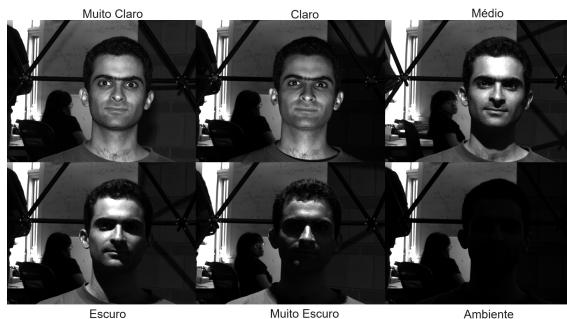


Figura 3. Exemplos separação dos dados. Imagens obtidas do trabalho de [Georghiades et al. 2001].

A preparação dos dados foi realizada com um *script* em Python para criação das *bounding boxes* para marcação dos rostos, sendo essa a localização real da face na imagem. A implementação desse *script* foi necessária em virtude da ausência de marcação das faces na base de dados original. O *script* permitiu gerar as coordenadas da caixa com dois cliques do mouse, o primeiro para as coordenadas do canto superior esquerdo e o segundo para o canto inferior direito. Isto é, a determinação das coordenadas que definem os *bounding boxes* foi feita manualmente. Ao realizar esse procedimento com todas as imagens, o código do *script* salva a nome da imagem juntamente com as duas coordenadas em um documento de texto. Além disso, foi implementado um outro *script* em Python para apresentar as caixas delimitadoras criadas e verificar se estão corretamente posicionadas, a Figura 4 apresenta o rosto de um indivíduo com a *bounding box* criada com o auxílio do código mencionado (delimitada em verde claro).

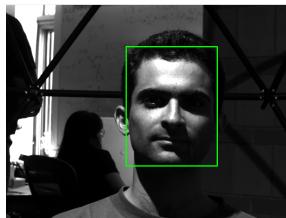


Figura 4. Exemplo da caixa delimitadora criada nas imagens. Imagem obtida do trabalho de [Georghiades et al. 2001].

Esse processo permitiu a posterior verificação e validação dos resultados obtidos pelos métodos de detecção, permitindo a obtenção de matrizes de confusão e cálculos de precisão, acurácia, por exemplo (maiores detalhes serão fornecidos a seguir).

3.1. Escolha dos algoritmos

Para o desenvolvimento deste trabalho foi utilizada a biblioteca OpenCV (Open Source Computer Vision Library). No contexto da detecção de faces, o OpenCV oferece módulos específicos, como o cv2.CascadeClassifier, utilizado neste trabalho. Este módulo usa o algoritmo de Viola-Jones, que aplica o modelo Haar-Cascaded, para detectar faces em imagens estáticas. O OpenCV oferece diferentes versões desses modelos para detecção facial, incluindo a padrão (default) e versões específicas para faces frontais, como Frontal Face Alt Tree e Frontal Face Alt 2, todos disponíveis no próprio OpenCV. De forma breve:

- Padrão (default) — HC-D: adequado para a detecção de faces em diferentes orientações e condições de iluminação, oferecendo um bom equilíbrio entre precisão e desempenho computacional;
- Frontal Face Alt Tree — HC-FFAT: projetado para detecção frontal de faces, concentrando-se principalmente na detecção de faces em perfil frontal e podendo ser mais rápido que a versão padrão em condições ideais de iluminação;
- Frontal Face Alt 2 — HC-FFA2: direcionado para melhorar a precisão em diferentes ângulos e rotações leves das faces. É útil em cenários onde as faces não estão perfeitamente alinhadas, proporcionando resultados mais confiáveis em detecções de faces.

Além das variações supracitadas, também foram utilizados:

- YOLO (*You Only Look Once*): O YOLO é capaz de realizar detecção de objetos em tempo real, além de conseguir detectar múltiplos objetos em uma única passagem pela imagem, possuindo equilíbrio entre velocidade e precisão na detecção de objetos, incluindo faces (Redmon *et. al.*, 2016). A versão do YOLO escolhida foi a mais recente, a versão 8 (YOLO v8) do ano de 2023. O modelo foi pré-treinado com a base de dados COCO [Lin *et al.* 2015];
- DNN (Deep Neural Network): As redes neurais profundas, DNNs, são usadas em tarefas de detecção de objetos, incluindo faces, e são capazes de aprender representações complexas de dados. Geralmente alcançam desempenho superior em comparação com métodos tradicionais, quando treinadas em conjuntos de dados substanciais, além da compatibilidade com o OpenCV, que oferece suporte à integração de modelos de DNN treinados para detecção de objetos, o que permite a utilização de redes pré-treinadas para a detecção de faces [LeCun *et al.* 2015].

- HOG (Histogram of Oriented Gradients): O HOG é especialmente útil em cenários onde a precisão é fundamental, na detecção de objetos, como faces. A biblioteca dlib utiliza o HOG para oferecer funcionalidades adicionais, como rastreamento de faces e identificação facial. Essas características fazem da dlib uma escolha adequada em aplicações de análise de vídeo e rastreamento de rostos. A dlib é uma biblioteca de código aberto com uma licença amigável que permite seu uso em uma variedade de projetos, devido a sua simplicidade, suporte a linguagem Python e C++, além de sua compatibilidade com aprendizado de máquina e visão computacional [King 2009];
- MTCNN (Multi-task Cascaded Convolutional Networks): O MTCNN é um modelo projetado especificamente para a detecção de faces, é preciso na localização de faces em imagens, pode detectar várias faces em uma única imagem e utiliza uma hierarquia de redes para a detecção de faces, incluindo a detecção de pontos-chave do rosto [Zhang et al. 2021]. Nesse algoritmo, como o próprio nome já implica, usa redes neurais convolucionais multitarefa em cascata e seu modelo foi pré-treinado com a base de dados CelebA.

A Figura 5 mostra, no campo superior esquerdo, a imagem de um indivíduo juntamente com a *bounding box* definida manualmente por meio do script em Python, enquanto as demais se referem às detecções realizadas pelos diferentes métodos.

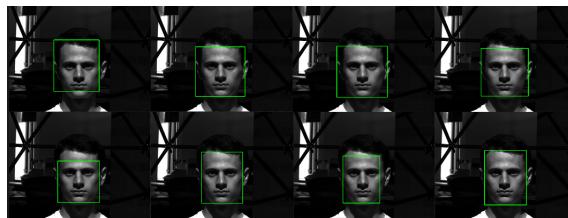


Figura 5. Demonstração das detecções feitas por diferentes métodos. Imagens obtidas do trabalho de [Georghiades et al. 2001].

3.2. Avaliação

A avaliação de algoritmos desempenha um papel crucial no desenvolvimento e na otimização de modelos. Dentre os diversos métodos de avaliação disponíveis, o método de avaliação base empregado neste trabalho é a matriz de confusão, que se destaca como uma ferramenta fundamental, fornecendo uma visão abrangente do desempenho do modelo ao comparar suas previsões com os resultados esperados.

Para verificar se os modelos avaliados detectaram as faces corretamente e as posicionaram de maneira precisa, empregou-se uma matriz que relacionou as *bounding boxes* (caixas delimitadoras) geradas pelos métodos com as ideais (definidas manualmente). Essas caixas são retângulos definidos por quatro coordenadas que definem as regiões de interesse, neste caso, as faces. O cálculo referente às qualidades das *bounding boxes* determinadas pelos métodos é realizado pela *Intersection Over Union* (IOU), que leva em conta as áreas e um *threshold*, de 50% por exemplo, da interseção dividida pela união das áreas das duas caixas. As equações para o cálculo das áreas da *bounding box*, intersecção, união e a IOU, respectivamente, são apresentadas pelas Equações (1), (2), (3) e (4).

$$Area = (x_2 - x_1) * (y_2 - y_1) \quad (1)$$

$$Intersection = (x_{2i} - x_{1i}) * (y_{2i} - y_{1i}) \quad (2)$$

$$Union = A_1 + A_2 - Intersection \quad (3)$$

$$IOU = \frac{Intersection}{Union} \quad (4)$$

Nestas, x_2, y_2, x_1 e y_1 representam as coordenadas do canto inferior direito e o canto superior esquerdo, que descrevem a *bounding box*, enquanto que x_{2i}, y_{2i}, x_{1i} e y_{1i} referem-se às mesmas coordenadas, mas da intersecção das duas caixas, já na Equação (3), A_1 e A_2 , são as áreas das caixas. Com as Equações (2) e (3) pode-se calcular a IOU, Equação (4).

A partir do cálculo da IOU e a comparação com o *threshold*, é possível realizar uma detecção. Como há imagens rotuladas para teste, as detecções podem ser avaliadas frente à resposta desejada, com o auxílio de uma matriz de confusão, uma ferramenta utilizada para avaliar o desempenho de algoritmos de classificação e/ou detecção. Esta é uma tabela que permite a visualização das previsões de um modelo em comparação com os resultados esperados, ajudando a quantificar a qualidade de classificação de um modelo. Essa matriz é particularmente relevante em problemas de classificação binária, mas pode ser estendida para problemas de classificação multiclasse [Muller e Guido 2016].

Desta forma, cada detecção pode ser atribuída a um dos quadrantes da matriz de confusão: verdadeiro positivo, há uma face e o algoritmo detectou uma face; verdadeiro negativo, não há uma face e o algoritmo não detectou uma face; falso positivo, não há uma face e o algoritmo detectou uma face; falso negativo, há uma face e o algoritmo não detectou uma face. A partir destas atribuições foram derivados valores de Acurácia, Precisão, Revocação e F1-Score, empregados na avaliação e comparação dos diferentes métodos de detecção facial.

4. Resultados

Os resultados das detecções de cada método foram coletados separadamente e registrados em arquivos de texto correspondentes. Em seguida, foi realizada uma comparação com as coordenadas das *bounding boxes* das faces, obtidas manualmente, para calcular a *Intersection Over Union* (IOU) para todos os métodos analisados. Feita a determinação dos valores de IoU, utilizou-se um *threshold* de 50% para determinar a validade das detecções (isso significa que só se contabilizou a detecção de uma face caso o IoU seja maior que 50%). Os resultados foram então classificados em Verdadeiros Positivos (VP), Falsos Positivos (FP), Falsos Negativos (FN) e Verdadeiros Negativos (VN), o que permitiu a construção das matrizes de confusão. Para cada método, foram geradas seis matrizes de confusão, uma para cada grupo de grau de iluminação avaliado e uma geral, que considera todas as detecções do método em questão. Com base nessas matrizes, foi realizada uma avaliação quantitativa do desempenho de cada método considerado.

A Figura 6 apresenta a comparação dessas quatro métricas para os sete métodos analisados. Note que esta figura apresenta uma visão geral, sem entrar em detalhes de

cada grau de iluminação abordado. Observa-se que os valores de precisão são geralmente mais altos devido à baixa incidência de falsos positivos, em contraste com a revocação, que considera os falsos negativos. Como há várias imagens com baixa iluminação, a ocorrência de falsos negativos é mais prevalente, indicando a dificuldade de detecção das faces por parte de alguns métodos. Para melhor visualização dos gráficos, os métodos estão abreviados, sendo Haar-Cascaded Default (HC-D), Haar-Cascaded FrontalFace Alt Tree (HC-FFAT), Haar-Cascaded FrontalFaceAlt 2 (HC-FFA2), dlib HOG (HOG). Os demais métodos seguem as siglas já apresentadas na seção anterior.

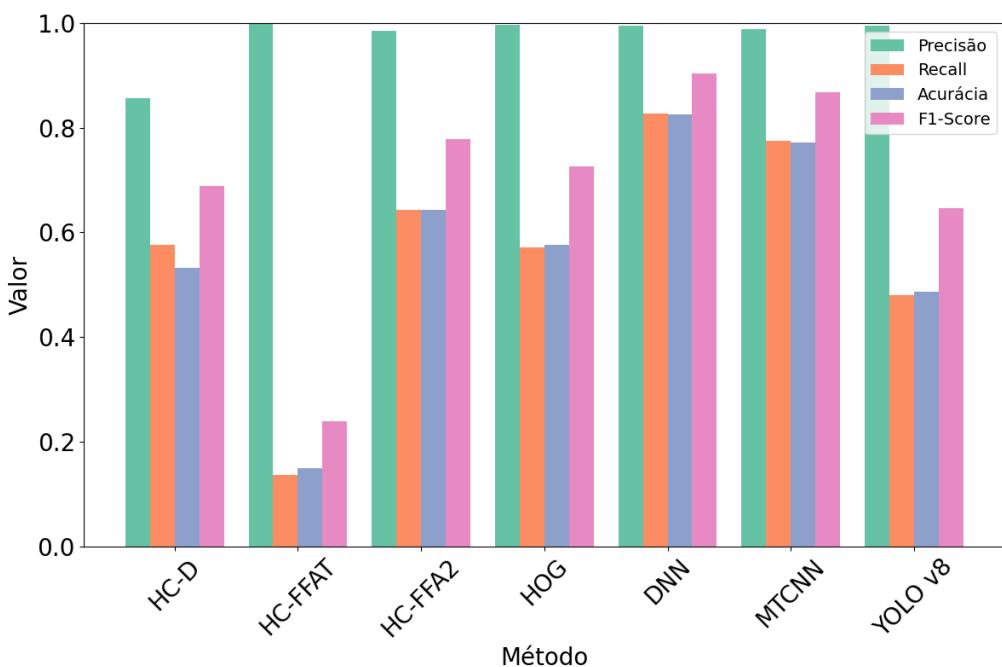


Figura 6. Comparação dos sete métodos sob diferentes métricas de qualidade.

Analizando a Figura 7 é possível observar que a acurácia teve valores próximos aos de revocação, em decorrência do pequeno número de falsos positivos e verdadeiros negativos em relação aos verdadeiros positivos e falsos negativos. O primeiro, FP, é menor, pois, em geral, detectou-se as faces no local correto quando houve uma detecção. O segundo, VN, é pequeno, pois a base de dados possuía apenas 9 imagens do ambiente (total escuridão — ausência de luz incidente na face) para cada indivíduo, que foram usadas para verdadeiros negativos. Portanto, há 252 imagens de ambiente, frente à um total de 16.380 imagens, um número pequeno. Vale ressaltar que todos os métodos não realizaram detecção de face nesses casos, portanto o valor de VN foi igual a 252 para todos.

Em geral, os métodos MTCNN e DNN demonstraram maior robustez à variação da iluminação, mantendo um desempenho consistente em todas as condições testadas. Eles alcançaram os maiores valores de acurácia e F1-Score, embora a precisão não seja a mais alta em comparação ao dlib com HOG e HC-FFAT. Excluindo-se o método Haar-Cascaded com o modelo Frontal Face Alt Tree (HC-FAT), que resultou no pior desempenho, apesar da sua maior precisão, e os dois melhores (DNN e MTCNN), os três métodos restantes, HC-D, HC-FFA2 e o YOLO v8, apresentaram resultados intermediários quanto à qualidade das detecções e próximos entre si (os métodos possuem performance similar).

Como sugere a análise de Kaur e Sharma 2023, o método MTCNN possuiu um desempenho superior se comparado ao método Haar (HC em suas diferentes variantes) e HOG considerando a revocação, acurácia e F1-Score para a detecção de faces com maiores restrições nas imagens, nesse caso, a variação na iluminação.

Uma vez que os métodos foram comparados no contexto global, isto é, sem uma segmentação por diferentes condições de iluminação, podemos agora partir para esta análise. Isto permite identificar métodos mais ou menos adequados em cenários particulares de iluminação. A Figura 7 apresenta a evolução da métrica F1-Score considerando as diferentes condições de iluminação para os sete métodos de detecção avaliados. Optou-se pela avaliação com a métrica F1-Score, uma vez que ela sumariza em um único valor, de forma combinada, os valores de precisão e revocação. O primeiro comportamento discernível no gráfico diz respeito à queda de qualidade dos modelos com o escurecimento das imagens, ainda que em diferentes taxas, conforme discutimos e detalhamos a seguir.

O método DNN apresenta os resultados mais consistentes frente à todas condições de iluminação avaliadas, sendo o melhor método em relação à métrica F1-Score. Seu desempenho é seguido de perto pelo método MTCNN, que apresenta uma queda mais acentuada de performance somente no último grau de baixa iluminação (muito escuro). Surpreendentemente, o método YOLO v8 (pertencente à mesma classe dos métodos supracitados, isto é, um método de aprendizado de máquina), apresenta uma queda acentuada de desempenho já a partir de iluminação média. Este mesmo método é o segundo pior método em cenários muito escuros, melhor apenas que o método HC-FFAT.

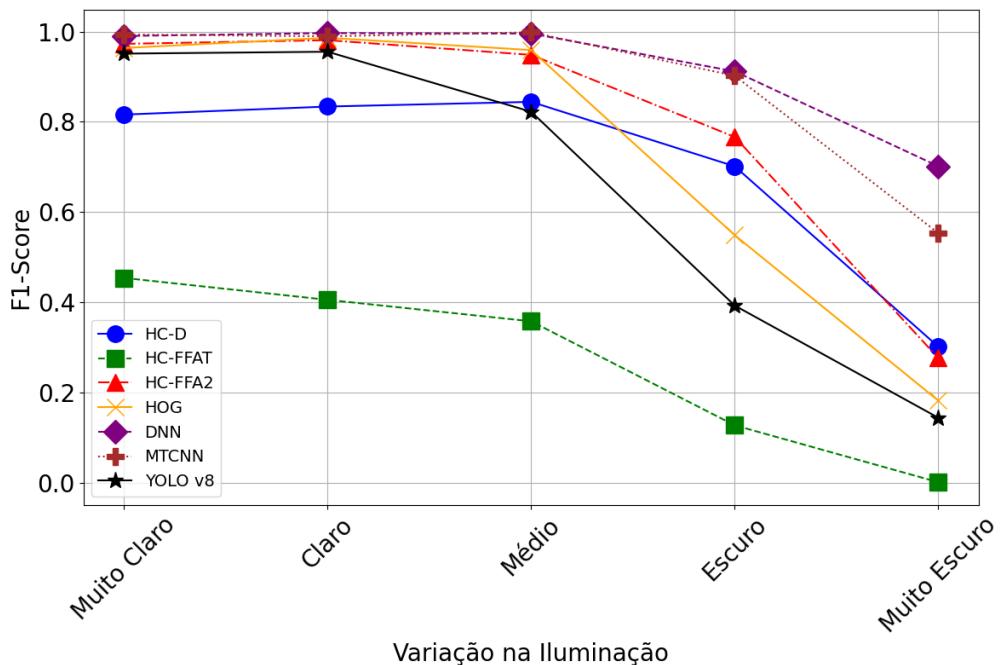


Figura 7. Comparação da evolução da F1-Score

5. Conclusões

Este estudo comparativo revelou que os métodos de detecção facial apresentam variações significativas de desempenho sob diferentes condições de iluminação. Os algoritmos ba-

seados em Haar-Cascaded e HOG mostraram-se robustos em ambientes com iluminação controlada, enquanto métodos baseados em Redes Neurais Artificiais demonstraram maior flexibilidade, embora ainda enfrentem desafios em condições de pouca luz. Esses achados destacam a necessidade de aprimoramento contínuo dos algoritmos de detecção facial, especialmente em cenários de iluminação adversa. O desenvolvimento de técnicas híbridas ou a integração de pré-processamento adaptativo de imagens podem ser caminhos promissores para superar essas limitações. Estudos futuros podem focar na otimização de algoritmos para maior resiliência em ambientes de baixa luminosidade, contribuindo para a criação de sistemas de detecção mais eficazes e confiáveis.

Referências

- Abate, A. *et al.* (2007). 2d and 3d face recognition: A survey. *Pattern Recognition Letters*, 28(14):1885–1906.
- Behling, A. (2019). Reconhecimento de emoções em vídeo utilizando redes neurais artificiais.
- Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's Journal of Software Tools*.
- Böhm, S. M. (2021). Análise de performance de um algoritmo de reconhecimento facial por visão computacional aplicado a sistemas embarcados.
- Dalal, N. e 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 vol. 1.
- Gamage, C. e Seneviratne, L. (2014). Development of a learning algorithm for facial recognition under varying illumination. In: Proceedings of the 7th internacional conference on information and automation for sustainability. 1(7):1–6.
- Georghiades, A., Belhumeur, P., e Kriegman, D. (2001). From few to many: illumination cone models for face recognition under variable lighting and pose. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6):643–660.
- Jain, A., Hong, L., e Pankanti, S. (2000). Biometric identification. *Communications of the ACM*, 43(2):90–98.
- Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., e Darrell, T. (2014). Caffe: Convolutional architecture for fast feature embedding. *arXiv preprint arXiv:1408.5093*.
- Kaur, S. e Sharma, D. (2023). Comparative study of face detection using cascaded haar, hog and mtcnn algorithms. In *2023 3rd International Conference on Advancement in Electronics & Communication Engineering (AECE)*, pages 536–541.
- King, D. E. (2009). Dlib-ml: A machine learning toolkit. *Journal of Machine Learning Research*, 10:1755–1758.
- Kumar, A., Kaur, A., e Kumar, M. (2019). Face detection techniques: a review. *Artificial Intelligence Review*, 52(2):927–948.
- LeCun, Y., Bengio, Y., e Hinton, G. (2015). Deep learning. 521(7553):436–444.

- Lee, H.-W., Peng, F.-F., Lee, X.-Y., Dai, H.-N., e Zhu, Y. (2018). Research on face detection under different lighting. In *2018 IEEE International Conference on Applied System Invention (ICASI)*, pages 1145–1148.
- Lin, T.-Y., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Perona, P., Ramanan, D., Zitnick, C. L., e Dollár, P. (2015). Microsoft coco: Common objects in context.
- Muller, A. C. e Guido, S. (2016). *Introduction to machine learning with Python*. O'Reilly, Gravenstein Highway North, Sebastopol.
- Ochoa-Villegas, M. A., Nolazco-Flores, J. A., Barron-Cano, O., e Kakadiaris, I. A. (2015). Addressing the illumination challenge in two-dimensional face recognition: a survey. *IET Computer Vision*, 9(6):978–992.
- Pankanti, S., Bolle, R. M., e Jain, A. (2000). Biometrics: The future of identification [guest editors' introduction]. *Computer*, 33(2):46–49.
- Rana, W., Pandey, R., e Kaur, J. (2022). Face recognition in different light conditions. In Smys, S., Balas, V. E., e Palanisamy, R., editors, *Inventive Computation and Information Technologies*, pages 839–850, Singapore. Springer Nature Singapore.
- Redmon, J., Divvala, S., Girshick, R., e Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788.
- Viola, P. e 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.
- Zafeiriou, S., Zhang, C., e Zhang, Z. (2015). A survey on face detection in the wild: Past, present and future. *Computer Vision and Image Understanding*, 138:1–24.
- Zhang, K., Zhang, Z., Li, Z., e Qiao, Y. (2016). Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE signal processing letters*, 23(10):1499–1503.
- Zhang, L., Wang, H., e Chen, Z. (2021). A multi-task cascaded algorithm with optimized convolution neural network for face detection. In *2021 Asia-Pacific Conference on Communications Technology and Computer Science (ACCTCS)*, pages 242–245.