

Adaptação de Domínio para Reconhecimento de Padrões Visuais

Aluno: João Renato Ribeiro Manesco

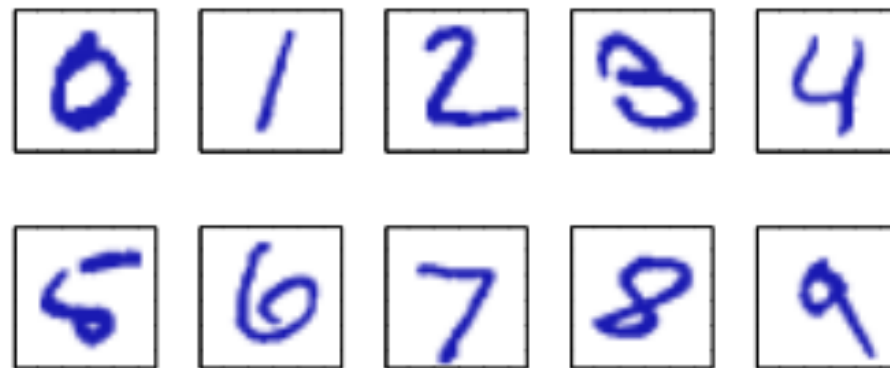
Orientador: Prof. Assoc. Aparecido Nilceu Marana

Trabalho de Conclusão de Curso
Bacharelado em Ciência da Computação
Faculdade de Ciências - UNESP - Bauru

Introdução

- Reconhecimento de Padrões (RP) visa descobrir automaticamente padrões em conjuntos de dados;
- Em RP os dados de treino (fonte) e de teste (alvo) costumam ter distribuições similares;
- Quando as distribuições fonte e alvo são distintas surge um problema chamado de deslocamento de domínio;
- Adaptação de domínio (AD) é uma forma de minimizar o problema deslocamento de domínio.

Introdução



Fonte: (BISHOP, 2006)



Fonte: (NETZER et al., 2011).

Introdução

- Exige-se desempenhos cada vez melhores em sistemas de RP
 - Principalmente em aplicações reais, cujos domínios fonte e alvo tendem a ser bastante distintos;
- Neste trabalho foram avaliadas quatro métodos de AD:
 - CORAL.
 - TCA
 - JDA
 - TKL
- Foi desenvolvido um aplicativo para autenticação de faces em dispositivos móveis com e sem adaptação de domínio.

Aplicação: Reconhecimento de Objetos

- Detectar e classificar objetos em imagens
- Descritores utilizados
 - SURF
 - DeCAF
- Problema: Imagens de objetos obtidas de diferentes fontes, com diferentes fundos, iluminação e qualidade de imagem.

Aplicação: Identificação Biométrica Facial

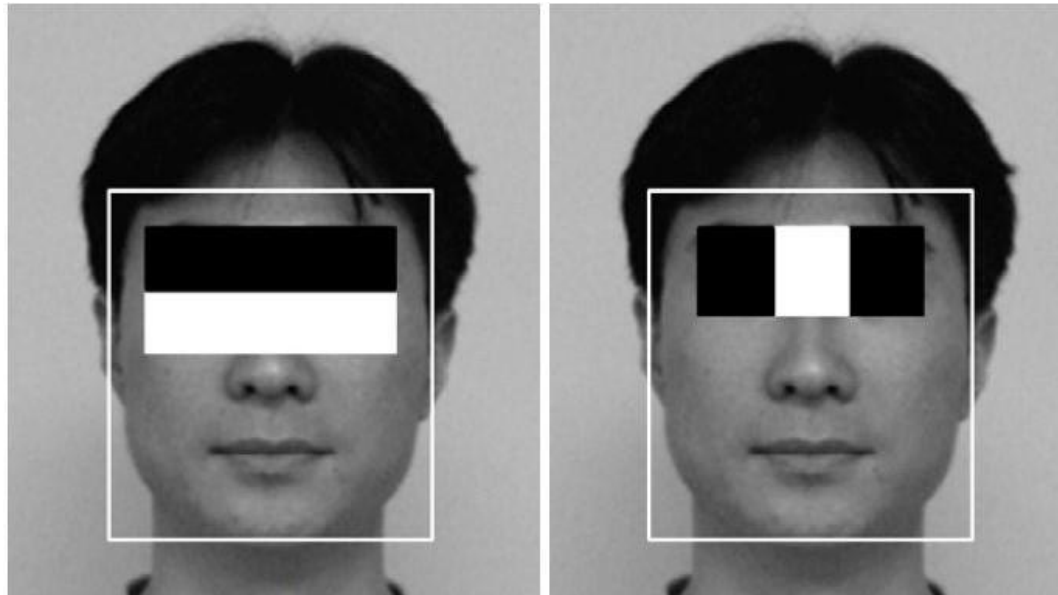
- Pode ser realizada em dois modos:
 - Autenticação (1:1)
 - Reconhecimento (1:N)
- Exige três etapas:
 - Detecção da face
 - Extração de características
 - Comparação (matching)
- Problema: podem existir variações de pose, iluminação, expressão, idade e oclusão.

Detecção de Faces

- Encontrar a região da imagem em que há uma face
- Neste trabalho, foram utilizados dois métodos:
 - Viola-Jones (baseado em filtros de Haar)
 - MTCNN (baseado em um conjunto de redes convolucionais em cascata)

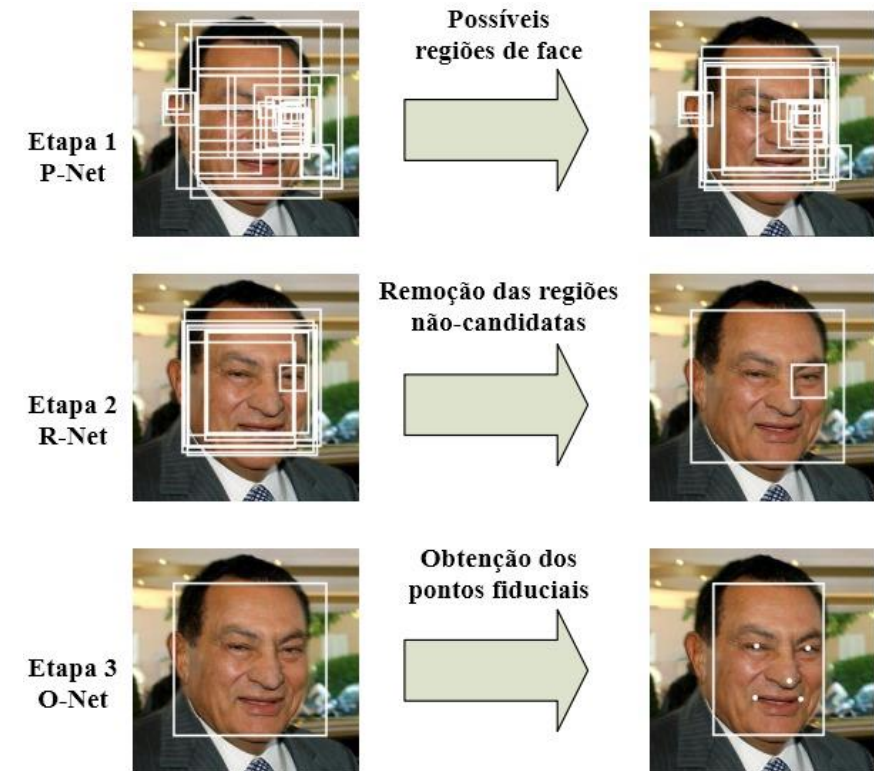
Detecção de Faces

VIOLA-JONES



Fonte: (JAIN; ROSS; NANDAKUMAR, 2011).

MTCNN



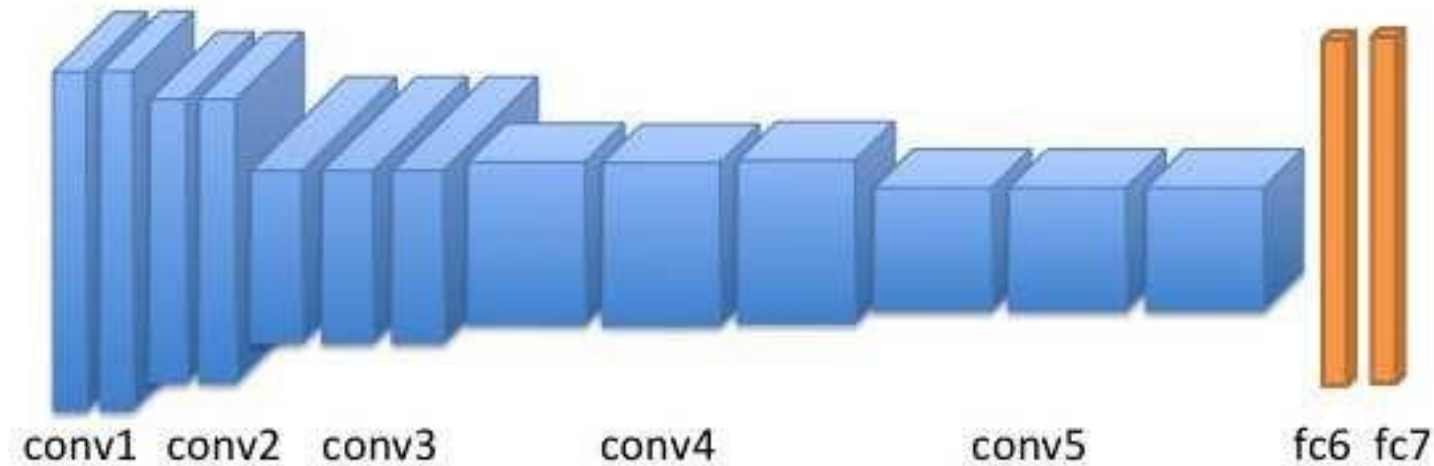
Fonte: (ZHANG et al., 2016).

Extração de Características

- Objetivos:
 - Reduzir a dimensionalidade dos dados
 - Encontrar melhores representações para a tarefa de classificação

Extração de Características Faciais

- Neste trabalho, a extração de características faciais foi feita por meio da rede VGG Face (as características foram extraídas da camada fc7).



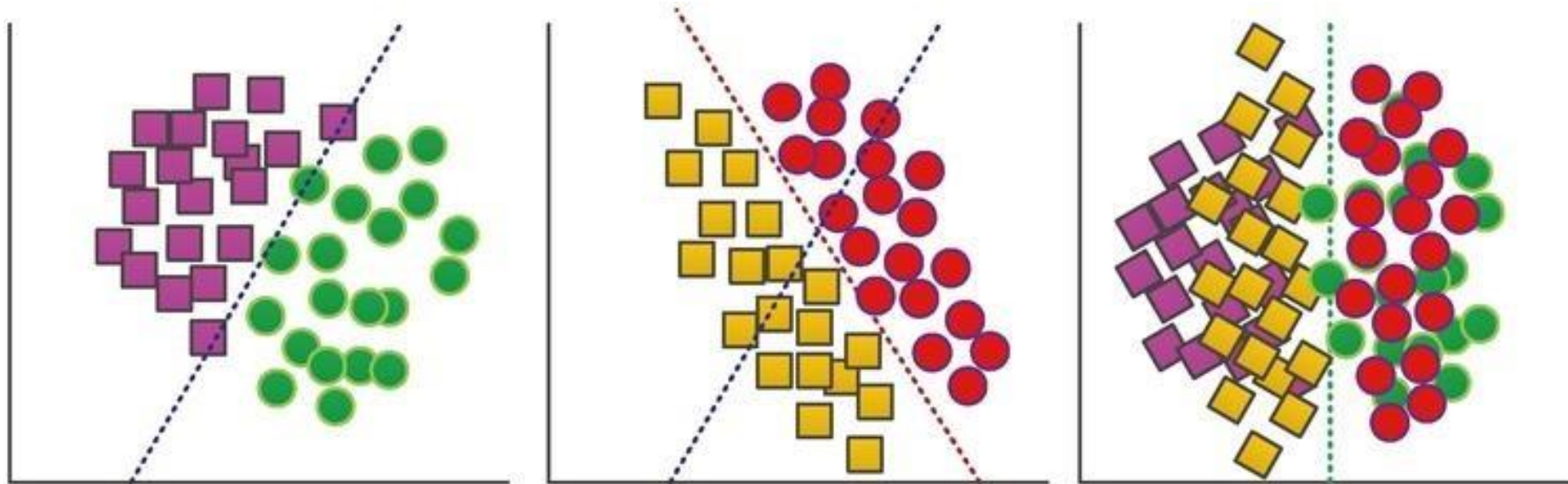
Fonte: (NAKADA; WANG; TERZOPOULOS, 2017).

Adaptação de Domínio

- Deslocamento de domínio ocorre quando as distribuições dos domínios fonte (treino) e alvo (teste) são diferentes.
- Pode-se utilizar técnicas de adaptação de domínio (AD) para lidar com o problema do deslocamento de domínio.
- Estratégia para AD:
 - Utilizar um conjunto de dados externo ao domínio fonte visando encontrar uma melhor representação dos dados.
 - É necessário que os dois domínios tratem do mesmo problema.

Adaptação de Domínio

Exemplo de uma situação em que é necessário utilizar adaptação de domínio



Fonte: (CHAI et al., 2018).

Necessidade de Adaptação de Domínio para o Reconhecimento Facial

Exemplos de variações de pose, iluminação e expressão facial



Fonte: Elaborada pelo autor.

Método CORAL

- Hipótese: o impacto do deslocamento de domínio é causado por uma relação linear entre as covariâncias.
- Objetivo: alinhar a covariância dos dados, minimizando a diferença.
- Possui por duas etapas:
 - Esbranquiçamento: “descorrelaciona” os dados do domínio fonte por meio da pseudo-inversa da matriz de covariância.
 - Recoloração: “re-correlaciona” os dados com a covariância do domínio alvo.
- Problema: Nem sempre a relação linear da covariância dos dados é a causa do deslocamento de domínio.

Método TKL

- Objetivo: obter um Kernel de domínio invariante, utilizado para alimentar uma SVM.
- O kernel deve minimizar a variação da distribuição entre os dois domínios.
- Extrapola os autovetores do domínio fonte, através dos autovetores do domínio alvo (Teorema de Mercer).
- Reconstrói o Kernel, em um problema que visa minimizar a distância entre o Kernel reconstruído e o Kernel original, relaxando os autovalores.
- Kernel do domínio fonte em relação ao fonte é utilizados no treino.
- Kernel do domínio alvo em relação ao fonte é utilizado no teste.

Método TCA

- Objetivo: obter uma transformação $\phi(\cdot)$ para um novo espaço de características, onde a métrica MMD entre os domínios é minimizada;
- Utiliza como base a métrica de distância MMD;
- Realiza o cálculo da MMD de forma matricial, para encontrar os parâmetros de uma matriz de transformação W ;
- Torna-se um problema similar ao do PCA (minimização são os m autovetores mais significativos).
 - Consequência: redução de dimensionalidade.

Método JDA

- Generalização do TCA;
- Usa os rótulos das amostras para melhorar a distribuição condicional dos dados.
 - Análise da matriz de MMD é feita para cada classe.
- Método não-supervisionado:
 - Inicia a adaptação de domínio por meio do TCA;
 - Gera pseudo-rótulos por meio de um classificador auxiliar;
 - Utiliza o conjunto de pseudo-rótulos na análise da matriz de MMD.
- Resultado passa a ser refinado iterativamente.

Metodologia

- Os métodos de adaptação de domínio foram comparados com dois métodos sem adaptação: 1NN e SVM;
- Para os métodos baseados em SVM foram utilizados:
 - Parâmetro de regularização $C = 10$;
 - Kernel RBF, com $\sigma = 1.0$.
- O TKL utilizou fator de amortecimento $\zeta = 1.1$;
- Nos métodos TCA e JDA:
 - Fator de regularização: $\mu = 1.0$
 - Número de componentes principais: $m = 20$

Metodologia

- Protocolo:
 - Treino realizado somente com dados do domínio fonte.
 - Teste realizado somente com dados do domínio alvo.
- Métricas utilizadas: acurácia e separabilidade.
- TKL foi avaliado somente em relação a acurácia
 - Não realizava transformação para um novo espaço de características.

Reconhecimento de Objetos

- Base de dados Office-Caltech
 - Imagens de quatro domínios distintos:
 - Dataset Caltech-256;
 - *Website* Amazon;
 - Capturadas com *webcam*;
 - Capturadas com câmara DSLR.
- Descritores SURF e DeCAF
 - Disponibilizados pelos autores da base de dados.

Reconhecimento de Objetos

Exemplos de imagens da base Office-Caltech



Caltech



Amazon



Webcam



DSLR

Reconhecimento Facial

- Base de Dados ARFace:
 - Detecção das faces → Viola-Jones
 - Quatro domínios:
 - (N): Neutro;
 - (I): Variação de iluminação;
 - (C): Oclusão da boca;
 - (O): Oclusão da região ocular.
- Base de Dados FRGC:
 - Detecção das faces → MTCNN
 - Dois domínios:
 - Ambiente controlado
 - Ambiente não-controlado.

Reconhecimento Facial

Exemplos de imagens da base ARFace



Fonte: (ZHOU;ZHANG, 2019).

Reconhecimento Facial

Exemplos de imagens da base de dados FRGC



Fonte: Elaborada pelo autor.

Desenvolvimento

- O trabalho foi desenvolvido em duas etapas:
 - Experimentos envolvendo adaptação de domínio
 - Linguagem *Python*, com as bibliotecas *Numpy*, *Scipy* e *Scikit-learn*;
 - Desenvolvimento de um aplicativo para reconhecimento facial em dispositivos móveis com adaptação de domínio
 - React Native (plataforma *Expo*).

Desenvolvimento

- Viola-Jones: *OpenCV* (função Cascade Classifier);
- MTCNN: pacote do mesmo nome (fornece acesso de alto nível às redes em cascata).
- VGGFace: pacote *keras_vggface* (com pesos pré-treinados).

Desenvolvimento

- TKL e CORAL: implementados utilizando a biblioteca *numpy*, com a criação do pacote *tklpy*;
- TCA e JDA: possuíam implementações públicas em um repositório. Foram realizadas melhorias para que o código pudesse ser executado.

Resultados – Office-Caltech (SURF)

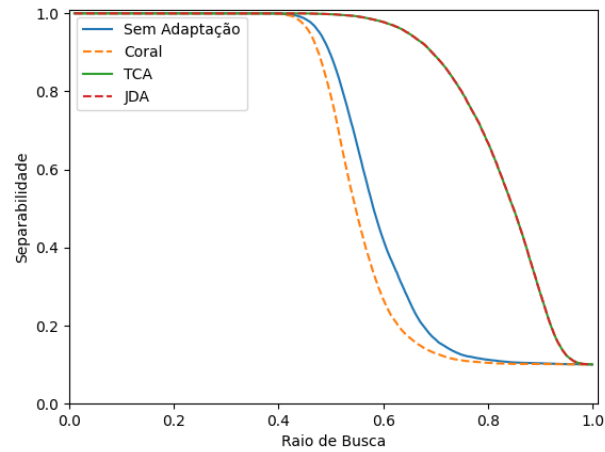
Acurácias - Office-Caltech - Descritor SURF

	$A \rightarrow C$	$C \rightarrow A$	$D \rightarrow W$	$W \rightarrow D$	$C \rightarrow D$	$A \rightarrow W$	$D \rightarrow C$	$W \rightarrow A$
1-NN	27.66%	21.88%	72.00%	60.12%	25.47%	29.83%	26.26%	22.96%
SVM	43.90%	51.98%	62.37%	84.71%	48.40%	38.30%	29.56%	32.15%
Coral 1NN	27.66%	23.69%	63.38%	60.12%	25.47%	29.83%	26.27%	22.96%
Coral SVM	43.90%	52.71%	47.79%	84.71%	48.40%	38.30%	29.56%	32.15%
TKL	45.94%	55.95%	86.10%	84.71%	45.22%	42.03%	39.01%	37.99%
TCA	40.60%	46.65%	88.47%	91.71%	47.77%	37.28%	32.14%	31.62%
JDA	40.60%	46.54%	89.83%	92.35%	47.13%	38.98%	32.05%	32.85%

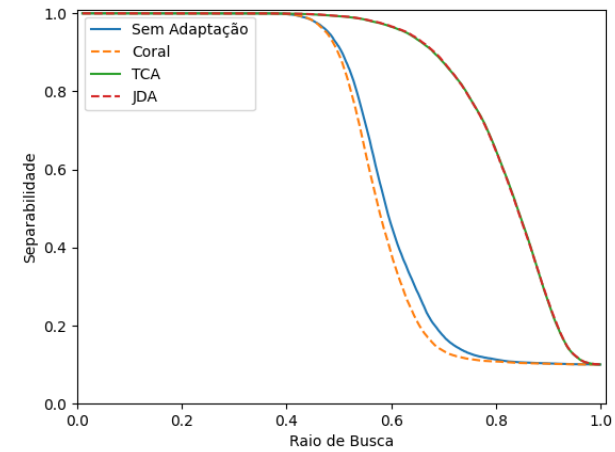
Fonte: Elaborada pelo autor.

Separabilidades – Office-Caltech (SURF)

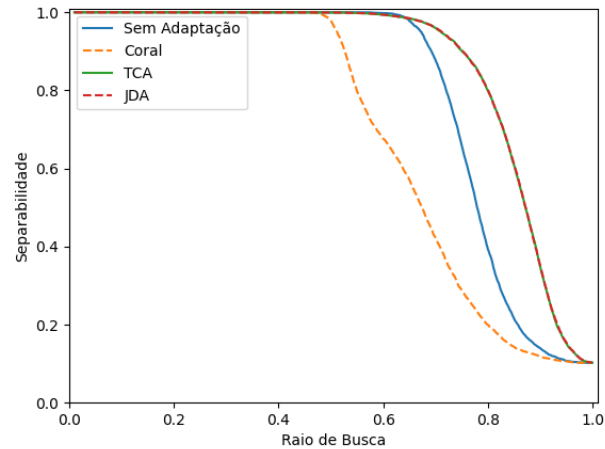
Amazon → Caltech



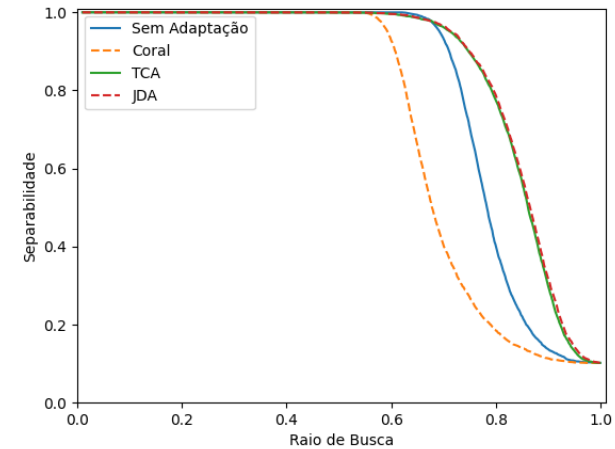
Caltech → Amazon



DSLR → Webcam

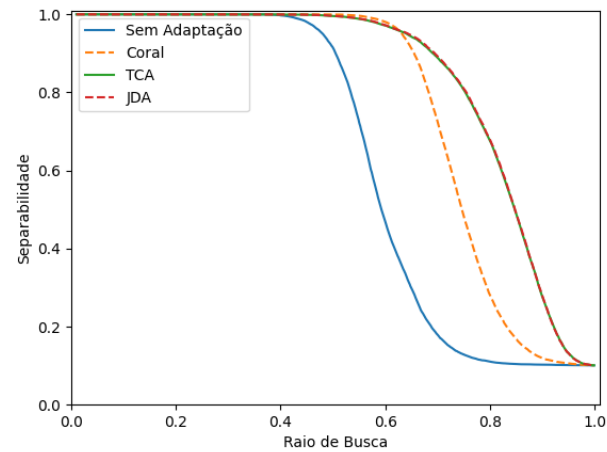


Webcam → DSLR

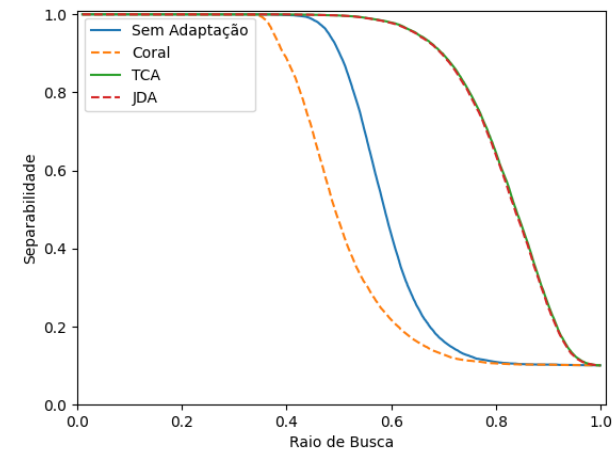


Separabilidades – Office-Caltech (SURF)

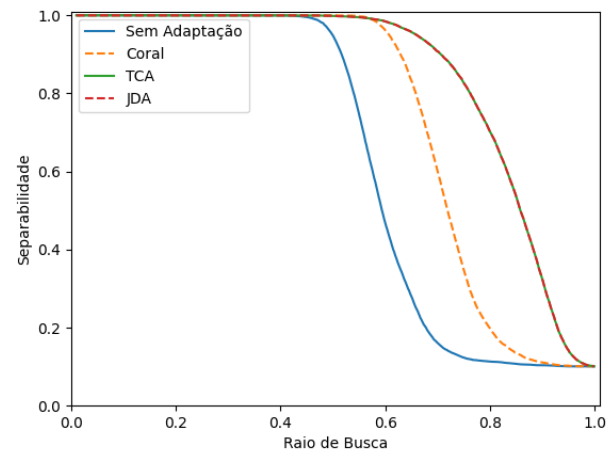
Caltech → DSLR



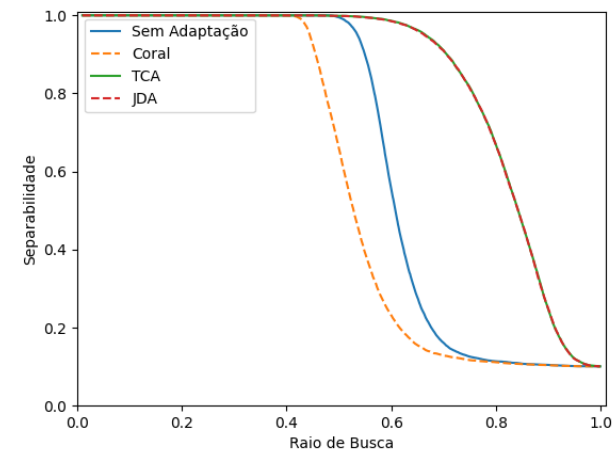
DSLR → Caltech



Amazon → Webcam



Webcam → Amazon



Resultados – Office-Caltech (DeCAF)

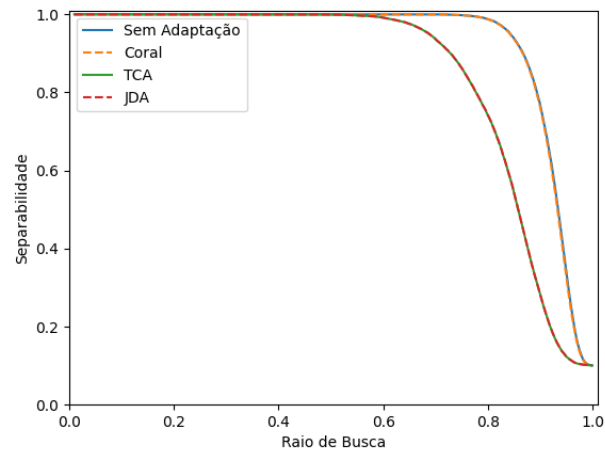
Acurácias - Office-Caltech - Descritor DeCAF

	$A \rightarrow C$	$C \rightarrow A$	$D \rightarrow W$	$W \rightarrow D$	$C \rightarrow D$	$A \rightarrow W$	$D \rightarrow C$	$W \rightarrow A$
1-NN	71.28%	88.54%	99.66%	87.50%	86.24%	74.57%	79.16%	77.13%
SVM	83.43%	91.54%	98.30%	98.73%	87.89%	74.91%	66.25%	72.65%
Coral 1NN	83.70%	89.97%	99.66%	87.50%	86.24%	74.57%	79.16%	77.13%
Coral SVM	81.74%	91.63%	96.61%	98.73%	87.89%	74.91%	58.05%	72.86%
TKL	82.81%	93.63%	97.96%	99.36%	87.89%	76.61%	74.44%	75.57%
TCA	83.79%	89.45%	99.32%	99.36%	85.98%	74.91%	81.30%	80.58%
JDA	83.79%	90.81%	99.32%	99.36%	87.89%	76.94%	83.52%	85.17%

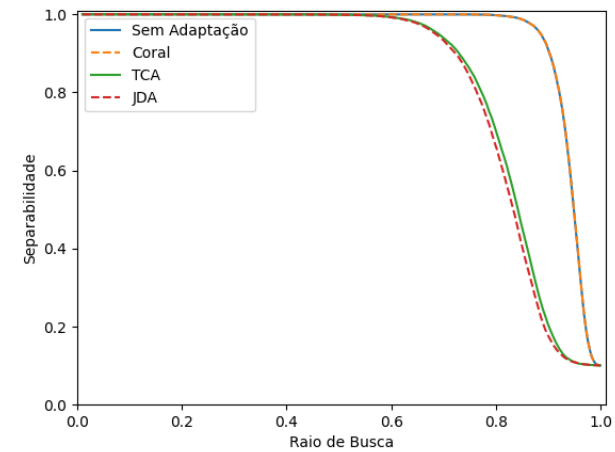
Fonte: Elaborada pelo autor.

Separabilidades – Office-Caltech (DeCAF)

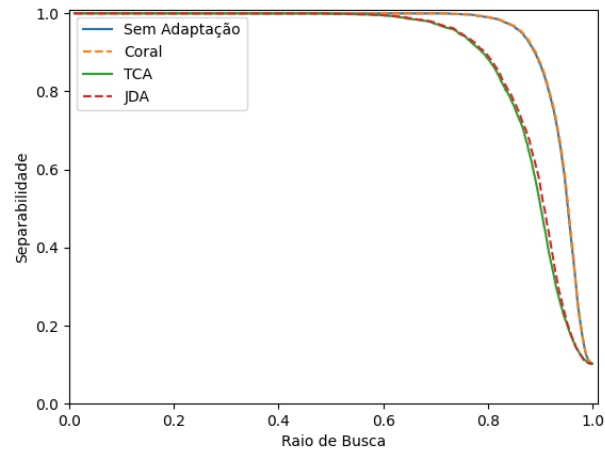
Amazon → Caltech



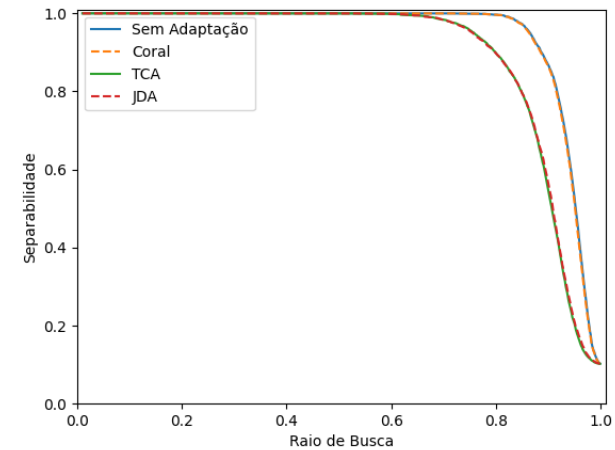
Caltech → Amazon



DSLR → Webcam

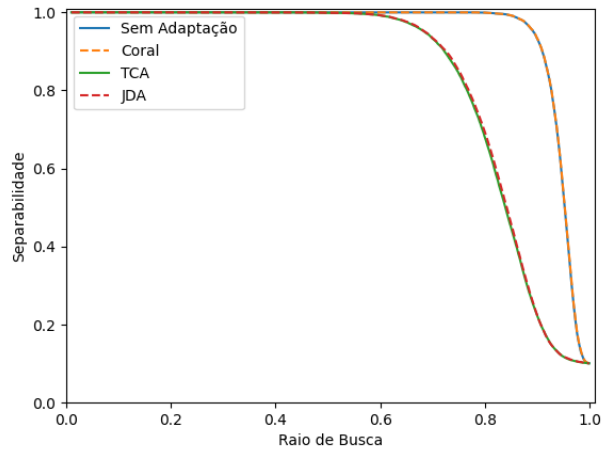


Webcam → DSLR

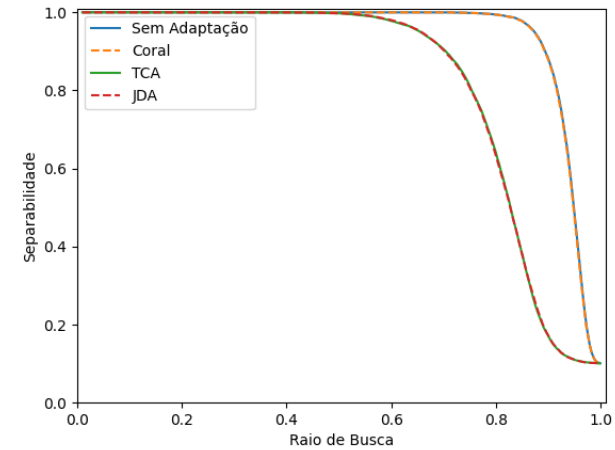


Separabilidades – Office-Caltech (DeCAF)

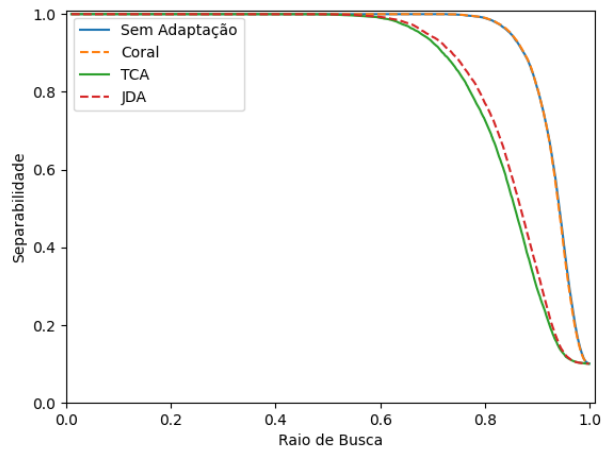
Caltech → DSLR



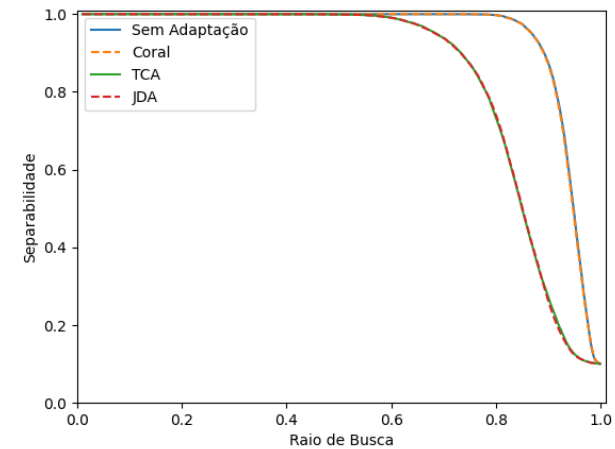
DSLR → Caltech



Amazon → Webcam



Webcam → Amazon



Resultados – Reconhecimento Facial

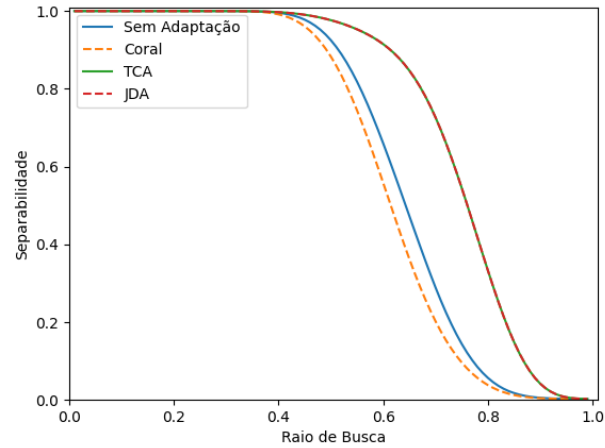
Acurácias obtidas na tarefa de reconhecimento facial

	FRGC	ARFace N → O	ARFace N → C	ARFace N → I
1-NN	75.33%	67.29%	95.89%	99.62%
SVM	76.35%	64.81%	94.52%	100%
Coral 1NN	76.82%	67.29%	95.89%	99.60%
Coral SVM	78.06%	70.41%	96.80%	100%
TKL	82.63%	89.73%	97.71%	99.76%
TCA	88.92%	85.60%	99.09%	100%
JDA	88.92%	84.40%	99.54%	100%

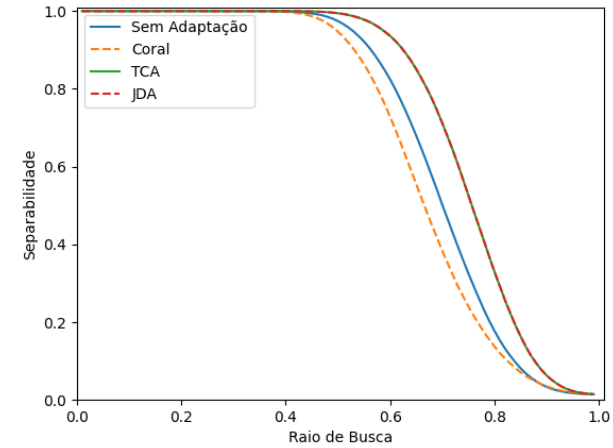
Fonte: Elaborada pelo autor.

Separabilidades – Reconhecimento Facial

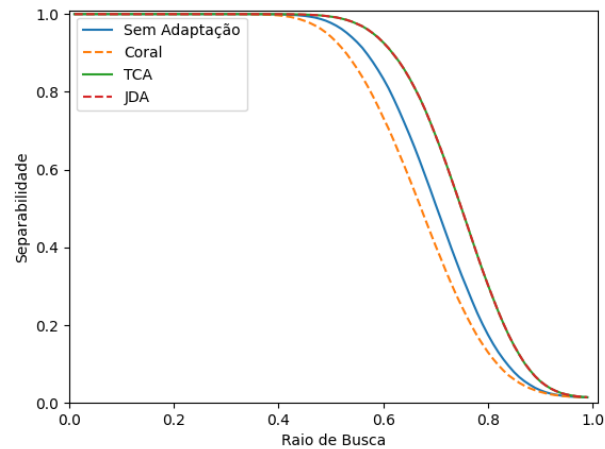
FRGC



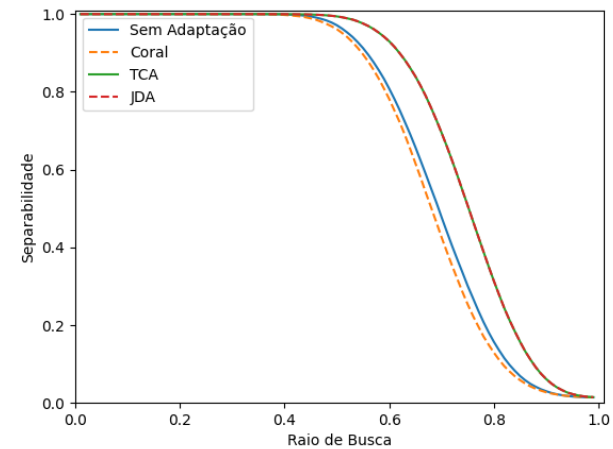
ARFace N→O



ARFace N→C

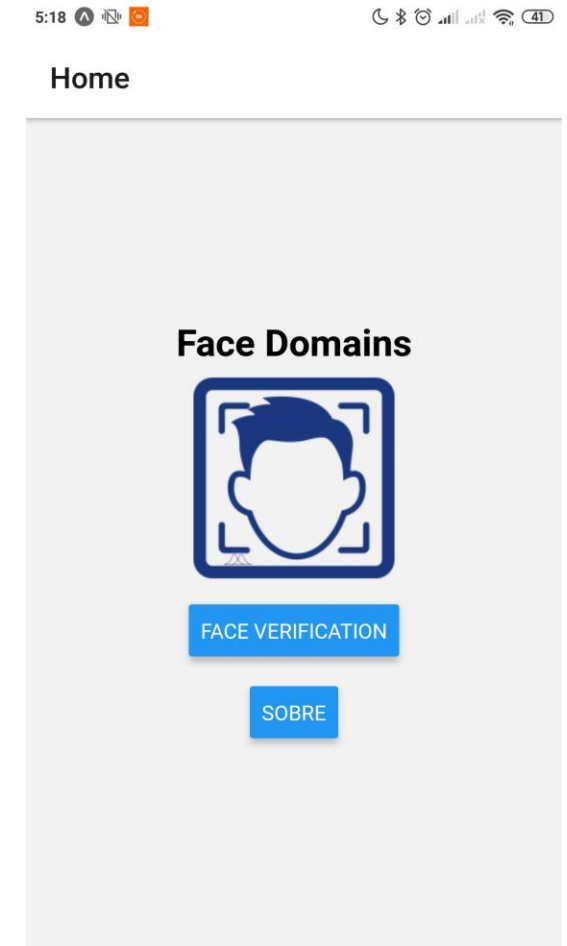


ARFace N→I



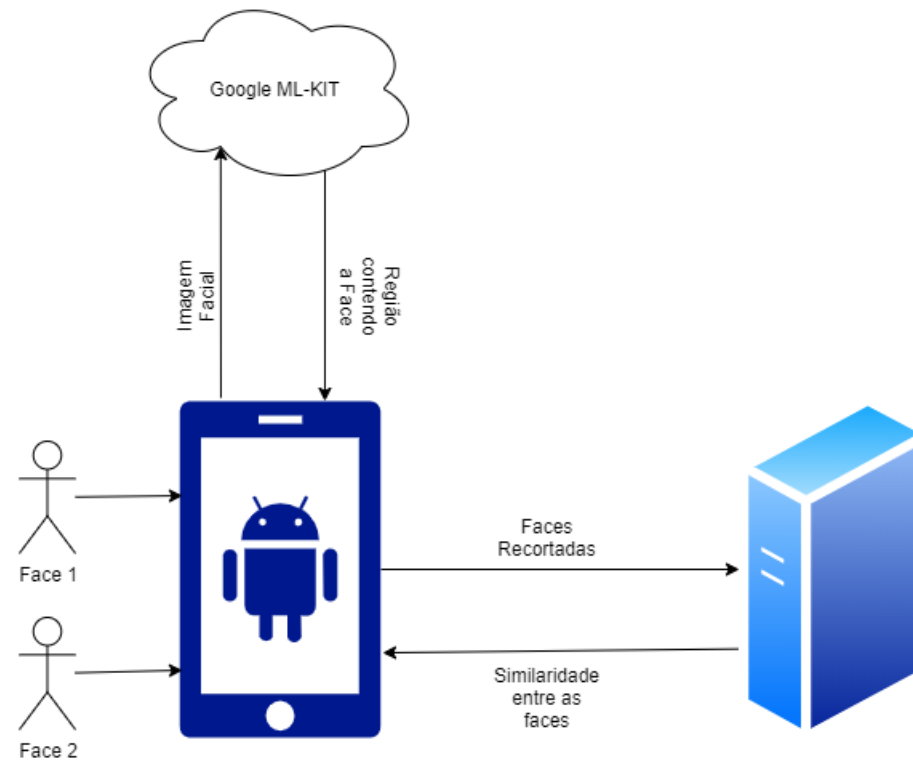
Aplicativo Face Domains

- Desenvolvido utilizando a ferramenta React Native (plataforma *Expo*)
- Opera no modo de autenticação facial;
 - Dadas duas imagens faciais, verifica se são do mesmo indivíduo;
- Detecção e recorte das faces: API do Google ML-Kit;
- Extração de características: realizada em um servidor, por meio de uma API REST.
- Valores padrão de limiar: associado ao Equal Error Rate (calculada via 10 mil comparações genuínas e 10 mil impostoras da base FRGC).



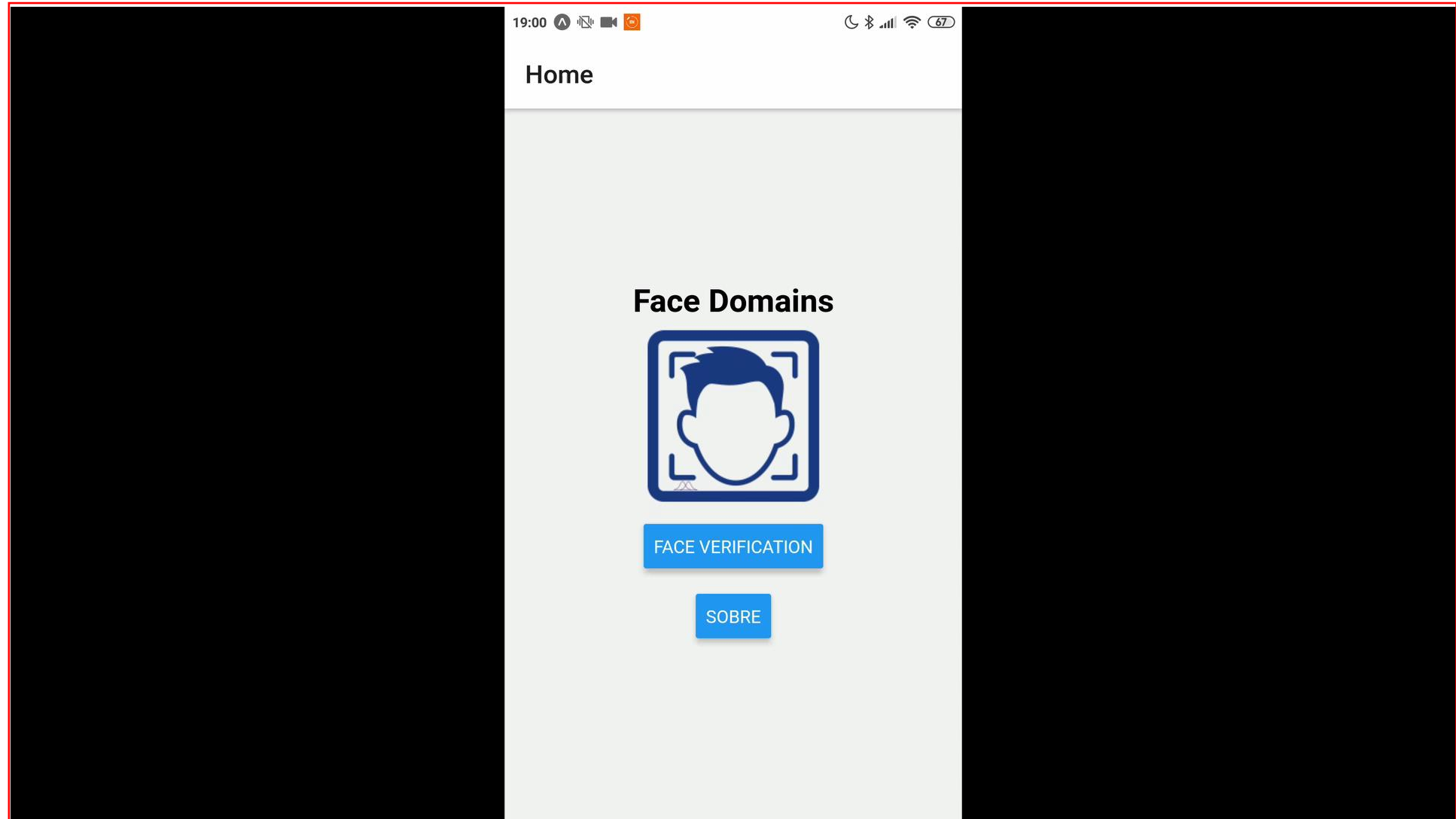
Aplicativo Face Domains

Funcionamento do aplicativo



Fonte: Elaborada pelo autor.

Aplicativo Face Domains - Vídeo



Conclusão

- Métodos de Adaptação de Domínio causam um impacto positivo em tarefas de reconhecimento de padrões visuais.
 - Melhor acurácia e obtenção de espaços mais separáveis.
- Uso de técnicas de adaptação de domínio é viável também em aplicações reais.
 - Pode-se utilizar adaptação de domínio de maneira simples até mesmo fora de ambientes experimentais (vide aplicativo implementado).

Trabalhos Futuros

- Explorar adaptação de domínio em outras tarefas de reconhecimento de padrões visuais;
- Explorar a utilização de técnicas de adaptação de domínio profundas e avaliar os pontos positivos e negativos de sua utilização em relação às técnicas tradicionais.

Referências

- BISHOP, C. M. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Berlin, Heidelberg: Springer-Verlag, 2006. ISBN 0387310738
- CHAI, Xin et al. Multi-subject subspace alignment for non-stationary EEG-based emotion recognition. *Technology and Health Care*, v. 26, n. S1, p. 327-335, 2018.
- JAIN, Anil K.; ROSS, Arun A.; NANDAKUMAR, Karthik. *Introduction to biometrics*. Springer Science & Business Media, 2011.
- NAKADA, Masaki; WANG, Han; TERZOPOULOS, Demetri. AcFR: Active face recognition using convolutional neural networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 2017. p. 35-40.
- NETZER, Yuval et al. *Reading digits in natural images with unsupervised feature learning*. 2011.
- ZHANG, Kaipeng et al. Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters*, v. 23, n. 10, p. 1499-1503, 2016.
- ZHOU, Jianhang; ZHANG, Bob. Collaborative Representation Using Non-Negative Samples for Image Classification. *Sensors*, v. 19, n. 11, p. 2609, 2019.

Agradecimentos

- Agradeço aos meus pais pelo apoio ao longo da faculdade, aos meus amigos e colegas pelo tempo que passamos juntos.
- Agradeço também aos professores por todo aprendizado ao longo da faculdade, em especial ao professor Aparecido Nilceu Marana por me orientar.