

Detecção de Indivíduos não Registrados em Galerias de Faces

Cassio E. dos Santos Jr. e William Robson Schwartz

Departamento de Ciência da Computação
Universidade Federal de Minas Gerais
Belo Horizonte - MG - Brasil

{cass,william}@dcc.ufmg.br

Abstract. *Face recognition is commonly used in tasks such as biometric and video surveillance. If one wants to track subjects using images captured from several cameras, the surveillance system must correctly associate identities of face images with those subjects appearing in images from other cameras. Along with the aforementioned problem, the re-enrollment of a subject (enrollment of a subject that was already in the gallery), decreases the overall accuracy of the face recognition, as it will be shown in this work. Motivated by those two issues, we propose five approaches to detect unenrolled subjects in face galleries. The approaches are evaluated on the FRGC and Pubfig83 datasets, which are standard and challenging datasets for face recognition.*

Resumo. *O uso de reconhecimento de faces é comum em tarefas tal como biometria ou vigilância por meio de vídeos. Se é desejável rastrear indivíduos utilizando câmeras distribuídas em um ambiente, é necessário determinar se uma imagem de face corresponde a algum indivíduo que aparece nas imagens de faces das demais câmeras. Além do problema citado anteriormente, o segundo registro de um indivíduo que já está registrado na galeria reduz a acurácia geral do reconhecimento de faces assim como será mostrado neste trabalho. Motivado por estes dois problemas, este trabalho descreve cinco novas abordagens para detectar indivíduos não registrados na galeria de faces. As abordagens são avaliadas nas bases de dados FRGC e Pubfig83.*

1. Introdução

Há três tarefas no reconhecimento de faces [Chellappa et al. 2010]: *verificação*, que determina se duas dadas imagens de face pertencem ou não ao mesmo indivíduo; *identificação*, que objetiva determinar a identidade de uma dada imagem de face de acordo com uma galeria de faces conhecidas; e *watch-list*, no qual a tarefa é determinar se uma dada imagem de face corresponde a algum indivíduo na galeria de faces identificá-lo caso corresponda.

Verificação de faces difere de *watch-list* no aspecto que, no *watch-list*, todos os indivíduos da galeria são considerados para determinar se uma face é desconhecida (não registrado na galeria de faces). De fato, a tarefa de verificação de faces é equivalente à tarefa de *watch-list* quando apenas um indivíduo está registrado na galeria de faces. Verificação de faces é tipicamente utilizado em sistemas de controle de acesso onde o

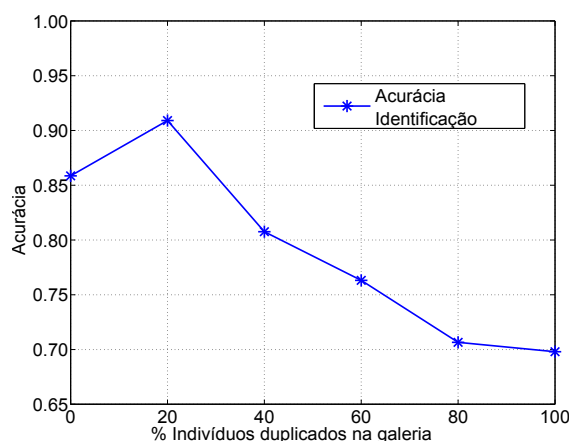


Figura 1. Taxa de reconhecimento da abordagem de identificação de faces proposta em [Schwartz et al. 2012] quando uma percentagem dos indivíduos na galeria são duplicados. Os experimentos são referentes a base de dados FRGC [Phillips et al. 2005], Experimento 4. O número de indivíduos registrados na galeria foi fixado para comparação justa uma vez que uma grande quantidade de indivíduos registrados pode reduzir a taxa de reconhecimento.

indivíduo se identifica para o sistema e este deve verificar a identidade declarada deste indivíduo de acordo com a imagem de sua face.

Watch-list é um caso particular de reconhecimento em conjunto aberto [Wechsler 2007]. No *watch-list* é esperado que a maioria das imagens de faces sendo testadas não representem um indivíduo na galeria. Um cenário típico de *watch-list* é a detecção de indivíduos suspeitos em imagens de câmeras de vigilância. Neste cenário, é mais provável que a maioria dos indivíduos que aparecem nas imagens não sejam suspeitos. Para o reconhecimento em conjunto aberto, o interesse pode ser rastrear indivíduos em vídeos capturados por várias câmeras. Dado que um indivíduo aparece em um vídeo, é necessário determinar se este indivíduo não aparece nos vídeos das outras câmeras antes de registrá-lo como um novo na galeria de faces conhecidas.

Cuidado deve ser tomado ao registrar novos indivíduos na galeria de faces uma vez que se o indivíduo sendo registrado já estiver na galeria, a adição de um segundo registro deste pode reduzir a acurácia geral do reconhecimento de faces. A Figura 1 ilustra este efeito quando uma percentagem dos indivíduos da galeria são duplicados. Embora o gráfico apresente uma tendência decrescente conforme o aumento do número de registros duplicados, o aumento na acurácia para 20% dos indivíduos duplicados na Figura 1 deve-se à escolha aleatória de um subconjunto de indivíduos da FRGC mais fácil de serem classificados.

Dentre as cinco abordagens descritas neste trabalho, uma abordagem é baseada na comparação da imagem de face de teste com um conjunto independente de faces comuns (descrito na Seção 3.1). As outras quatro abordagens são baseadas na resposta de identificação para cada indivíduo da galeria. Três dessas abordagens são baseadas na separação entre essas respostas (descritas nas Seções 3.2, 3.3 e 3.4) e uma abordagem é baseada na distribuição dessas respostas (descrita na Seção 3.5). Essas abordagens variam quanto às características exploradas nos dados e quanto à escalabilidade com relação ao

número de indivíduos registrados na galeria. Para identificação utiliza-se a abordagem proposta em [Schwartz et al. 2012] por esta ser robusta a mudanças na expressão facial, iluminação, envelhecimento e apresentar escalabilidade [Carlos et al. 2013]. As abordagens foram avaliadas nas bases de dados FRGC e Pubfig83, que são bases de dados padrão para reconhecimento de faces.

O restante deste trabalho está organizado da seguinte forma: a Seção 2 descreve trabalhos relacionados na literatura; a Seção 3 descreve as abordagens utilizadas neste trabalho; experimentos estão na Seção 4; a conclusão está na Seção 5.

2. Trabalhos Relacionados

Esta Seção descreve resumidamente alguns dos trabalhos relacionados na literatura. A Seção 2.1 descreve abordagens para reconhecimento de face. Descrição de características está descrito na Seção 2.2. Avanços no reconhecimento de faces em conjunto aberto estão descritos na Seção 2.3.

2.1. Reconhecimento de Faces

Abordagens para reconhecimento de faces na literatura podem ser divididos em dois grupos [Zou et al. 2007]: abordagens que descrevem imagens de faces de forma *holística* ou abordagens que descrevem apenas *pontos de interesse* da imagem. Abordagens baseadas na descrição holística [Turk and Pentland 1991, Etemad and Chellappa 1997] da imagem de face geralmente apresentam problemas de generalização uma vez que mudanças de pose e iluminação podem afetar significativamente a aparência geral da imagem. Estas abordagens geralmente requerem várias amostras de um indivíduo em diferentes condições para modelar corretamente as variações de aparência daquele indivíduo. Abordagens baseadas na descrição de pontos de interesse [Cao et al. 2010, Kumar et al. 2009] generalizam mais facilmente do que abordagens baseadas em descrição holística. Visto que estas abordagens focam na descrição de alguns pontos da face (como canto dos olhos, ponta do nariz, boca, entre outros), se alguns destes pontos estiverem ocultos devido a variação de pose, efeitos de sombra ou oclusão, os demais pontos podem ser utilizados para identificar a face. No entanto, abordagens baseadas em pontos de interesse ainda apresentam desafios quanto ao modo de representar uma face de forma completa.

Abordagens baseadas na descrição holística da face, tal como em [Turk and Pentland 1991] e [Etemad and Chellappa 1997], representam imagens de face utilizando subespaços. A vantagem de utilizar a representação em subespaço é a representação compacta que resulta em redução de ruído, baixo consumo de espaço e baixo custo computacional. Assim como demonstrado em [Etemad and Chellappa 1997] e [Schwartz et al. 2012], selecionar subespaços discriminativos entre os indivíduos da galeria aumenta a acurácia em comparação com métodos como Análise de Componentes Principais, *Principal Component Analysis* (PCA) [Turk and Pentland 1991].

A abordagem descrita em [Schwartz et al. 2012] utiliza *Partial Least Squares* (PLS). PLS é uma técnica de projeção e regressão supervisionada capaz de considerar descritores com elevado número de dimensões. Esta técnica consiste em projetar descritores e respostas para um subespaço semelhante ao PCA. Este subespaço é escolhido de forma a maximizar a covariância entre descritores e respostas. Em seguida, regressão simples utilizando mínimos quadrados é utilizado entre descritores e respostas projetados.

Para cada indivíduo na galeria, um modelo PLS é calculado considerando amostras deste indivíduo, com valor de resposta +1, e amostras dos demais indivíduos, com valor de resposta -1. Em seguida, amostras de teste são submetidas a cada um dos modelos PLS e a identidade da amostra é associada ao modelo que retornar maior resposta. Para reduzir o custo de testar uma amostra em cada um dos modelos PLS, uma estrutura de árvore é sugerida. Em [Carlos et al. 2013], os autores sugerem o uso de um subconjunto de amostras de indivíduos para reduzir o custo em re-treinar todos os modelos PLS quando uma nova amostra é registrada na galeria de faces.

2.2. Descritores de Características

Enquanto pesquisadores desenvolvem descritores de características eficientes para diferentes objetivos, estes descritores podem ser utilizados com sucesso no reconhecimento de faces. Exemplos de reconhecimento de faces utilizando SIFT e HOG podem ser encontrados em [Luo et al. 2007] e [Albiol et al. 2008], respectivamente. Os descritores mais comuns utilizados no reconhecimento de faces são baseados em padrões binários locais (este descritor é mais conhecido na literatura como *Local Binary Pattern* (LBP) [Ahonen et al. 2004]), variações do LBP [Wolf et al. 2008] e filtros de Gabor [Zhang et al. 2007]. Existem também trabalhos que combinam LBP e filtros de Gabor [Tan and Triggs 2007] além de vários outros descritores de características [Zhou and Sadka 2011].

Quando vários descritores de características são utilizados, técnicas de projeção em subespaço se tornam necessárias para contornar problemas relacionados ao elevado número de dimensões do espaço de descritores. O trabalho em [Schwartz et al. 2012] utiliza a redução de dimensionalidade realizada pelo PLS em uma combinação de vários descritores.

2.3. Reconhecimento de Faces em Conjunto Aberto

Não há muitos trabalhos na literatura que abordam o problema de reconhecimento de faces em um conjunto aberto. A abordagem proposta em [Stallkamp et al. 2007] considera valores relacionados à confiança de classificação de indivíduos registrados na galeria para determinar um limiar usado para decidir se uma amostra de teste está registrada na galeria ou não. A confiança de classificação é calculada testando amostras de um indivíduo na galeria em modelos treinados utilizando amostras dos demais indivíduos da galeria. A desvantagem desta abordagem é que esta depende de um número considerável de indivíduos na galeria durante o treinamento. Este número de indivíduos deve ser suficiente para garantir acurácia da abordagem. Esta desvantagem pode inviabilizar tarefas como *watch-list*, onde a abordagem deve ser capaz de trabalhar com poucos indivíduos registrados na galeria.

Em [Li and Wechsler 2005], os autores utilizam a confiança de classificação da amostra de teste considerando k amostras mais próximas da galeria, que não sejam do indivíduo considerado. Se a confiança de classificação da amostra de teste não for diferente o suficiente entre os valores das k amostras da galeria, é provável que o indivíduo da face de teste não esteja na galeria. Para garantir que a abordagem apresente acurácia para poucos indivíduos na galeria, os autores utilizam de um conjunto extra de imagens de faces que não correspondem a nenhum indivíduo no treinamento ou teste. O uso deste

conjunto extra é comum em abordagens que lidam com tarefas de verificação de faces como em [Guo et al. 2011] e [Wolf et al. 2011a].

Em [Wolf et al. 2011b], os autores utilizam um classificador para cada indivíduo na galeria para decidir se uma imagem de face de teste pertence àquele indivíduo ou não. Um conjunto denso de amostras de faces é necessário para definir adequadamente a fronteira de decisão de cada classificador. Os autores utilizam amostras sintéticas geradas interpolando cor e posição dos pixels de imagens de faces de um indivíduo em direção a imagens de faces dos demais indivíduos da galeria.

3. Abordagens

Entre as abordagens propostas neste trabalho, uma é baseada na utilização de um *conjunto independente* e quatro são baseadas em um *conjunto de validação*.

Na abordagem baseada em um conjunto independente, a ideia é representar faces desconhecidas comuns que não sejam aquelas que estão na galeria por meio de um conjunto independente de faces que não correspondem a nenhum indivíduo no treinamento ou teste. A motivação do uso de um conjunto independente segue outras abordagens de verificação de faces, como em [Guo et al. 2011] e [Wolf et al. 2011a].

A estimativa de parâmetros de quatro das abordagens propostas neste trabalho dependem da estimativa das respostas de identificação de faces no teste. Esta estimativa é obtida utilizando um subconjunto das amostras de treinamento como teste para a abordagem de identificação de faces treinada com o restante das amostras de treinamento. O objetivo dessa separação é evitar testar as mesmas amostras utilizadas no treinamento do modelo, o que geraria respostas não correspondentes aquelas esperadas para as amostras de teste. Entretanto, é necessário um número suficiente de amostras de treinamento (pelo menos quatro por indivíduo nos experimentos deste trabalho) para realizar tal separação. O conjunto de validação é construído seguindo o esquema de validação cruzada. As amostras de cada indivíduo são divididas aleatoriamente em um número específico de partições independentes. A cada iteração, uma partição é utilizada como teste e as demais para o treinamento. Dessa forma é possível obter um número de respostas estimadas igual ao número de amostras de treinamento.

As próximas subseções descrevem as cinco abordagens propostas neste trabalho para detectar indivíduos não registrados em galerias de faces. A abordagem baseada no conjunto independente está descrita na Seção 3.1. Abordagens baseadas em vetores de suporte, mais conhecido como *Support Vector Machine* (SVM) na literatura, estão nas Seções 3.2 e 3.3. A Seção 3.4 descreve a abordagem baseada em mínimos quadrados. Enfim, a abordagem baseada na distribuição das respostas de identificação de faces está descrita na Seção 3.5.

3.1. Abordagem Baseada no Conjunto Independente

Esta abordagem treina um classificador para detectar se o indivíduo de uma amostra apresentada para o sistema está registrado na galeria de forma que um teste neste classificador seja suficiente para determinar se a face de teste pertence a um indivíduo conhecido ou não. A vantagem desta abordagem com relação as demais abordagens descritas neste trabalho é que ela não requer que a imagem de face de teste seja apresentada em cada um dos modelos de identificação de faces.

O classificador consiste em um modelo PLS, referido neste trabalho como *Modelo de Conjunto Independente* (MCI), semelhante a aquele utilizado na identificação de faces em [Schwartz et al. 2012]. São utilizadas no MCI as amostras do conjunto independente, descrito no início da Seção 3, com respostas +1 e amostras da galeria com respostas -1 no treinamento. Dessa forma, o objetivo desta abordagem é fazer com que o MCI aprenda a distinguir características que discriminem os indivíduos da galeria daqueles presentes no conjunto independente. Para avaliar uma amostra de teste, esta amostra é submetida ao MCI e se este retornar um valor de resposta maior do que um limiar τ , é assumido que a face de teste é desconhecida.

3.2. Abordagem Baseada em SVM para Todos Indivíduos

O objetivo desta abordagem é separar respostas de indivíduos conhecidos das respostas de indivíduos desconhecidos utilizando o critério de máxima margem do classificador SVM.

Como as respostas de indivíduos desconhecidos tendem a apresentar valores baixos, é esperado que estas respostas se aproximem de um agrupamento em torno de um ponto comum no espaço de respostas. Indivíduos conhecidos, por outro lado, tendem a apresentar valores de resposta elevados para o modelo PLS que corresponde a sua respectiva identidade e valores baixos para os demais modelos. Dessa forma, ao utilizar o SVM, é necessário utilizar funções de *kernel* para mapear as respostas em um espaço de dimensionalidade elevada que permita separação linear entre respostas de indivíduos conhecidos e desconhecidos.

Para o SVM, é comum o uso da técnica *kernel-trick* para evitar realizar cálculos no espaço de dimensionalidade elevada. A técnica de *kernel-trick* consiste em aplicar uma função $K(A, B)$ que retorne o produto interno entre dois vetores, A e B , sem que seja necessário projetá-los para o espaço de dimensionalidade elevada. Neste contexto, para classificar uma amostra de teste, é utilizada a Equação 1, onde x_i denota o i -ésimo vetor de suporte do SVM, α_i denota o respectivo peso do vetor de suporte e b representa o deslocamento do hiperplano com relação à origem do espaço de respostas. A Equação 1 equivale a equação de classificação do SVM sem que seja aplicada a função sinal. A Equação 1 também é relacionada com a aproximação de verossimilhança descrita em [Schölkopf et al. 2001]. Foram utilizadas as funções *kernel* radial e polinomial nos experimentos deste trabalho.

$$\sum_{i=0}^N \alpha_i K(x_i, x) + b \geq \tau \quad (1)$$

Visto que todas as respostas de validação, utilizadas para treinar o SVM, incluem uma resposta para um indivíduo da galeria conforme discutido no início da Seção 3, é necessário gerar amostras sintéticas representando exemplos negativos para o SVM. Amostras sintéticas são geradas substituindo a resposta para o respectivo indivíduo por outra resposta aleatória daquela amostra.

3.3. Abordagem Baseada em SVM por Indivíduo

O objetivo desta abordagem é separar respostas de um indivíduo das outras respostas (do conjunto independente ou de outros indivíduos na galeria) utilizando o critério de máxima margem semelhante a abordagem que utiliza o SVM considerando todos os indivíduos descrita na Seção 3.2.

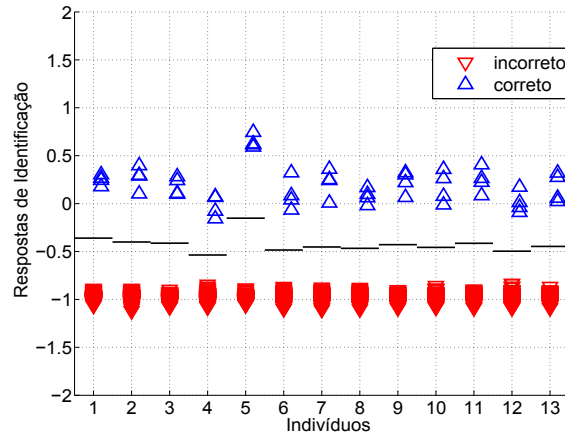


Figura 2. Respostas para 13 indivíduos do conjunto de validação da base de dados FRGC, Experimento 2. Respostas corretas para o respectivo indivíduo estão representadas por triângulos azuis apontados para cima enquanto que triângulos vermelhos apontados para baixo representam respostas incorretas. Segmentos horizontais pretos representam limiares calculados utilizando SVM. Respostas corretas foram deslocadas um pouco para a direita para facilitar a visualização.

Esta abordagem diferente daquela abordagem que considera todos os indivíduos no aspecto que nesta abordagem são utilizados classificadores para cada indivíduo na galeria. Neste contexto, para testar uma amostra é utilizada a equação do hiperplano do SVM descrita na Equação 2. Para o classificador SVM do i -ésimo indivíduo na galeria, x_j denota o j -ésimo vetor de suporte, α_j denota o respectivo peso do vetor de suporte e b é relacionado com o desvio do hiperplano com relação à origem do espaço de respostas. O limiar de decisão para o i -ésimo indivíduo, τ_i , é dado pela Equação 3. Exemplos de limiares de decisão calculados por esta abordagem estão ilustrados na Figura 2.

$$\sum_{j=1}^N x_j \alpha_j + b = 0 \quad (2)$$

$$\tau_i = -b / \left(\sum_{j=1}^N x_j \alpha_j \right) \quad (3)$$

3.4. Abordagem Baseada em Mínimos Quadrados

O objetivo desta abordagem é aproximar os limiares calculados pela abordagem SVM por indivíduo (Seção 3.3) de forma a reduzir o custo computacional.

O limiar é calculado encontrando a raiz de uma reta no espaço de respostas de identificação e classes. É atribuído o valor +1 à classe de uma resposta de identificação se esta resposta é referente ao indivíduo. Caso contrário, é atribuído o valor -1. Neste contexto, seja Y o vetor de classes e $X_{2 \times T}$ uma matriz onde a primeira coluna é preenchida com as respostas de identificação do i -ésimo modelo PLS e a segunda coluna preenchida com valores 1. A segunda coluna de X representa o deslocamento de uma determinada resposta com relação à origem do espaço e está associada com o peso da

amostra no treinamento. O objetivo é calcular uma reta na forma $\beta x + c = y$ que melhor separe as classes +1 e -1. Uma vez conhecidos β e c , o limiar é dado por $\tau = -c/\beta$. $[\beta, c]$ podem ser calculados resolvendo a Equação 4 por meio de mínimos quadrados conforme a Equação 5.

$$\arg \min_{\beta, c} f(\beta, c) = \sum_{i=1}^T (\beta X_i + c - Y_i)^2 \quad (4)$$

$$[\beta, c]^t = (X^t X)^{-1} X^t Y \quad (5)$$

Como mínimos quadrados pode escolher parâmetros que favoreçam uma classe que contenha mais amostras, a subamostragem é utilizada escolhendo n respostas incorretas (que não se referem ao indivíduo sendo testado) mais próximas de +1 e n respostas corretas são amostradas onde o valor de n é fixado para o mínimo entre o número de respostas corretas e incorretas. Ao utilizar subamostragem, o objetivo é obter a reta que melhor separe regiões de fronteira entre as classes +1 e -1, similar ao que é feito na abordagem que utiliza SVM para cada indivíduo da galeria.

3.5. Abordagem Baseada em Desigualdade

O objetivo desta abordagem é detectar indivíduos desconhecidos estimando um modelo probabilístico para respostas que não correspondem a um indivíduo. Esta abordagem difere daquelas baseadas no critério de máxima margem visto que esta considera a distribuição das respostas.

O limiar de decisão é calculado para respostas de cada indivíduo $r_i, i = [1, 2, \dots, n]$ utilizando a desigualdade de Samuelson [Barnett and Dragomir 2008] como mostrado na Equação 6. A média estimada é representada por μ e σ representa o desvio padrão estimado. Dada a hipótese de que respostas do conjunto de teste demonstrarão aproximadamente o mesmo padrão de respostas do conjunto de validação, o limiar é escolhido como sendo $\mu + \sigma\sqrt{n-1}$.

$$\mu - \sigma\sqrt{n-1} \leq r_i \leq \mu + \sigma\sqrt{n-1} \quad (6)$$

Alguns modelos probabilísticos paramétricos foram testados¹. No entanto, ao utilizar desigualdades ao invés de distribuições paramétricas, esta abordagem se torna independente de qualquer que seja a distribuição presente nos dados, mesmo que as respostas r_i possam ficar distantes do limiar comparado com o limiar calculado por meio de distribuição paramétricas.

Nos experimentos da Seção 4, será mostrado que nem sempre é viável utilizar a desigualdade de Samuelson. De fato, o termo $\sqrt{n-1}$ da Equação 6 pode ficar acima da ordem de valores representados pelo desvio padrão σ . Isto faz com que pequenas diferenças entre desvios estimados entre indivíduos façam com que alguns indivíduos sejam mais considerados durante o teste do que os demais. Nestes casos (que ocorrem quando o número de amostras por indivíduo é elevado), podemos utilizar a desigualdade de Chebyshev.

¹Foram testadas as distribuições Normal, de Gumbel e Valor Extremo Generalizado estimadas utilizando respostas de validação e avaliadas utilizando testes χ^2 nas respostas de validação e teste.

4. Experimentos

As abordagens propostas na Seção 3 são avaliadas nas bases de dados FRGC versão 1 [Phillips et al. 2005] e Pubfig83 [Pinto et al. 2011]. A descrição das bases de dados, experimentos e configuração de parâmetros estão na Seção 4.1. A Seção 4.2 avalia as abordagens em termos de Característica de Operação do Receptor, ou *Receiver Operating Characteristic* (ROC). A Seção 4.3 avalia a influência do número de dimensões, o único parâmetro de identificação de faces necessário. Uma comparação quanto ao número de indivíduos registrados na galeria está na Seção 4.4.

4.1. Configuração Experimental

Esta Seção descreve experimentos, bases de dados e configuração de parâmetros.

FRGC. A base de dados FRGC é composta por 152 indivíduos e seis experimentos envolvendo imagens com informação de profundidade, alta resolução, capturadas sob controle de iluminação ou não. Este trabalho considera os Experimentos 2 e 4 da base de dados FRGC uma vez que os experimentos 3, 4 e 6 são relativos a imagens de profundidade e o Experimento 1 é considerado o mais fácil dentre os demais. O Experimento 2 consiste em quatro amostras de imagens de face por indivíduo capturadas sob condições de iluminação controlada e várias imagens de teste capturadas sob as mesmas condições. O Experimento 4 é considerada a mais desafiadora desta base de dados e consiste em uma amostra de imagem de face por indivíduo capturada sob iluminação controlada e várias imagens de teste capturadas sem controle de iluminação. As imagens foram dimensionadas para 138 pixels de largura por 160 pixels de altura. Os descritores de características utilizados foram HOG, LBP, cor média e filtro de Gabor de acordo com os parâmetros utilizados em [Schwartz et al. 2012]. O vetor que descreve cada imagem possui 162.172 dimensões.

Amostras Extras. Visto que o Experimento 4 da base de dados FRGC disponibiliza apenas uma amostra por indivíduo para treinamento, amostras adicionais são usadas de [Schwartz et al. 2012]. Estas amostras adicionais são necessárias para construir o conjunto de validação, descrito no início da Seção 3. Tais amostras foram obtidas interpolando posição e cor de cada amostra de face de treinamento com uma “face média”. A “face média” é obtida calculando a imagem média de várias amostras de treinamento na FRGC. A amostra original é interpolada 20%, 30% e 40% em direção a “face média”, resultando em três amostras sintetizadas adicionais.

Pubfig83. A base de dados Pubfig83 é composta por 830 imagens de 83 celebridades baixadas pela internet. São disponibilizadas 100 amostras por indivíduo. Os experimentos foram feitos seguindo o protocolo em [Carlos et al. 2013] que consiste em dividir amostras de indivíduos aleatoriamente em 90 para treinamento e 10 para teste. Este processo é repetido 10 vezes e a média e desvio da Área Sob a curva ROC, ou *Área Under the ROC Curve* (AUC), são calculados. A descrição das imagens consiste em dimensionar as imagens para 100 pixels de largura por 100 pixels de altura. Os descritores de características utilizados foram HOG, LBP e cor média com os mesmos parâmetros utilizados no trabalho em [Schwartz et al. 2010] resultando em um vetor de descrição com 6.039 dimensões.

Treinamento e Teste. Os indivíduos de cada base de dados são divididos em dois conjuntos complementares: *registrado* e *não-registrado*. Apenas indivíduos no conjunto re-

gistrado são utilizados como galeria para a identificação de faces. O teste consiste em avaliar se as abordagens determinaram corretamente se uma amostra de teste pertence ao grupo registrado ou não-registrado. A acurácia da identificação das faces não é comparada pois esta foi previamente discutida em [Schwartz et al. 2012]. Na base de dados Pubfig83 foram separados 8 indivíduos para compor o conjunto independente descrito no início da Seção 3.

Identificação de Faces Apenas o número de dimensões para cada modelo PLS é necessário para a abordagem de identificação de faces. O número de dimensões é escolhido empiricamente conforme descrito na Seção 4.3. Cerca de dez dimensões são suficientes para o Experimento 4 da FRGC. Duas dimensões são usadas no Experimento 2 e vinte na base de dados Pubfig83.

SVM. Foram utilizadas funções *kernel* radial e polinomial (assim como discutido na Seção 3.2) para a abordagem que utiliza SVM considerando todos os indivíduos. O parâmetro σ , relacionado com o raio da função *kernel* radial, foi empiricamente fixado em 0,1, embora alterações neste parâmetro não alteraram significativamente os resultados. O grau do polinômio foi fixado em 3 visto que valores maiores não melhoram os resultados. O parâmetro de custo foi empiricamente fixado em 0,001 na FRGC e 0,5 para Pubfig83. Para a abordagem de vetores de suporte por indivíduo, foi utilizado *kernel* linear com função de custo empiricamente fixado em 1.

Abordagens Baseadas em Limiares. Para calcular curvas ROC para abordagens que utilizam um limiar τ para cada indivíduo (abordagens baseadas em mínimos quadrados, SVM por indivíduo e desigualdade), a menor diferença entre cada τ e sua respectiva resposta é utilizada como resposta da abordagem.

4.2. Receiver Operating Characteristic (ROC)

Curvas ROC para os Experimentos 2 e 4 do FRGC são mostradas nas Figuras 3(a) e 3(b), respectivamente. Embora o Experimento 2 seja o mais fácil, as abordagens baseadas em conjunto independente e SVM considerando todos os indivíduos com função *kernel* polinomial não apresentam alta acurácia. A abordagem SVM com *kernel* radial apresenta acurácia razoável no Experimento 2, porém, este apresenta baixa acurácia no Experimento 4. Estas três abordagens se aproximam de respostas aleatórias representadas em curvas ROC pela reta de identidade. Isto acontece uma vez que respostas do modelo de conjunto independente estão distribuídas quase que da mesma forma que de indivíduos na galeria de faces. As demais abordagens apresentam aproximadamente a mesma acurácia.

A intuição sobre a baixa acurácia da abordagem SVM considerando todos os indivíduos é exemplificada na Figura 4: a abordagem retorna valores próximos, representados pelas curvas de nível, para indivíduos conhecidos ou não. A Figura 4 também demonstra a razão das abordagens que consideram indivíduos separadamente apresentarem melhores resultados. A separação entre respostas de indivíduos conhecidos e desconhecidos é maior, para os dois indivíduos mostrados na Figura 4, quando se considera cada dimensão separadamente.

A Tabela 4.2 mostra média e desvio padrão da AUC para as 10 partições da base de dados Pubfig83. As linhas estão ordenadas de forma decrescente, de cima para baixo, de acordo com a média de AUC. Como pode ser observado na Tabela 4.2, a abordagem baseada na desigualdade de Samuelson não apresenta bons resultados. Isto ocorre devido

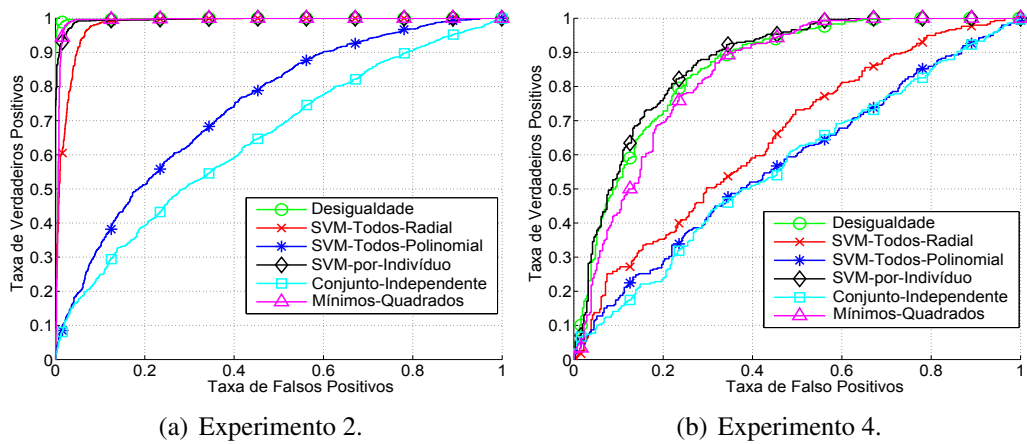


Figura 3. Curvas ROC para várias abordagens avaliadas na base de dados FRGC.

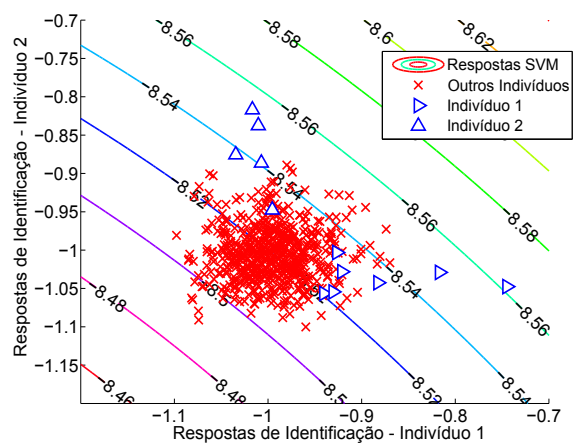


Figura 4. Curvas de resposta do classificador de vetores de suporte, com função radial, para o indivíduo 1 (triângulos apontados para direita) e para o indivíduo 2 (triângulos apontados para cima). Respostas referentes ao Experimento 4 da base de dados FRGC. Marcas cruzadas vermelhas representam respostas de outros indivíduos.

Abordagem	Média (AUC)	Desvio Padrão (AUC)
Mínimos Quadrados	0,792	0,027
SVM por Indivíduo	0,790	0,029
Desigualdade de Chebyshev	0,755	0,033
SVM Radial	0,667	0,035
Conjunto Independente	0,567	0,037
SVM Polinomial	0,490	0,035
Desigualdade de Samuelson	0,484	0,035

Tabela 1. Média e desvio padrão da Área Sob a Curva ROC (AUC) para as 10 partições de treinamento e teste da base de dados Pubfig83.

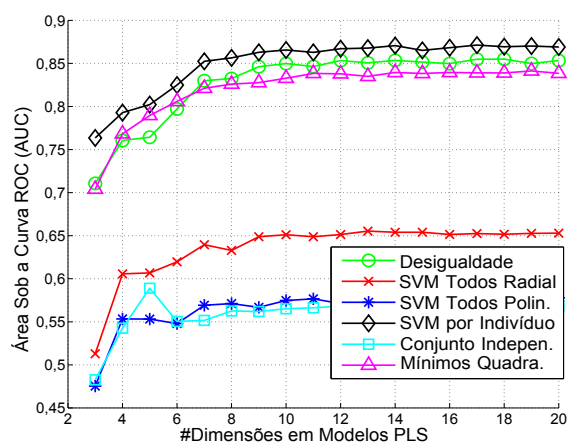


Figura 5. Área Sob a Curva ROC (AUC) para o Experimento 4, FRGC, por número e dimensões nos modelos PLS (de 3 a 20 dimensões).

a grandes diferenças entre limiares calculados por meio da Equação 6. Neste caso, o termo $\sqrt{n-1}$ da Equação 6 fica próximo de 10 enquanto que as variações de desvio padrão σ não são superiores a 0,1. Dessa forma, pequenas diferenças entre o desvio padrão no cálculo de um indivíduo e demais indivíduos são amplificadas fazendo com que apenas alguns limiares sejam utilizados no teste. No entanto, ao utilizar a desigualdade de Chebyshev ao invés da desigualdade de Samuelson, obteve-se resultados similares a aqueles observados na base de dados FRGC.

4.3. Número de Dimensões nos Modelos PLS

A Figura 5 mostra AUC para várias curvas obtidas variando o número de dimensões nos modelos PLS no Experimento 4. A variação no número de dimensões entre 3 e 20 não altera significativamente a acurácia observada na Figura 3(a) para o Experimento 2. Para todas as abordagens, um número de dimensões maior do que aproximadamente nove não aumenta significativamente a acurácia. Isto é esperado uma vez que os autores em [Schwartz et al. 2012] não observaram grandes variações na acurácia quando variam o número de dimensões entre 13 e 21.

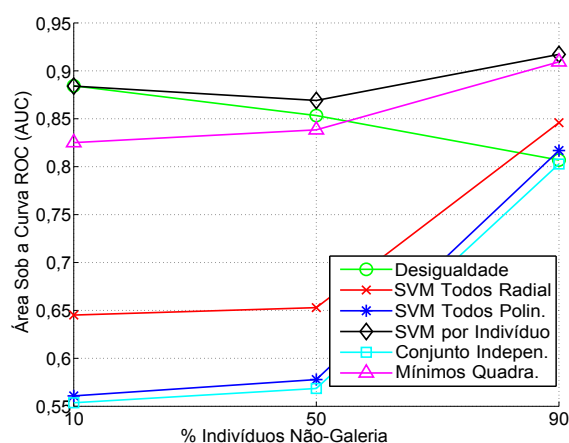


Figura 6. Área Sob a Curva ROC (AUC) para a base de dados FRGC, Experimento 4, quando 10%, 50% e 90% dos 152 indivíduos (arredondando para cima) estão no conjunto *não-galeria*.

4.4. Número de Indivíduos na Galeria

A Figura 6 mostra AUC para várias curvas obtidas variando o número de indivíduos registrados na galeria de faces no Experimento 4. O número de indivíduos registrados na galeria não altera significativamente a acurácia observada na Figura 3(a) para o Experimento 2. Como pode ser observado na Figura 6, todas as abordagens baseadas no conjunto de validação, exceto a abordagem baseada na desigualdade de Samuelson, apresentam maior acurácia quando há 15 indivíduos na galeria. Isto ocorre uma vez que o número de respostas de indivíduos para serem classificados é menor, o que reduz a chance de que estas abordagens classifiquem incorretamente um indivíduo como conhecido. Para a abordagem baseada em conjunto independente, a justificativa do aumento de acurácia para menor número de indivíduos na galeria ocorre devido a facilidade de se encontrar características comuns entre estes indivíduos. Se apenas um indivíduo estiver registrado na galeria, o modelo de conjunto independente tende a funcionar de forma semelhante a abordagens de verificação como em [Guo et al. 2011].

A acurácia da abordagem baseada na desigualdade de Samuelson depende do número de respostas que não são do indivíduo sendo considerado no conjunto de validação para estimar média e variância. Dessa forma, poucas amostras de indivíduos na galeria resultam em média e variância estimadas distantes da média e variância real dos dados. Quando o número de amostras de indivíduos aumenta, a AUC para a abordagem baseada na desigualdade de Samuelson se aproxima dos melhores resultados alcançados pela abordagem SVM por indivíduo e mínimos quadrados.

Aumentando o número de indivíduos de 76 para 136 não altera a AUC significativamente. A abordagem SVM por indivíduo e desigualdade de Samuelson aumentam a acurácia, de 76 para 136 indivíduos na galeria, enquanto que este possui maior acréscimo (0,04 aproximadamente). As demais abordagens apresentam menor acurácia (decréscimo de 0,02 na AUC, aproximadamente).

5. Conclusões

Este trabalho propôs cinco abordagens para determinar se uma imagem de face de teste corresponde ou não a um indivíduo na galeria de faces conhecidas. As abordagens foram avaliadas nas bases de dados FRGC e PubFig83. Três destas abordagens, uma baseada em SVM, outra em modelos estatísticos e a terceira em mínimos quadrados, mostraram melhores resultados. As abordagens baseadas em SVM por indivíduo e mínimos quadrados apresentam acurácia considerável para poucos ou vários indivíduos registrados na galeria. Para a abordagem baseada em modelos estatísticos, foi mostrado um caso onde a abordagem não funcionaria corretamente. Para este caso, foi proposta uma correção utilizando desigualdade de Chebyshev ao invés da desigualdade de Samuelson. Os baixos valores de acurácia das demais abordagens foram explicadas pela dificuldade em encontrar características comuns a vários indivíduos registrados na galeria.

Referências

- Ahonen, T., Hadid, A., and Pietikäinen, M. (2004). Face recognition with local binary patterns. In *Computer Vision-ECCV 2004*, pages 469–481. Springer.
- Albiol, A., Monzo, D., Martin, A., Sastre, J., and Albiol, A. (2008). Face recognition using hog+ebgm. *Pattern Recognition Letters*, 29(10):1537–1543.
- Barnett, N. S. and Dragomir, S. S. (2008). *Advances in Inequalities from Probability Theory and Statistics*. Nova Publishers.
- Cao, Z., Yin, Q., Tang, X., and Sun, J. (2010). Face recognition with learning-based descriptor. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference*, pages 2707–2714. IEEE.
- Carlos, G. P., Pedrini, H., and Schwartz, W. R. (2013). Fast and scalable enrollment for face identification based on partial least squares. In *Automatic Face and Gesture Recognition (FG), 2013 10th IEEE International Conference and Workshops*, pages 1–8.
- Chellappa, R., Sinha, P., and Phillips, P. J. (2010). Face recognition by computers and humans. *Computer*, 43(2):46–55.
- Etemad, K. and Chellappa, R. (1997). Discriminant analysis for recognition of human face images. *JOSA A*, 14(8):1724–1733.
- Guo, H., Robson Schwartz, W., and Davis, L. S. (2011). Face verification using large feature sets and one shot similarity. In *Biometrics (IJCB), 2011 International Joint Conference*, pages 1–8. IEEE.
- Kumar, N., Berg, A. C., Belhumeur, P. N., and Nayar, S. K. (2009). Attribute and simile classifiers for face verification. In *Computer Vision, 2009 IEEE 12th International Conference*, pages 365–372. IEEE.
- Li, F. and Wechsler, H. (2005). Open set face recognition using transduction. *Pattern Analysis and Machine Intelligence, IEEE Transactions*, 27(11):1686–1697.
- Luo, J., Ma, Y., Takikawa, E., Lao, S., Kawade, M., and Lu, B.-L. (2007). Person-specific sift features for face recognition. In *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference*, volume 2, pages 593–596. IEEE.

- Phillips, P. J., Flynn, P. J., Scruggs, T., Bowyer, K. W., Chang, J., Hoffman, K., Marques, J., Min, J., and Worek, W. (2005). Overview of the face recognition grand challenge. In *Computer vision and pattern recognition (CVPR), 2005. IEEE computer society conference*, volume 1, pages 947–954. IEEE.
- Pinto, N., Stone, Z., Zickler, T., and Cox, D. (2011). Scaling up biologically-inspired computer vision: A case study in unconstrained face recognition on facebook. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2011 IEEE Computer Society Conference*, pages 35–42. IEEE.
- Schölkopf, B., Platt, J. C., Shawe-Taylor, J., Smola, A. J., and Williamson, R. C. (2001). Estimating the support of a high-dimensional distribution. *Neural computation*, 13(7):1443–1471.
- Schwartz, W. R., Guo, H., Choi, J., and Davis, L. S. (2012). Face identification using large feature sets. *Image Processing, IEEE Transactions*, 21(4):2245–2255.
- Schwartz, W. R., Guo, H., and Davis, L. S. (2010). A robust and scalable approach to face identification. In *Computer Vision–ECCV 2010*, pages 476–489. Springer.
- Stallkamp, J., Ekenel, H. K., and Stiefelhagen, R. (2007). Video-based face recognition on real-world data. In *Computer Vision–ICCV 2007. IEEE 11th International Conference*, pages 1–8. IEEE.
- Tan, X. and Triggs, B. (2007). Fusing gabor and lbp feature sets for kernel-based face recognition. In *Analysis and Modeling of Faces and Gestures*, pages 235–249. Springer.
- Turk, M. and Pentland, A. (1991). Eigenfaces for recognition. *Journal of cognitive neuroscience*, 3(1):71–86.
- Wechsler, H. (2007). *Reliable face recognition methods: system design, implementation and evaluation*, volume 7. Springer.
- Wolf, L., Hassner, T., and Maoz, I. (2011a). Face recognition in unconstrained videos with matched background similarity. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference*, pages 529–534. IEEE.
- Wolf, L., Hassner, T., and Maoz, I. (2011b). Face recognition in unconstrained videos with matched background similarity. pages 529–534.
- Wolf, L., Hassner, T., Taigman, Y., et al. (2008). Descriptor based methods in the wild. In *Workshop on Faces in 'Real-Life' Images: Detection, Alignment, and Recognition*, pages 1–14.
- Zhang, B., Shan, S., Chen, X., and Gao, W. (2007). Histogram of gabor phase patterns (hgpp): A novel object representation approach for face recognition. *Image Processing, IEEE Transactions*, 16(1):57–68.
- Zhou, H. and Sadka, A. H. (2011). Combining perceptual features with diffusion distance for face recognition. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions*, 41(5):577–588.
- Zou, J., Ji, Q., and Nagy, G. (2007). A comparative study of local matching approach for face recognition. *Image Processing, IEEE Transactions*, 16(10):2617–2628.