

Reconhecimento facial utilizando imagens 2D

LIMA, K. A. B., PIRES, L. M., RIBAS, A.

I. INTRODUÇÃO

Em situações cotidianas, indivíduos normalmente sentem-se mais confortáveis em passar pelo processo de aquisição de imagens faciais do que de outras características biométricas mais eficientes, como íris e retina [1], [2], [3]. Esse comportamento se deve ao fato do reconhecimento por meio da face ser um processo bastante comum utilizado pelos seres humanos para distinguir indivíduos. A utilização de biometrias faciais torna possível ainda a identificação de indivíduos sem sua colaboração direta. De forma geral, a coleta de imagens é um processo simples, não-intrusivo, que não requer contato físico com o sensor de coleta, e, muitas vezes, não requer sensores complexos e de alto custo.

O reconhecimento facial, como tema de pesquisa na visão computacional, está em constante evolução e é, atualmente, importante para o desenvolvimento de aplicações de segurança e gestão administrativa. Na área de segurança pública, o reconhecimento facial é utilizado como ferramenta de apoio para identificar possíveis reincidentes, ou seja, suspeitos que já praticaram algum tipo de ato ilícito e possuem antecedentes criminais [4].

Os sistemas de biometria facial apresentam problemas que inviabilizam sua utilização em determinadas situações. Embora a face apresente padrões, como olhos, nariz e boca, não é simples definir o padrão de cada característica individualmente devido sua alta variabilidade para o mesmo indivíduo e para indivíduos diferentes. Diferentemente do que ocorre no reconhecimento de impressões digitais, onde as minúcias, a característica biométrica para essa tipo de biometria, são bem definidas. Outros pontos negativos da biometria facial são:

- É afetado por variações em pose, expressões faciais, poses, e iluminações, podendo alterar o resultado esperado [5].
- Pessoas parecidas (e.g. gêmeos) podem enganar o sistema.

Computacionalmente, um dos métodos pioneiros de biometria facial foi a utilização de análise geométrica [6]. Nessa abordagem os resultados dependem da precisão da detecção dos pontos biométricos, como olhos, nariz e

boca. Por requererem um sistema de detecção para cada uma destas características, o custo computacional pode ser elevado [7]. Nestes casos, é comum a substituição dos múltiplos detectores por alternativas que possibilitem a detecção de mais de uma classe de padrão.

Como indivíduos diferentes podem facilmente ter padrões geométricos similares, a análise geométrica por si só não possui alta confiabilidade. Por isso se faz necessário, identificar nuances na face que diferenciam indivíduos, como rugas, por processos de análise de textura [8]. Esses processos requerem a localização e normalização da face, sendo que abordagens baseadas em cores apresentam alto desempenho na localização de faces, como o seguidor de face elíptico [9], ou a análise da distribuição geométrica juntamente com padrão de cor de pele [10].

Uma abordagem que pode ser utilizada para vários problemas e está sendo utilizada na biometria facial é a representação esparsa [11]. Os autores propuseram um algoritmo de reconhecimento facial eficiente via representação esparsa em 2D, transformando a imagem representações *Fisherface*. Essa abordagem também é estendida ao tratamento de *pixels* corrompidos, onde matrizes suplementares são usadas. Os resultados experimentais mostraram desempenho satisfatório com precisão e robustez.

Diversas soluções de biometria facial utilizando imagens 3D surgiram. Entretanto, o custo de um sensor para captura eficaz desse tipo de imagem ainda é caro, inviabilizando a popularização desses sistemas. Com a utilização de sensores de baixo custo (menos eficiente), é possível obter informações 3D na forma de imagens RGB-D. Diversas abordagens utilizam métodos de PCA (*principal component analysis*) e LDA (*Linear discriminant analysis*) para auxiliar no reconhecimento biométrico [5]. Elas propõem a utilização de informações da combinação de entropia e características de saliência, juntamente com características geométricas extraídas da face.

A. Benchmark

Diversas bases de dados e sistemas automatizados podem ser usadas para analisar o desempenho de soluções

e algoritmos. Essas bases de dados estão disponíveis nas mais diversas configurações de intensidade de luz, diferentes expressões faciais [12]. Uma base muito utilizada no problema de reconhecimento de face é a Labeled Faces in the Wild (LFW) [13]. O conjunto de dados, contém 13.233 imagens de rostos recolhidos da internet. Cada face foi rotulada com o nome da pessoa retratada, e cada uma das 1680 pessoas retratadas possui duas ou mais fotos distintas no conjunto de dados. A única restrição sobre as faces é que elas foram detectadas pelo detector de face *Viola-Jones*. Na LFW existem quatro conjuntos diferentes de imagens incluindo os tipos originais e três diferentes de imagens já normalizadas. Uma característica que desperta interesse nessa base são as variações disponíveis nas imagens que incluem variação de pose, iluminação, expressões faciais, diferentes fundos de imagem, variação de raça, etnia, idade, sexo, roupas, cabelos, qualidade da câmera, saturação de cor, foco, dentre outros parâmetros.

O banco de dados LFW contém 13.233 imagens de faces. Sendo que algumas imagens contêm mais de uma face, mas é a face que contém o pixel central da que é considerado o “rosto de definição” [13] enquanto as outras podem ser descartadas ou tratadas como *background*, cada imagem recebe um nome único que é associado à pessoa que a face foi coletada, como exemplo: “GeorgeW-Bush” corresponde ao ex presidente dos EUA enquanto GeorgeHWBush corresponde ao pai também ex-presidente, existem 5749 indivíduos nessa base sendo que 1680 possuem duas ou mais imagens o restante de 4069 possuem apenas uma imagem dentro do conjunto de dados.

Diversas análises foram realizadas em [14] em diferentes conjuntos de dados como imagens individuais, grupo de imagens, vídeos, conjuntos abertos ou fechados com dados controlados e não controlados. As buscas foram feitas com protocolos de 1:N e 1:1. A análise de desempenho quantitativo deste trabalho foi realizada através das métricas de falsa aceitação (FAR, *False Acceptance Rate*) e falsa rejeição (FRR, *False Rejection Rate*).

Dentre as métricas de avaliação quantitativa de desempenho de algoritmos, as taxas de verdadeiro-positivo (TPR-*True Positive Rate*) e falso-positivo (FPR-*False Positive Rate*) são importantes [15] e refletem melhor o desempenho do sistema do que a acurácia. Elas são utilizadas para compor a curva ROC, que reflete o desempenho de classificador como o processo de reconhecimento de face.

Para o problema de treinamento e validação de algo-

ritmos para face 2D, diversos *datasets* estão disponíveis na WEB, um exemplo muito utilizado e citado em alguns artigos é o *Bosphorus database* [16] que além de faces 2D trabalha também com exemplos em 3D. Ambas as imagens são coloridas e possuem 24 marcadores para *Ground truth*, são exibidos na Figura 1:

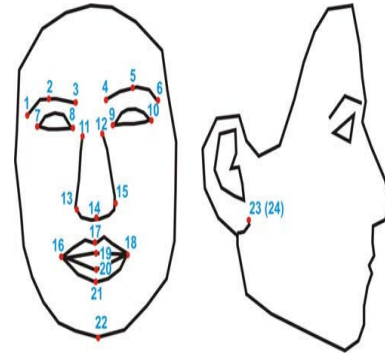


Figure 1: Áreas marcadas para geração do Template

Outro *dataset* gratuito também encontrado na literatura [17], [18], [19] é o BioID[20]. Esta base consiste em 1521 imagens em tons de cinza com resoluções de 384x286 pixels. As imagens foram obtidas de diversos indivíduos, com variações de pose e expressões faciais, condições de iluminação e escala, em quantidades aleatórias dos mesmos, com plano de fundo retratando a realidade do cotidiano.

B. Desafios

Embora o reconhecimento facial seja uma tarefa simples para os seres humanos, é extremamente complexo implementar esse processo em uma máquina, pois não se sabe ao certo como o cérebro humano realiza essa tarefa. O cérebro humano pode identificar corretamente uma pessoa a partir de sua imagem facial mesmo estando em condições desfavoráveis para o reconhecimento, como, por exemplo, em ambientes com variações de iluminação, ou até mesmo com distorções ou deformações [4]. Sistemas biométricos que sejam invariantes a essas mudanças inerentes ao ambiente compõem um dos desafios.

Além das mudanças provocadas por fatores externos, como ambiente, há as mudanças inerentes aos próprios indivíduos. A boca, por exemplo, é um elemento não rígido com alto grau de variação. Mudanças estéticas, como barba e cabelo, variam em escalas de dias e o processo de envelhecimento muda as características biométricas em meses ou anos. Robustez a essas variações das características biométricas é outro desafio a ser solucionado pela biometria.

O reconhecimento facial era uma tecnologia restrita a ambientes de alta segurança. Atualmente é utilizado em vários nichos de mercados, como indústrias e pequenas empresas. Cada um desses contextos possuem necessidades diferentes: minimizar falsa aceitação, no caso de ambientes com altos níveis de segurança, minimizar a falsa rejeição, para ambientes como pequenas empresas. Normalmente as soluções focam em um destes contextos, soluções genéricas buscam minimizar ambos os erros, a fim de obter um sistema mais robusto.

Na maioria dos ambientes, os sistemas biométricos utilizam sensores que capturam uma única pessoa por vez em ambientes com algum grau de controle sobre o ambiente. Recentemente, muitos estudos tem se dedicado a propor soluções que funcionem para ambientes externos não controlados com multidões. Essa situação específica apresenta alto índice de variação nos dados que facilmente prejudicam a precisão dos sistemas biométricos.

II. MARCAÇÕES BIOMÉTRICAS

A priori, não há informação da localização da região da face nas imagens obtida por uma câmera. É necessário que o sistema biométrico encontre a região de interesse, no caso, as faces presentes na imagem. Após a detecção dessa região de interesse, o processo de reconhecimento do indivíduo é realizado. Esse trabalho propõe a o reconhecimento facial limitado a região acima e abaixo dos olhos. A variação de expressão facial dificulta o reconhecimento utilizando a boca como região de interesse. Essa etapa deve ser treinada com base de dados genéricas para maximizar a robustez da escalabilidade da base de indivíduos.

Este estudo levantou 3 pontos biométricos: centróides de ambos os olhos e o ponto mais inferior do nariz da imagem. De uma base de 50 indivíduos diferentes com 5 imagens para cada indivíduo, foram marcadas todas as imagens da base utilizando a média de 3 especialistas sendo que cada um marcou os 5 pontos: ponto mais a esquerda, mais a direita e ponto central mais inferior do nariz. Foi calculado a média das marcações de cada olho a fim de encontrar a respectiva localização de cada centróide. Uma vez gerados os centróides extraiu-se a média das marcações dos três especialistas, sendo que essa marcação média foi utilizada como parâmetro para testar a sensibilidade do reconhecimento biométrico.

Para detectar os olhos, nariz e boca, utilizamos a detecção por meio de Haar-like features pela implementação do Haar Cascade[15]. O arquivos de treinamento para estes elementos foi feito em outros

trabalhos [21], [22] e posteriormente foi avaliado[23]. O ponto médio marcado pelos especialistas como *ground-truth* e a detecção dos elementos através das Haar-like features são aproximações. Desse modo a detecção é considerada correta se a distância d obedece a restrição, $d(\text{groundTruth}, \text{predito}) \leq \text{thresh}$, para o limiar thresh .

Para avaliar o desempenho do detector dessas marcações biométricas, utilizamos detecção de olhos, nariz e boca. De modo que, a marcação biométrica é estimada com as coordenadas do retângulo de detecção dado pelas equações,

$$\begin{aligned} Olho_x &= \frac{Rect_{olho,x} + Rect_{olho,width}}{2}, \\ Olho_y &= \frac{Rect_{olho,y} + Rect_{olho,height}}{2}, \end{aligned} \quad (1)$$

correspondendo ao centróide do retângulo da detecção dos olhos. Para a marcação biométrica do ponto inferior, detectamos a média das coordenadas x e coordenada y mais baixa para a detecção do nariz como,

$$\begin{aligned} Nariz_x &= \frac{Rect_{nariz,x} + Rect_{nariz,width}}{2}, \\ Nariz_y &= Rect_{nariz,y} + Rect_{nariz,height}, \end{aligned} \quad (2)$$

ou através da detecção da boca, pela média das coordenadas x e coordenada y mais alta,

$$\begin{aligned} Boca_x &= \frac{Rect_{boca,x} + Rect_{boca,width}}{2}, \\ Boca_y &= Rect_{boca,y}. \end{aligned} \quad (3)$$

A análise de desempenho utilizando os verdadeiros positivos de forma isolada não reflete a qualidade das respostas da detecção. Há dois olhos para cada indivíduo e a metodologia de análise anterior, não verifica se ambos os olhos estão sendo detectados. Assim como não verificamos a porcentagem de indivíduos que falham na detecção das outras métricas. A Figura 2 mostra a análise a porcentagem de indivíduos cujo o detector acerta considerando o par dos olhos na detecção de cada característica biométrica, contudo sem validação.

Há um baixo número de indivíduos que não é possível detectar um dos dois olhos, de modo que a região de interesse para identificação do indivíduo não possa ser obtida. Considerando as características biométricas de nariz e boca, a precisão dessas detecções não são altas, as falhas nelas se reduzem com valores mais altos para o limiar τ , revelando a imprecisão dessa métrica no sistema biométrico.

As marcações biométricas referente aos olhos podem ser validadas para eliminar detecções errôneas. Faces

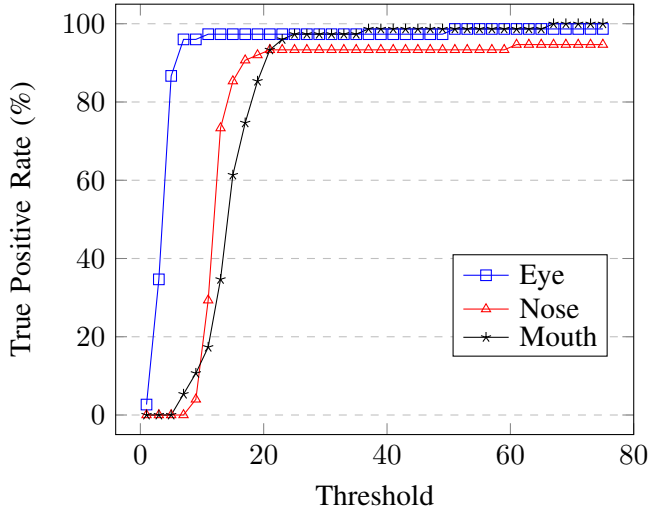


Figure 2: Detection Rate

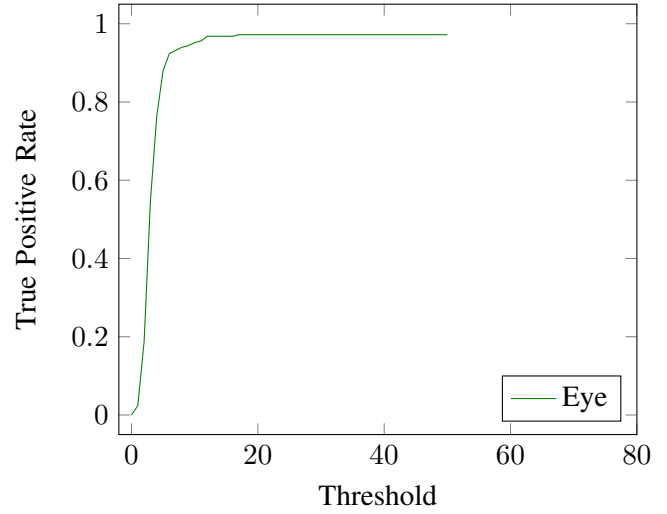


Figure 3: Detection Rate

são compostas por dois olhos com ângulo entre as marcações previsível se a face estiver na posição usual. De modo que é possível eliminar detecções de olhos cuja a permutação com as outras estimações de olhos não formem uma face válida. Primeiro, agrupamos os pontos em um raio τ e as médias μ_{eyes} . A validação dessas médias é dada pela equação,

$$v_{e_i, e_j} = \begin{cases} 1, & \text{se } -0.25 \leq \arctan2(e_{j,y} - e_{i,y}, e_{j,x} - e_{i,x}) \leq 0.25 \text{ e} \\ & ||e_j - e_i||_2 > \tau * 1.5 \\ 0, & \text{caso contrário} \end{cases} \quad (4)$$

de modo que $v_{i,j}$ indica as permutações de olhos válidas, $v_{i,j} = 1$.

Para a aprimorar o processo de detecção das marcações e extração das *features*, a equalização do canal de iluminação pode ser feita. Neste processo a imagem é convertida para o modelo de cor YCbCr, de modo que o canal Y é o canal de iluminação enquanto Cb e Cr são canais de cor. Y é normalizado equalizando o seu histograma e a imagem é convertida para o modelo de cor RGB novamente.

De modo que a marcação biométrica dos olhos é a mais confiável para extrair a região de interesse que é utilizada nos processos biométricos. A detecção mostrada na Figura 3 é feita utilizando o processo de validação juntamente com equalização do canal de iluminação. Nela podemos observar uma pequena melhora na detecção proveniente da equalização da iluminação.

III. NORMALIZAÇÃO E EXTRAÇÃO DA REGIÃO DE INTERESSE

Devido às variações de poses nas imagens da base de dados proposta, como pode ser visto na imagem 4a. É necessária a normalização das poses para auxiliar e melhorar a eficácia do processo de classificação. Uma vez localizados os dois olhos na imagem, a imagem é rotacionada afim de que o ângulo formado por esses pontos seja mínimo.

Uma vez feito esse processo para todas as imagens, a extração de interesse é feita utilizando a distância d entre os dois pontos dos olhos. A região de interesse corresponde a vizinhança $d/2$ dos pontos dos olhos conforme é demonstrado na imagem 4b.

IV. BIOMETRIA

A descrição da região de interesse é feita pelo Padrão Binário Local (LBP-*Local Binary Pattern*) [24] que pode ser visto na Figura 5a. Após o cálculo do padrão local de cada *pixel* da região de interesse, visto de forma genérica na Figura 5b, o histograma é feito utilizando 26 agrupamentos. A identificação de um indivíduo é feita pela comparação dos histogramas ordenando pelos indivíduos mais semelhantes. A comparação é feita pela distância Qui-quadrado χ^2 .

A avaliação do sistema é dada por duas formas de validação. Primeiramente fizemos uma validação de forma que todos os indivíduos sejam conhecidos. Com 3 imagens de cada indivíduo no conjunto de treinamento e duas imagens no conjunto de teste. Esse processo é realizado três vezes e os resultados obtidos são as médias dos experimentos de cada par treino-teste.

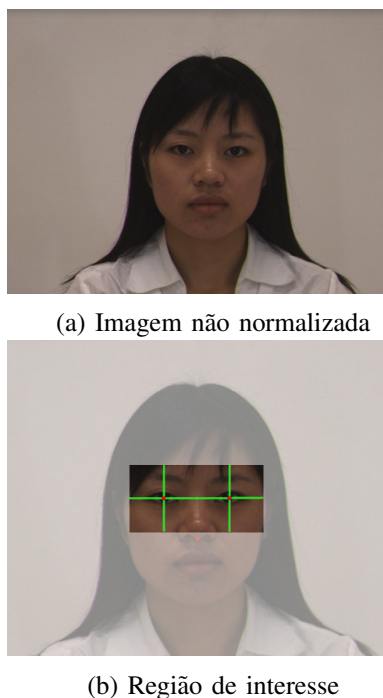
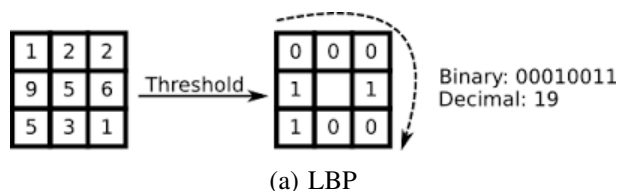


Figure 4: Imagem de face e Região de interesse



(b) LBP de uma região de face genérica

Uma forma de avaliar o desempenho do sistema biométrico é verificando se a identidade verdadeira do indivíduo consta em uma lista dos N indivíduos mais similares. O resultado para essa métrica de desempenho pode ser observado na Figura 6. Ela mostra que o sistema apresenta uma acurácia acima de 90% considerando unicamente o mais semelhante e acima de 97% considerando os cinco mais semelhantes.

Outra forma de avaliar o sistema biométrico é a curva ROC. A curva ROC avalia as taxas de *Recall* e *Precision* com as variações do limiar separação. Essa curva mostra o comportamento de classificação de um problema de uma classe, desse modo fazemos um *matching* “um-

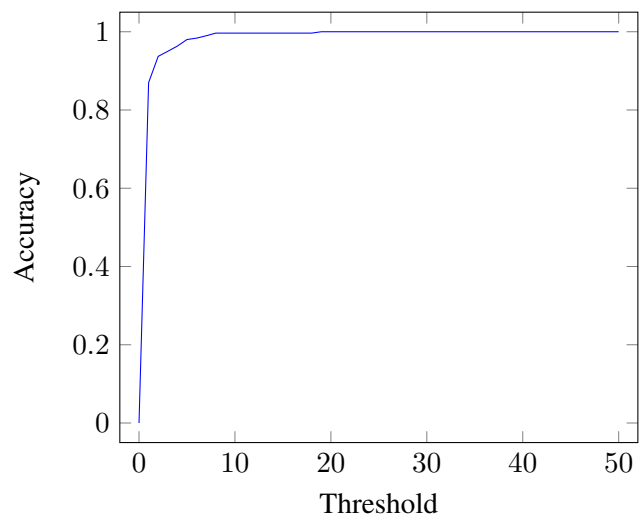


Figure 6: Rank Top N

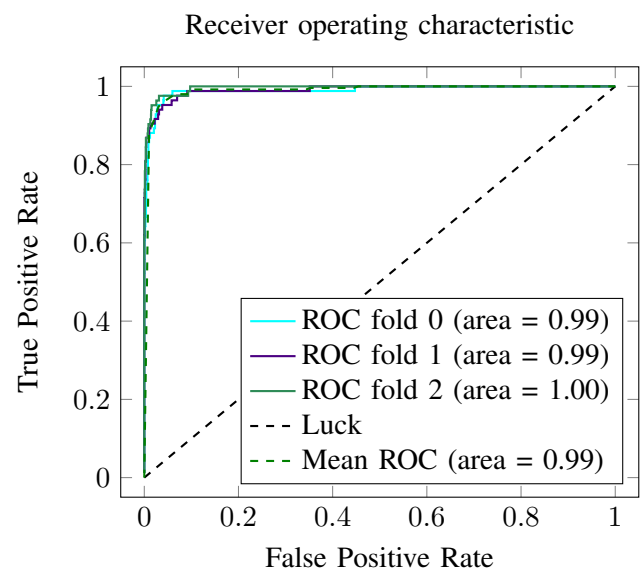


Figure 7: Curva ROC

contra-todos”. O resultado dessa métrica de avaliação é mostrado na Figura 7.

V. MULTIBIOMETRIA

Alguns sistemas consolidam as evidências de múltiplas fontes de informação biométrica para determinar de forma confiável a identidade de um indivíduo, esses sistemas são conhecidos como multibiométricos. A abordagem multibiométrica alivia as limitações de sistemas baseados em uma única biometria, utilizando para isso, as variações de várias fontes para identificar corretamente os indivíduos.

Para aplicar a multibiometria no nosso sistema, recebemos uma base contendo as distâncias entre as

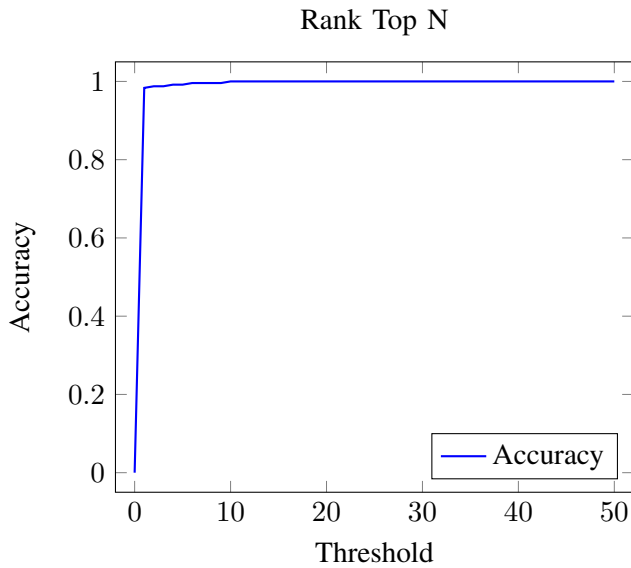


Figure 8: Rank Top N Multibiometria

amostras, seguindo o formato " $m_n p_q distance$ ", onde m e p são referentes ao identificador do indivíduo, e n e q referente a uma amostra do indivíduo. Utilizando mapeamentos $m \rightarrow p$ e $n \rightarrow q$, a distancias para cada biometria podem ser combinadas. Os valores de distancias são normalizados seguindo a equação,

$$dist = 1 - \frac{distance}{\sum distance}. \quad (5)$$

Depois de normalizadas, as distâncias $dist$ para cada biometria utilizando nos mapeamentos $m \rightarrow p$ e $n \rightarrow q$ são somadas. Os resultados são gerados de forma similar ao da biometria simples. A primeira análise é a curva de ranqueamento Top N. O resultado para essa métrica de desempenho pode ser observado na Figura 8. Ela mostra que o sistema, após a inserção da distância de *fingerprinth* melhorou seu desempenho. Considerando o indivíduo mais próximo no Rank, a acurácia foi de 98%.

Outra avaliação é a curva ROC, a curva ROC permite analisar um único limiar de separação para predizer a classe sem que seja necessário avaliar todas as classes. A Curva ROC para o sistema multibiométrico pode ser observada na Figura 9, onde observase uma redução da área sob a curva. A redução da capacidade de classificação ocorre devido à abordagem escolhida para combinação das distâncias não ter muito efetiva para separar os dados.

REFERENCES

[1] T. E. de Campos, *Técnicas de seleção de características com aplicações em reconhecimento de faces*. PhD thesis, Universidade de São Paulo, 2001.

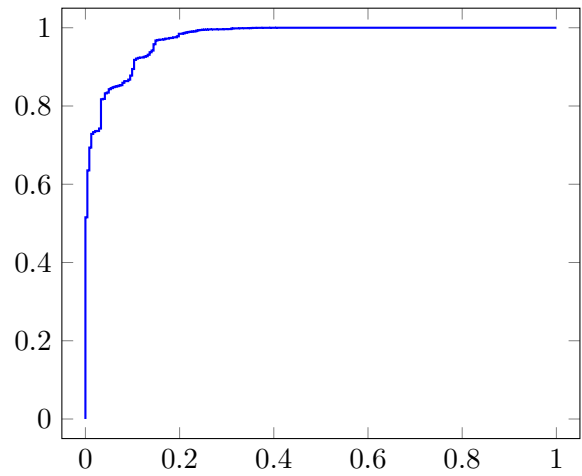


Figure 9: Curva ROC Multibiometria

- [2] S. Pankanti, R. M. Bolle, and A. Jain, "Biometrics: The future of identification [guest eeditors' introduction]," *Computer*, vol. 33, no. 2, pp. 46–49, 2000.
- [3] N. K. Ratha, A. Senior, and R. M. Bolle, "Automated biometrics," in *International Conference on Advances in Pattern Recognition*, pp. 447–455, Springer, 2001.
- [4] R. M. GUIMARÃES, "Desenvolvimento de um protótipo de software de reconhecimento facial de tempo real para registro eletrônico de ponto em ambientes indoor com utilização do dispositivo kinect.," Master's thesis, Universidade FUMEC, 2015.
- [5] R. S. Gaurav Goswami, Mayank Vatsa, "Rgb-d face recognition with texture and attribute features," *IEEE Transactions on Information Forensics and Security*, vol. 9, october 2014.
- [6] Y. B. Brudnyi, Alexander, *Methods of Geometric Analysis in Extension and Trace Problemns*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2012.
- [7] C. J. P. Passarinho, *Uma abordagem dinâmica para detecção e seguimento de face em vídeos coloridos em ambientes não controlados*. PhD thesis, Universidade Federal do Espírito Santo, 2012.
- [8] R. C. Gonzalez and R. E. Woods, *Digital Image Processing (3rd Edition)*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2006.
- [9] S. Birchfield, "An elliptical head tracker," in *Signals, Systems amp; Computers, 1997. Conference Record of the Thirty-First Asilomar Conference on*, vol. 2, pp. 1710–1714 vol.2, Nov 1997.
- [10] K. Nummiaro, E. Koller-Meier, and L. V. Gool, "An adaptive color-based particle filter," 2002.
- [11] G. Cheng and Z. Song, "Robust face recognition based on sparse representation in 2d fisherface space," *Optik - International Journal for Light and Electron Optics*, vol. 125, no. 12, pp. 2804–2808, 2014.
- [12] A. K. Jain and S. Z. Li, *Handbook of Face Recognition*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2005.
- [13] G. B. Huang, M. Mattar, T. Berg, and E. Learned-miller, "E.: Labeled faces in the wild: A database for studying face recognition in unconstrained environments," tech. rep., 2007.
- [14] S. Mau, F. Dadgostar, I. Cullinan, A. Bigdeli, and B. C. Lovell, "A face biometric benchmarking review and characterisation,"

- in *Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on*, pp. 2120–2127, Nov 2011.
- [15] P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” pp. 511–518, 2001.
 - [16] E. A. et al:14, “The bosphorus database,” 2009.
 - [17] O. Jesorsky, K. J. Kirchberg, and R. W. Frischholz, “Robust face detection using the hausdorff distance,” pp. 90–95, Springer, 2001.
 - [18] G. M. ARAÚJO, “Algoritmo para reconhecimento de características faciais baseado em filtros de correlação,” Master’s thesis, Universidade Federal do Rio de Janeiro, Fevereiro 2010.
 - [19] M. R. I. Rizoan Toufiq, “Face recognition system using pca-ann technique with feature fusion method,” *Electrical Engineering and Information E Communication Technology (ICEEICT), 2014 International Conference on*, Abril 2014.
 - [20] “Bioid face database.” Último acesso em setembro de 2016.
 - [21] M. Castrillón Santana, O. Déniz Suárez, M. Hernández Tejera, and C. Guerra Artal, “Encara2: Real-time detection of multiple faces at different resolutions in video streams,” *Journal of Visual Communication and Image Representation*, pp. 130–140, April 2007.
 - [22] M. Castrillón Santana, O. Déniz Suárez, D. Hernández Sosa, and J. Lorenzo Navarro, “Using incremental principal component analysis to learn a gender classifier automatically,” in *1st Spanish Workshop on Biometrics*, (Girona, Spain), June 2007.
 - [23] “Face and facial feature detection evaluation,” in *Third International Conference on Computer Vision Theory and Applications, VISAPP08*, January 2008.
 - [24] T. Ahonen, A. Hadid, and M. Pietikainen, “Face description with local binary patterns: Application to face recognition,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, pp. 2037–2041, Dec. 2006.