

# Estudo de caso de um Sistema de reconhecimento facial utilizando o *OpenFace* para identificação de faces em bancos de imagens

Lucas Ricardo Orso  
Departamento de Ciência da  
Computação  
Universidade do Extremo Sul  
Catarinense - UNESC  
Av. Universitária, 1105  
Criciúma – Santa Catarina  
lucasorso@gmail.com

Sergio Coral  
Departamento de Ciência da  
Computação  
Universidade do Extremo Sul  
Catarinense - UNESC  
Av. Universitária, 1105  
Criciúma – Santa Catarina  
sergiocoral@unesc.net

Mariane de Oliveira Filastro  
Departamento de Matemática  
Universidade do Extremo Sul  
Catarinense - UNESC  
Av. Universitária, 1105  
Criciúma – Santa Catarina  
mariyfilastro@outlook.com

Kristian Madeira  
Departamento de Matemática  
Universidade do Extremo Sul  
Catarinense - UNESC  
Av. Universitária, 1105  
Criciúma – Santa Catarina  
kristian@unesc.net

Carolina Michels  
Departamento de Matemática  
Universidade do Extremo Sul  
Catarinense - UNESC  
Av. Universitária, 1105  
Criciúma – Santa Catarina  
michelsscarol@gmail.com

Paulo João Martins  
Departamento de Ciência da  
Computação  
Universidade do Extremo Sul  
Catarinense - UNESC  
Av. Universitária, 1105  
Criciúma – Santa Catarina  
pjm@unesc.net

## ABSTRACT

*Searching a bigger knowledge about facial recognition, the objective of this work is to evaluate the Support Vector Machines classifier used by the OpenFace library, in the classification of images through the Cross-Validation method. For this were used the Labeled Faces in the Wild image bank, and as image bases, the MS-CELEB-1M and the FEI DATABASE, where with the OpenFace ANN (Artificial Neural Network) the characteristics of each image were extracted, to create a Support Vector Machines classifier. The evaluation was by the cross-validation method. The results confirmed that the use of artificial intelligence for the problems found in researched area is a valid option to be studied and used.*

## RESUMO

*Buscando um maior conhecimento sobre o reconhecimento facial, este artigo tem por objetivo avaliar o classificador Support Vector Machines utilizado pela biblioteca OpenFace, na classificação das imagens por meio do método Cross-Validation. Para isso, utilizou-se o banco de imagens Labeled Faces in the Wild, e como bases de imagens, o MS-CELEB-1M e o FEI DATABASE, onde com a RNA (Rede Neural Artificial) do OpenFace foram extraídas as características de cada imagem, para criar um classificador Support Vector Machines. A avaliação se deu então pelo método de validação cruzada. Os resultados confirmaram que, a utilização da inteligência artificial para os problemas encontrados na área pesquisada é uma opção válida a ser estudada e utilizada.*

## CCS Concepts

• **Computing Methodologies**→**Artificial Intelligence**→**Imagem Recognition.**

## Keywords

Biometria. Reconhecimento facial. *OpenFace*. Inteligência Artificial.

## 1. INTRODUÇÃO

Estudos relacionados a sistemas de reconhecimento facial estão em evidência atualmente, devido a possibilidade desses sistemas em reconhecer rostos a partir de vídeos ou imagens, tornando-se cada vez mais uma realidade a ser utilizada. Ainda, a visão computacional proporciona a construção de um sistema de visão artificial que se assemelha a capacidade do sistema visual humano, no qual, nos últimos anos houve significativos avanços, em razão do desenvolvimento de equipamentos computacionais mais rápidos [7].

A área da visão computacional engloba a análise de imagens, e tem como princípio diferenciar os objetos de interesse, sem a intervenção humana [17]. Visto isso, a visão computacional surgiu para facilitar a resolução de problemas mais complexos e busca imitar a cognição e habilidade do ser humano em tomar decisões de acordo com informações contidas na imagem [17].

Quando se discute sobre visão computacional, normalmente são estipulados dois níveis de abstração: processamento de imagens (baixo nível) e análise de imagens (alto nível) [4, 11]. “O processamento digital de imagens consiste em um conjunto de técnicas para capturar, representar e transformar imagens com o auxílio de computador” [8, p.29]. Isso envolve segmentação de imagem, descrição, representação e reconhecimento de objetos [4].

A área de reconhecimento facial está presente em nosso dia a dia, como por exemplo, em sistemas de segurança e entretenimento [19, 21]. Segundo Bombardelli [2], devido a popularização das redes sociais originou-se uma grande quantidade de dados, simultaneamente com o aumento da capacidade de processamento

de computadores e com novas técnicas no processamento de imagens e no aprendizado de máquina, e que influenciaram na precisão dos métodos de reconhecimento facial.

O método de reconhecimento facial é realizado por meio de imagens e filmagens de câmera de segurança [16, 20]. O mesmo, não é considerado um dos mais precisos no quesito de autenticação, devido as variações de poses, expressões faciais e iluminação. Ainda assim, é um método que possui suas vantagens, pois por não ser um método intrusivo, não é necessário que o usuário interaja ou toque em algum dispositivo por um período, propiciando assim a identificação do indivíduo sem a autorização do mesmo [13, 16].

Porém, as câmeras estão se tornando presentes em todos os lugares, em relação a Internet das coisas<sup>1</sup>, e podem usar tecnologias de reconhecimento facial para melhorar o contexto de reconhecimento facial. Entretanto, há uma grande lacuna de precisão entre os sistemas de reconhecimento facial disponíveis ao setor público e ao privado, de última geração [1].

Há poucos projetos comerciais que utilizam esse método, pelo fato de ainda existir pouco conhecimento sobre reconhecimento facial e as ferramentas são onerosas, limitando o uso a um pequeno grupo de empresas. Muitos projetos buscam um aprimoramento e uma aplicação precisa de reconhecimento facial, para reforçar a segurança de um determinado ambiente [14].

Na área de segurança vêm sendo muito utilizado a biometria, que favorecendo empresas e sistemas públicos que são restritos quanto ao acesso de pessoas, bem como o uso em investigações criminais e policiais, tornando-se assim uma contribuição válida nessa área [14].

Segundo Amos, Ludwiczuk e Satyanarayanan [1], atualmente a maioria das técnicas de reconhecimento facial utilizam técnicas de inteligência artificial, a qual tem o intuito de aumentar a precisão de um sistema de reconhecimento facial. Uma dessas técnicas é o Máquina de Vetores de Suporte (do inglês *Support Vector Machines* – SVM), o qual é um conjunto de métodos de aprendizado de máquina, que são usados para classificação [15].

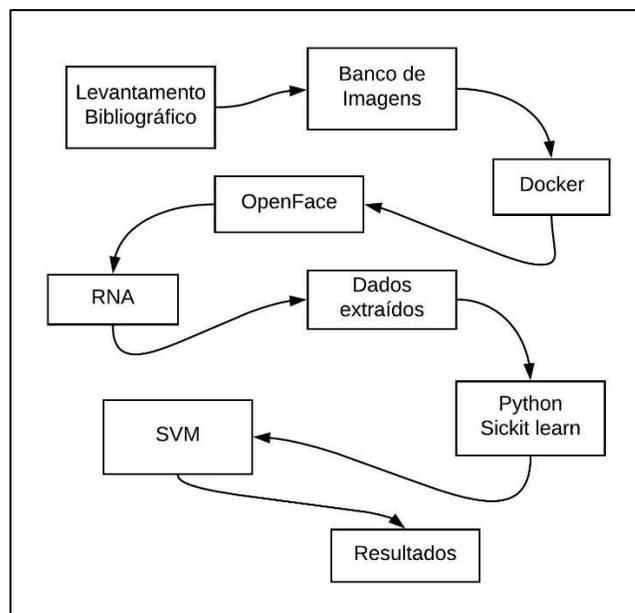
Toda via, a métrica acurácia é amplamente utilizada para avaliar a eficácia de um sistema de reconhecimento facial. Por isso, busca-se um maior conhecimento sobre qual o melhor método e eficiência de reconhecimento facial.

Deste modo, apresenta-se que o objetivo deste artigo é o de avaliar o SVM utilizado pela biblioteca *OpenFace* na classificação das imagens por meio do método *Cross-Validation*.

Este artigo está estruturado da seguinte forma: a Seção 2 descreve o reconhecimento facial, as bases de dados utilizadas, o classificador SVM, a RNA e o *OpenFace*. A Seção 3 descreve os testes e os seus resultados obtidos. A Seção 4 apresenta a discussão dos experimentos gerados, e a Seção 5 apresenta as conclusões.

## 2. MÉTODOS

A Figura 1 exemplifica as etapas realizadas nesta pesquisa. Inicialmente, fez-se necessário realizar um levantamento bibliográfico acerca da temática do reconhecimento facial e inteligência artificial.



**Figura 1 - Fluxograma experimental. Fonte: Dos autores (2018)**

Após, foi realizada uma pesquisa sobre os bancos de imagens existentes na Internet. Dentre os diversos bancos de imagens disponíveis para teste, foram selecionados três bancos:

- Labeled Faces in The Wild (LFW)<sup>2</sup>: possui cerca de 13 mil imagens diversas, com pessoas em diferentes poses, expressões e iluminação, além do fundo diferente de cada imagem, ou seja, em ambientes não controlados;
- MS-CELEB-1M<sup>3</sup>: contém 10 milhões de imagens para cerca de 100 mil celebridades, possuindo tamanho superior a 100Gb, devido a isto optou-se por trabalhar com um subconjunto de exemplos disponíveis para download. Deste subconjunto foram escolhidas cinco celebridades (Jennifer Lawrence, Justin Bieber, Ming-Na Wen, Luiz Fabiano e Sasha Pieterse), totalizando 566 imagens para o experimento. Assim como a LFW, o MS-CELEB-1M apresenta imagens em ambientes não controlados;
- FEI DATABASE<sup>4</sup>: é um banco de imagens brasileira que possui 2800 imagens, contendo 200 pessoas com 14 imagens cada. Diferentemente das outras bases, esta apresenta um ambiente controlado, com imagens em um fundo branco e todas as fotos são obtidas da mesma forma para todas as pessoas.

Primeiro, reproduziu-se o experimento de teste de acurácia da LFW, para validar os resultados apresentados no artigo “*OpenFace: A general purpose face recognition library with mobile applications*”. Os resultados da LFW utilizando o *OpenFace* são obtidos calculando a distância euclidiana ao quadrado nos pares e de rotulagem sob um limiar como sendo a mesma pessoa, acima do limiar como pessoas diferentes e abaixo como pessoas iguais. O melhor limite nos *folds* de treinamento é usado como limite na dobra restante [1].

<sup>2</sup> <http://vis-www.cs.umass.edu/lfw/>

<sup>3</sup> <https://www.msceleb.org/>

<sup>4</sup> <https://fei.edu.br/~cet/facedatabase.html>

<sup>1</sup> IOT: interconexão digital de objetos cotidianos com a Internet. [22, 23]

Posteriormente, empregou-se o uso das duas bases restantes (MS-CELEB-1M e FEI DATABASE) que possuem diferentes tamanhos, para avaliar o classificador SVM em diferentes cenários e tamanhos de exemplos. Para a utilização destas bases é necessária uma estrutura de diretórios, onde as imagens de cada indivíduo estejam separadas em diretórios diferentes (Figura 2).

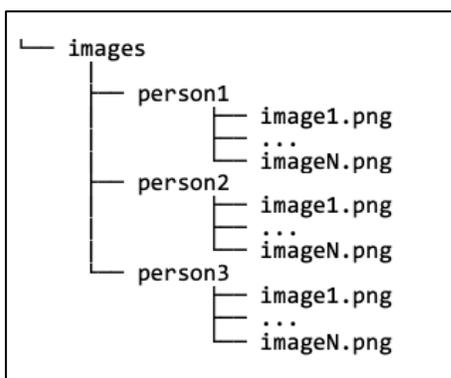


Figura 2 - Estrutura de diretórios em subdiretórios diferentes.  
Fonte: Dos autores (2018)

Diferentemente da LFW e MS-CELEB-1M, a FEI DATABASE não possui uma estrutura de pastas para armazenamento de imagens para cada pessoa. Para isso, foi elaborado um algoritmo (Figura 3), para que a separação não fosse realizada de forma manual. Então, para cada indivíduo foi criado uma pasta, resultando na separação das imagens.

Segundo Amos, Ludwiczuk e Satyanarayanan [1], a maneira mais fácil de executar o *OpenFace* é utilizando a imagem do ambiente para *Docker* (disponibilizado pelos criadores), que contém todas as bibliotecas e *frameworks* para o seu funcionamento. Escolheu-se utilizar a imagem do *container*, a fim de evitar problemas de compatibilidade com o sistema operacional, pois a instalação manual de todas as bibliotecas e *frameworks*, além de algumas como *OpenCV*, necessita realizar a compilação, podendo causar interrupções e demora no experimento.

Dessa forma, baixou-se a imagem do *OpenFace* e o mesmo foi posteriormente instalado no *Docker*. Com o container em execução e a imagem do *OpenFace* em funcionamento, iniciou-se o procedimento de pré-processamento das imagens, no qual foi utilizado o detector de pontos de referência com o método *dlib* [6] para centralizar e rotacionar a face na imagem (com base no nariz e olhos) (Figura 3). A partir deste momento, as imagens foram cortadas com base nos pontos de referência externo e redimensionadas em 96x96 pixels para aparecerem em locais semelhantes na entrada da rede neural<sup>5</sup>.

<sup>5</sup> Campo da ciência da computação ligado à inteligência artificial, buscando implementar modelos matemáticos que se assemelhem às estruturas neurais biológicas. [24]

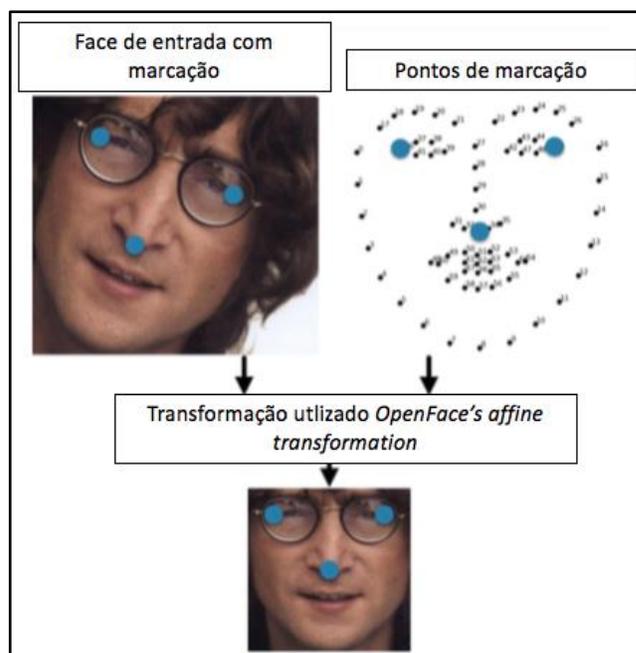


Figura 3 - Exemplo de pré-processamento. Fonte: Amos; Ludwiczuk; Satyanarayanan [1].

A rede neural escolhida foi “nn4.small2.v1”, que usa o alinhamento que consiste em três pontos principais: a parte externa dos olhos e o nariz, totalizando 68 pontos de referência em uma face. Ainda, “nn4.small2.v1” é pré-treinada e foi escolhida com base no desempenho apresentado no desenvolvimento no artigo do *OpenFace*.

A tabela 1 demonstra os resultados de desempenho das RNAs do *OpenFace*, utilizando um CPU de 8 núcleos de 3.70 GHz e uma GPU (Unidade de Processamento Gráfico, *Graphics Processing Unit*) Tesla K40 [1].

Tabela 1 - Performance da rede neural

Modelo	Tempo de Execução (CPU) (ms)	Tempo de Execução (GPU) (ms)
nn4.v1	75.67 ± 19.97	21.96 ± 6.71
nn4.v2	82.74 ± 19.96	20.82 ± 6.03
nn4.small1.v1	69.58 ± 16.17	15.90 ± 5.18
nn4.small2.v1	58.9 ± 15.36	13.72 ± 4.64

Fonte: Amos; Ludwiczuk; Satyanarayanan (2016) [1].

Neste trabalho, o *Docker* foi necessário para executar a RNA do *OpenFace* a qual é responsável por extrair as características das faces, já as demais etapas do AM (Aprendizado de Máquina) foram executadas diretamente no sistema operacional, sendo necessária a instalação de algumas bibliotecas citadas anteriormente, como *scikit-learn*, *numpy*, *matplotlib*.

Finalizado o pré-processamento das imagens, elas foram utilizadas na RNA para extração das representações faciais. Para sua aplicação, executou-se a linha de comando: `./batch-represent/main.lua -outDir "feature dir" -data "aligned IMG dir"`,

na qual a rede neural está implementada na linguagem Lua [2], para este comando são necessários dois parâmetros, “-outDir”, local onde são gerados os arquivos no formato CSV e “-data”, no qual encontram-se as imagens pré-processadas, que foram utilizadas pela RNA.

Para criação e avaliação do SVM foram usados métodos disponíveis na biblioteca *scikit-learn*. Onde, por meio do *cross\_val\_score* juntamente com o *K-fold*, dividiu-se o conjunto de dados em um subconjunto de treinamento e teste utilizando  $k=5$ . Ainda, devido ao fato de existir mais de uma classe usou-se o padrão *StratifiedKFold*, onde os *folds* são feitas preservando a porcentagem de amostras para cada classe.

A Figura 4 mostra o código-fonte implementado, que utiliza os arquivos CSV (*Comma Separated Values*) gerados pela RNA e que consistem nas representações faciais e suas respectivas classes. Após a criação do SVM com *kernel* linear é utilizado o método *cross\_val\_score*, passando por parâmetro, o classificador criado, dados de exemplos, classes e a quantidade de *folds*.

```

35 def crossVal(args):
36
37     fname = "{}/labels.csv".format(args.workDir)
38     labels = pd.read_csv(fname, header=None).as_matrix()[1:, 1]
39     labels = [os.path.split(os.path.dirname(label))[1] for label in labels]
40     le = LabelEncoder()
41     le.fit(labels)
42     labelsNum = le.transform(labels)
43     nClasses = len(le.classes_)
44
45     fname = "{}/reps.csv".format(args.workDir)
46     embeddings = pd.read_csv(fname, header=None).as_matrix()
47
48     X, y = embeddings, labelsNum
49
50     ln_clf = SVC(C=1, kernel='linear', probability=True)
51     clf = OneVsRestClassifier(ln_clf)
52
53     print("\nObtaining results cross-validation")
54     t_start = time.clock()
55     scores = cross_val_score(clf, X, y, cv=5, n_jobs=2)
56     t_end = time.clock()
57     print("Acuracy: %0.4f (+/- %0.2f)" % (scores.mean(), scores.std() * 2))
58     print("Cross validation time: {:0.4f}s.".format(t_end-t_start))
59     print("Scores: {}".format(scores))
60     print("=====")

```

**Figura 4 - Demonstração do código fonte das representações faciais e suas respectivas classes. Fonte: Dos autores (2018)**

Para que seja executado este método, a linha de comando: *Lucass-MacBook-Pro:evaluation lucasorso\$ python3 ROC\_CROSS.py crossVal \$(pwd)*, precisou ser executada. Neste comando consta a versão do *Python*, o nome do arquivo e o nome do método desenvolvido. Para executar este método, é necessário informar como parâmetro o caminho onde se encontram os arquivos no formato CSV.

Assim, ao executar o algoritmo desenvolvido é retornado um vetor de pontuações do classificador para cada execução da validação cruzada.

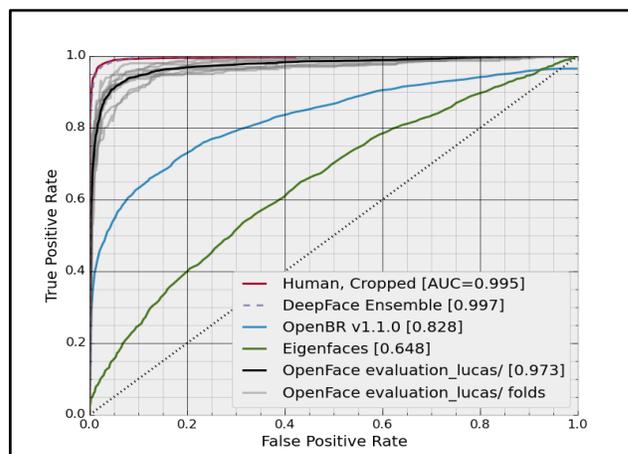
Com o intuito de complementar esta pesquisa, foram gerados, ainda, dois gráficos correspondentes aos dados das bases MS-CELEB-1M (Figura 7) e FEI DATABASE (Figura 8), utilizando-se a técnica t-SNE para demonstrar a matriz de 128 dimensões dos dados extraídos pela RNA.

A técnica t-SNE é uma técnica para redução de dimensionalidade que é particularmente adequada para a visualização de conjuntos de dados de alta dimensão. A técnica pode ser implementada via aproximações do Barnes-Hut, permitindo que ela seja aplicada em grandes conjuntos de dados do mundo real (t-SNE, MAATEN; HILTON, 2008). Ela converte semelhanças entre pontos de dados para probabilidades conjuntas e tenta minimizar a divergência de Kullback-Leibler<sup>6</sup> entre as probabilidades [15].

## 3. RESULTADOS

### 3.1 Reprodução do experimento de acurácia na base LFW

Nesse experimento, os resultados (Figura 5), demonstraram uma alta precisão na identificação de faces, quando utilizada a base de dados LFW.



**Figura 5 - Reprodução do experimento. Fonte: Dos autores (2018).**

Com esta reprodução foi possível compreender a forma que o *OpenFace* utiliza as técnicas de inteligência artificial para o processo de reconhecimento facial, com isso ficou esclarecido a forma de como utilizá-lo em outras bases de imagens.

A biblioteca do *OpenFace* contém o algoritmo também em *Python*, responsável por realizar o experimento na base da LFW. Porém, este algoritmo é somente para a base LFW, já que ela contém suas regras para a realização do experimento. Sendo assim, não foi possível testar e plotar a curva ROC utilizando o mesmo algoritmo em outros bancos de imagens como foi presumido no início desta pesquisa.

### 3.2 Validação cruzada com duas bases de imagens

Antes de iniciar o treinamento do classificador se fez necessário realizar o pré-processamento das imagens. Este pré-processamento consiste em identificar uma face em uma imagem, bem como alinhá-la. A partir disso, obtiveram-se os dados listados na Tabela 2 para cada uma das bases.

<sup>6</sup> Kullback-Leibler: É uma medida de como uma distribuição de probabilidades diverge de outra distribuição de probabilidades.

**Tabela 2 - Pós processamento dos dados**

	MS-CELEB-1M	FEI DATABASE
<b>Total</b>	566	2800
<b>Alinhadas</b>	549	2627
<b>Diferença</b>	17	173

Fonte: Dos autores (2018).

Conforme mostrado em nossos resultados, o algoritmo *dlib OpenFace* não conseguiu identificar algumas imagens. Para o subconjunto da base de dados MS-CELEB-1M não foram identificadas 17 imagens, e para a base de dados FEI DATABASE não foram identificadas 173 imagens.

Após o pré-processamento, a RNA é utilizada nas imagens alinhadas, onde foram obtidos dois arquivos no formato CSV. O arquivo “**reps.csv**” contém as representações das faces, que são definidas como um vetor bidimensional, em que  $x$  é a quantidade de imagens alinhadas e  $y$  é a quantidade de características que a RNA extraiu de cada face. Para o *OpenFace* a saída da RNA é um vetor com 128 posições para cada imagem, e como foram aplicadas várias imagens, o resultado é uma matriz, conforme demonstrado na tabela 3. Já o “**labels.csv**” também é um vetor bidimensional, no qual encontra-se o caminho completo de cada imagem utilizada, bem como seu índice.

**Tabela 3 - Matriz de representações**

	MS-CELEB-1M	FEI DATABASE
<b>Matri</b>	[ ]	[ ]
<b>z</b>	[ ]	[ ]

Fonte: Dos autores (2018).

A tabela 4 demonstra os resultados da validação cruzada nas bases MS-CELEB-1M e FEI DATABASE, respectivamente, que foram obtidos com o método *cross\_val\_score*.

**Tabela 4 - Resultados da validação cruzada**

	MS CELEB-1M	FEI DATABASE
<i>Fold1</i>	1.0	0.9
<i>Fold2</i>	0.97	0.91
<i>Fold3</i>	0.99	0.88
<i>Fold4</i>	0.99	0.89
<i>Fold5</i>	1.0	0.88
<b>Acurácia</b>	<b>0.99</b>	<b>0.89</b>

Fonte: Dos autores (2018).

A MS-CELEB-1M contém mais exemplos por cada classe e contém menos classes. Os resultados demonstraram uma maior acurácia com 99%, ao contrário dos resultados da FEI DATABASE, que mesmo contendo mais classes do que a MS-CELEB-1M e com menos exemplos em cada classe, obteve uma acurácia de 89%. Com este resultado compreende-se que, conforme maior o número de exemplos utilizados para o treinamento de um classificador, ele predirá com uma maior acurácia.

Nossos resultados demonstraram que os dados de exemplos da MS-CELEB-1M obtiveram uma acurácia de 99%, já utilizando a FEI DATABASE obteve uma acurácia de 89%. Para ambas as bases de dados o classificador foi avaliado com a estratégia da validação cruzada, ficando claro que a RNA que o *OpenFace* utiliza para a extração de características de uma face, são dados adequados de exemplos para a geração de classificadores.

### 3.3 Testes Adicionais

Para complementação do trabalho, utilizou-se a técnica t-SNE para demonstrar a matriz de 128 dimensões dos dados extraídos pela RNA. Sendo assim, dois gráficos foram gerados, correspondentes aos dados das bases MS-CELEB-1M (Figura 6) e FEI DATABASE (Figura 7).

Os resultados demonstraram na base MS-CELEB-1M (Figura 6) que a RNA conseguiu segregar as classes distantes entre si. Isso se dá pelo fato de haver menos classes e mais quantidade de exemplos do que a base FEI DATABASE.

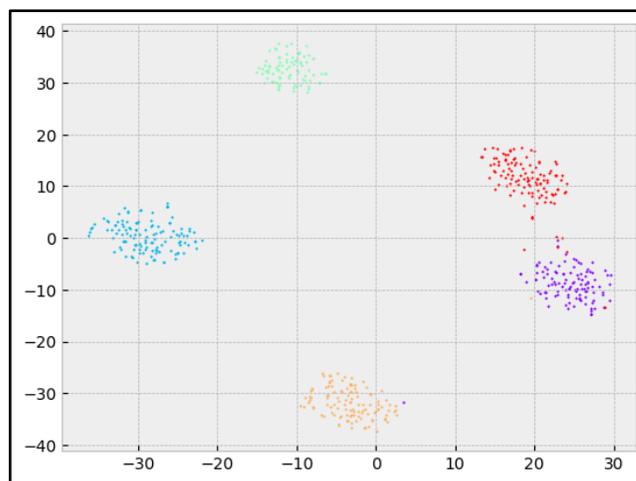
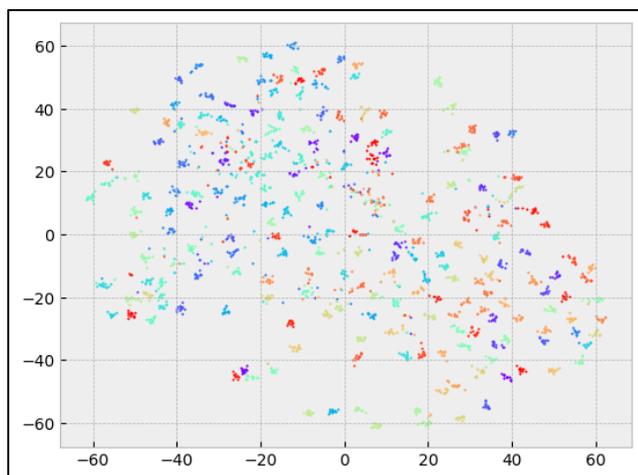


Figura 6 - Resultado t-SNE MS-CELEB-1M. Fonte: Dos autores (2018).

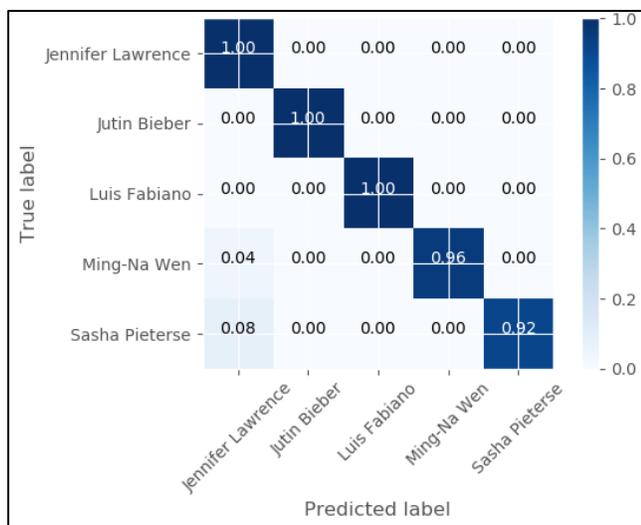
Já na base FEI DATABASE que contém mais classes e menos exemplos, apresentou uma maior aproximação entre as classes no seu resultado (Figura 7). Mesmo assim, é possível distinguir a segregação entre as classes.



**Figura 7 - Resultado t-SNE FEI DATABASE. Fonte: Dos autores (2018).**

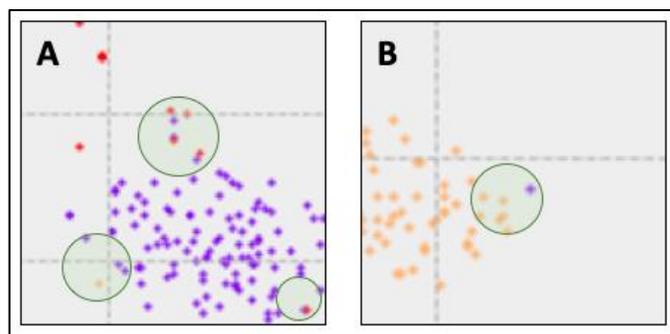
Para finalizar, foi realizado o teste de matriz de confusão<sup>7</sup>, que mostra como o classificador criado conseguiu distinguir cada classe. A Figura 8, é uma amostragem do que foi obtido a partir de um dos *folds*, da base MS-CELEB-1M, utilizando método *confusion\_matrix*, o qual está disponível na biblioteca *scikit-learn*. É possível verificar qual classe verdadeira foi predita erroneamente. Exemplo é a classe real Sasha Pieterse, que foi classificada com uma pontuação de 0.08 como sendo Jennifer Lawrence, observando assim um erro.

Para a FEI DATABASE a matriz também foi gerada, porém como ela é obtida por meio de classes reais, preditas, e por conter 200 classes ficou inviável a sua visualização.



**Figura 8 - Amostra da matriz de confusão para cada classe. Fonte: Do autor (2018).**

A diferença que o classificador encontrou na Figura 9, deu-se pelo fato da RNA extrair características semelhantes para classes distintas (Figura 9 A e B).



**Figura 9 - Características semelhantes entre as classes. Fonte: Dos autores (2018).**

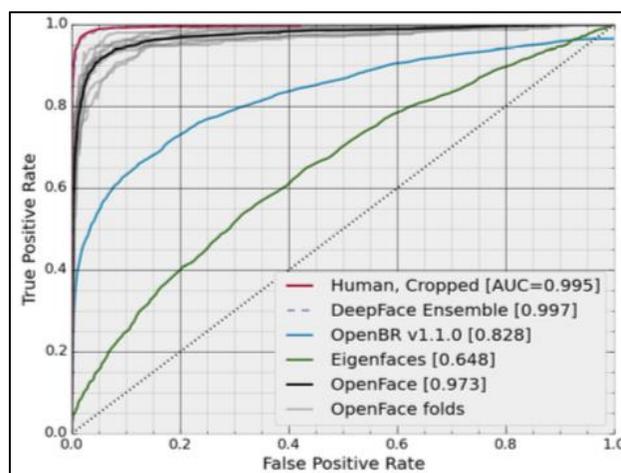
Visto isso, pode-se dizer que os resultados apresentados demonstraram uma acurácia significativa no reconhecimento facial, o que corrobora que a biblioteca *OpenFace* é uma opção relevante de estudo para área de reconhecimento facial.

## 4. DISCUSSÃO

### 4.1 Reprodução do experimento de acurácia na base LFW

Analisando os resultados conforme a Figura 5, obtivemos que estes coincidiram com os mesmos resultados apresentado por Amos, Ludwiczuk e Satyanarayanan [1], que demonstram uma alta precisão na identificação de faces, quando utilizaram a base de dados LFW.

A Figura 10 representa o gráfico obtido do trabalho de Amos, Ludwiczuk e Satyanarayanan [1]. Quando comparados esses resultados aos do trabalho aqui apresentado, é possível afirmar que a construção do experimento é replicada, obtendo-se os mesmos resultados e acurácia.



**Figura 10 - Acurácia do OpenFace no teste LFW. Fonte: Amos, Ludwiczuk e Satyanarayanan [1]**

<sup>7</sup> Ferramenta usada para avaliações de modelos de classificação em aprendizado de máquina. Aqui ela foi utilizada para distinguir as classes.

## 4.2 Validação cruzada com duas bases de imagens

Como já citado anteriormente neste artigo, conclui-se que conforme maior o número de exemplos utilizados para o treinamento de um classificador, maior a predição da acurácia.

Maia [10], cita em seu trabalho o desempenho da RNA em termos da capacidade de generalização, que está diretamente ligada a fatores como qualidade da base de treinamento, distribuição das classes e representatividade dos padrões.

Presume-se ainda, que a RNA utilizada pelo *OpenFace* conseguiu gerar com qualidade as representações de cada classe, gerando assim uma base adequada ao treinamento e ser utilizado no AM e consequentemente, conseguindo uma boa pontuação no modelo do classificador. Da mesma forma, Delbiaggio [3] demonstrou que teve êxito em relação aos outros algoritmos de reconhecimento facial, mesmo utilizando poucos exemplos para treinamento. Segundo o mesmo autor, ainda que alcance resultados melhores que os demais testados em seu trabalho, o *OpenFace* encontrou alguns erros, mas estes foram minimizados com dados de treinamento maiores. Sendo assim, fica evidenciado que quanto mais exemplos utilizados no AM para criação de classificadores, melhor será a pontuação deste.

No estudo de Meloni [12], foi utilizado o SVM para classificar as imagens de sensoriamento remoto por meio da validação cruzada, onde foi utilizado o método *Leave-One-Out*, diferentemente do método deste estudo (*OneVsRest*). Ele obteve resultados que confirma a aplicabilidade dos classificadores, e o uso da técnica de aprendizado de máquina, mesmo que as imagens sejam usadas para diferentes aplicações o comportamento do classificador apresenta uma evolução em seu desempenho, ou seja, com a adição de características, contribuam no progresso dos resultados.

Já no trabalho de Horta et al. [5], utilizou-se várias técnicas para classificação de empresas, como Regressão Logística (RL), SVM, MLP e Árvore de Decisão (AD). Os autores destacam que os classificadores com melhores resultados foram AD e o SVM. Neste caso, o AD obteve um desempenho melhor que o SVM. Contudo, Santos [18], em sua dissertação sobre o uso de SVM para aprendizagem e reconhecimento de objetos baseado na aparência, apresentou que o tipo de *Kernel* geralmente não interfere na precisão final do classificador. Entretanto, Amos, Ludwiczuk e Satyanarayanan [1] citam em seu trabalho sobre a escolha do *Kernel* Linear para ser utilizado como padrão no *OpenFace*.

Santos [18] também realizou outro experimento comparando MLP e SVM, e considerando apenas 10 classes, os resultados da SVM apresentaram uma pequena superioridade em relação à MLP, que é uma técnica totalmente viável para problemas de classificação de padrões, em especial para problemas de reconhecimento baseado na aparência, o que corrobora ainda mais com nossos resultados.

Meloni [12] em seu estudo complementa que a adição de características favorece na evolução dos resultados, criando classificadores com um desempenho maior, e ressalta que as escolhas de suas características colaboraram na classificação das imagens de satélite.

## 5. CONCLUSÃO

A replicação do teste por meio da LFW na biblioteca *OpenFace* e a utilização de dois bancos de imagens (MS-CELEBE-1M e FEIDATABASE) nos possibilitou compreender melhor como a IA (Inteligência Artificial). é utilizada na área de reconhecimento facial e como ocorre a extração das características faciais, bem como o A.M. para o treinamento e classificação de novos exemplos. Assim, todos os experimentos realizados nesta pesquisa contribuíram para uma compreensão prática da técnica, de reconhecimento facial, bem como de seu uso.

A validação cruzada apresentou uma melhor acurácia na MS-CELEB-1M (99%) do que na FEI DATABASE (89%), indicando que a utilização de mais exemplos por classe pode interferir no classificador.

A utilização da rede neural pré-treinada do *OpenFace*, para a extração de características, juntamente com um classificador SVM obteve resultados considerados satisfatórios. De modo que, evidentemente mais testes sejam necessários, diante dos resultados apresentados pode-se considerar o *OpenFace* como uma boa opção para o uso do reconhecimento facial.

## REFERÊNCIAS

- [1] AMOS, B.; LUDWICZUK, B.; SATYANARAYANAN, M. (2016) "OpenFace: A general-purpose face recognition library with mobile applications". 20 f. TCC (Graduação) - School Of Computer Science, Carnegie Mellon University, Pittsburgh, Pa 15213, Eua.
- [2] BOMBARDELLI, F. G. (2015) "Estudo sobre Reconhecimento Facial", 5 f. Monografia (Especialização) - Curso de Ciência da Computação, Universidade Federal do Paraná, Paraná.
- [3] DELBIAGGIO, N. (2017) "A comparison of facial recognition's algorithms", 41f. Trabalho de Conclusão de Curso (Graduação em Tecnologia da Informação Empresarial) - Universidade de ciências aplicadas, Finlândia.
- [4] GONZALEZ, R. C.; WOODS, R. E. (2006) "Digital Image Processing", New Jersey: USA, p.103.
- [5] HORTA, R. A. M.; BORGES, C. C. H.; CARVALHO, F. A. A. DE.; ALVES, F. J. S.; (2011) "Previsão de Insolvência: Uma Estratégia para Balanceamento da Base de Dados Utilizando Variáveis Contáveis de Empresas Brasileiras", Sociedade, Contabilidade e Gestão. v. 6, n. 2.
- [6] KING, D. E. Dlib-ml: "A Machine Learning Toolkit" (2009) Journal of Machine Learning Research. v. 10, p. 1755–1758.
- [7] LOPES, E. C. (2017) "Detecção de Faces e Características Faciais". Relatório Técnico, Rio Grande do Sul. 2016, <http://www.pucrs.br/facim-prov/wp-content/uploads/sites/19/2016/03/tr045.pdf>.
- [8] LOPES, R. G. C.; MARIO, M. C.; SILVA-FILHO, J. I. (2012) "Estudo para medição de peças através de imagens utilizando Lógica Paraconsistente Anotada", Science and Technology. v. 1, n. 1, p. 28-32.
- [9] MAATEN, L. V. D.; HINTON, G. (2008) "Visualizing High-Dimensional Data Using t-SNE", Journal of Machine Learning Research. v. 9, p. 2579-2605.
- [10] MAIA, H. L. F. (2016) Detecção e reconhecimento facial por meio de aprendizado de máquina. 50f. Trabalho de

- Conclusão de Curso (Graduação em Engenharia de Redes de Comunicação) - Universidade de Brasília, Brasília.
- [11] MARQUES-FILHO, O.; VIEIRA-NETO, H. (1999) “Processamento Digital de Imagens. Rio de Janeiro: Brasport”. p. 307.
- [12] MELONI, R. B. da S. (2009) “Classificação de Imagens de Sensoriamento Remoto usando SVM”, 64 f. Dissertação (Mestrado) - Curso de Ciência da Computação, Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro.
- [13] MENEZES, G. G. S. (2009) “Reconhecimento de faces usando análise de componentes principais e morfologia matemática”. 43f. Trabalho de Conclusão de Curso (Graduação em Engenharia da Computação) - Universidade de Pernambuco.
- [14] NUNES, F. T.; SILVA, J. B.; NICOLETE, P. C.; PEREIRA, J.; CRISTIANO, M. A. S. (2016) “Um Estudo sobre Técnicas de Biometria Baseadas em Padrões Faciais e sua Utilização na Segurança Pública”, p. 113 -129. In: **Tecnologias da Informação e Comunicação na Segurança Pública e Direitos Humanos**, Edited by SPANHOL, F. J.; LUNARDI, G. M.; SOUZA, M. V. de. São Paulo: Blucher.
- [15] PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V.; VANDERPLAS, J.; PASSOS, A.; COURNAPEAU, D.; BRUCHER, M.; PERROT, M.; DUCHESNAY, E. (2011) “Scikit-learn: Machine learning in python”, The Journal of Machine Learning Research. v. 12, p. 2825–2830.
- [16] PRAZERES, S. G. (2010) “Reconhecimento de Faces Utilizando Redes Neurais Autoassociativas”, 39f. Trabalho de Conclusão de Curso (Graduação em Engenharia da Computação) - Escola Politécnica de Pernambuco, Universidade de Pernambuco, Recife.
- [17] RODRIGUES, B. U. (2015) “Reconhecimento do tipo de cachaca utilizando visão computacional e reconhecimento de padrões”, 119 f. Dissertação (Mestrado em Ciência da Computação), Universidade Federal de Goiás, Goiânia.
- [18] SANTOS, E.M. DOS. (2002) “Teoria e Aplicação de Support Vector Machines à Aprendizagem e Reconhecimento de Objetos Baseado na Aparência”, 111 f. Dissertação (Mestrado) - Curso de Ciência da Computação, Departamento de Informática, Universidade Federal da Paraíba, Campina Grande.  
[http://docs.computacao.ufcg.edu.br/posgraduacao/dissertacoes/2002/Dissertacao\\_EulandaMirandadosSantos.pdf](http://docs.computacao.ufcg.edu.br/posgraduacao/dissertacoes/2002/Dissertacao_EulandaMirandadosSantos.pdf).
- [19] SORTE, L. X. B. (2011) “Reconhecimento de Faces: Aplicação de algoritmos de Redes Neurais”, In: SEMINÁRIO INTERNO DE COGNIÇÃO ARTIFICIAL, III, 2011, 3., 2011, Campinas. Campinas: Unicamp. p.5.
- [20] TRAMONTIN, E. D. (2016) Análise e aplicação de reconhecimento facial em sistema embarcado.75f. Trabalho de Conclusão de Curso (Graduação em Engenharia da Computação) - Universidade Federal de Santa Catarina, Araranguá.
- [21] VINODHINI, M; AISHWARYA, C.; SAFRINA, K. J.; RAMYA, K. (2016) “A comparative analysis of linear discriminant, principal component and evolutionary pursuit for recognition of human faces”, International Journal of Modern Trends in Engineering and Research (IJMTER). v. 3.
- [22] MOHAMMADI ZANJIREH, MORTEZA & LARIJANI, HADI. (2015). A Survey on Centralised and Distributed Clustering Routing Algorithms for WSNs. IEEE Vehicular Technology Conference. VTC 2015. Glasgow, Scotland.
- [23] CONNER, MARGERY (14 de Agosto de 2018). «Sensors empower the "Internet of Things"» (10): 32-38. ISSN 0012-7515.
- [24] Ferneda, E. (2006). Redes neurais e sua aplicação em sistemas de recuperação de informação. *Ciência da Informação*, 35(1), 25–30. <https://doi.org/10.1590/s0100-19652006000100003>