

RECONHECIMENTO DE FACES UTILIZANDO PADRÃO BINÁRIO LOCAL E KNN

Vinicius Santos Andrade

Centro Universitário Sagrado Coração – UNISAGRADO - Bauru – São Paulo
vinicius.andrade@unisagrado.edu.br

Renan Caldeira Menechelli

Centro Universitário Sagrado Coração – UNISAGRADO - Bauru – São Paulo
renan.menechelli@unisagrado.edu.br

Elvio Gilberto da Silva

Centro Universitário Sagrado Coração – UNISAGRADO - Bauru – São Paulo
egilberto@uol.com.br

Patrick Pedreira Silva

Centro Universitário Sagrado Coração – UNISAGRADO - Bauru – São Paulo
patrickpsilva@gmail.com.br

Resumo. É comum a utilização de técnicas de Processamento de Imagens para solucionar diversos problemas computacionais, isso é reflexo do crescimento de dados oriundos de imagens. Este trabalho tem como foco a área de segurança da informação e, assim, tem como principal objetivo, apresentar um algoritmo em Matlab que seja capaz de analisar imagens de faces humanas (biometria), utilizando o método Padrão Binário Local (LBP) para extrair informações da face (textura), e com auxílio do K vizinhos mais próximos (KNN), efetuar o reconhecimento da face.

Palavras-chave: Análise de Imagens. Visão Computacional. Padrão Binário Local. KNN. Biometria.

Abstract. *It is common to use Image Processing techniques to solve several computational problems, this is a reflection of the growth of data from images. This work focuses on the area of information security and, therefore, its main objective is to present an algorithm in Matlab that is capable of analyzing images of human faces (biometrics), using the Local Binary Standard (LBP) method to extract information face (texture), and with the help of K nearest neighbors (KNN), perform face recognition.*

keywords: image analysis; computational vision; local binary standard; kNN; Biometry.

1 INTRODUÇÃO

Aos poucos a tradicional senha de n dígitos, muitas vezes só número, em alguns casos com letras também e até mesmo caracteres especiais, vem perdendo espaço para senhas baseadas em biometria. Atualmente, inúmeros dispositivos celulares, até mesmo os mais simples, já contam com reconhecimento digital. Inúmeros bancos já possuem também seus caixas eletrônicos com opções de biometria digital.

Além da questão da segurança, o reconhecimento de face tem recebido recentemente uma atenção significativa, especialmente durante os últimos anos. Pelo menos duas razões explicam essa tendência: a primeira é a ampla gama de aplicações comerciais e de aplicação das leis, e a segunda é a disponibilidade de tecnologias viáveis após 30 anos de pesquisa. O reconhecimento facial também pode ter diversas outras aplicações como, por exemplo, videogames, realidade virtual, programas de treinamentos, acesso à internet, segurança em aeroportos e investigação criminal (ZHAO *et al.*, 2003).

Este trabalho sugere um algoritmo de classificação de faces humanas, utilizando o método Padrão Binário Local (LBP) para extrair informações da face e o kNN para classificação.

2 PROCESSAMENTO DIGITAL DE IMAGENS

Gonzales e Woods (2010, p.1) descrevem uma imagem digital sendo:

Uma função bidimensional, $f(x, y)$, em que x e y são coordenadas espaciais (plano), e a amplitude de f em qualquer par de coordenadas (x, y) é chamada de intensidade ou nível de cinza da imagem nesse ponto.

Quando essas coordenadas e os valores de intensidade de f são quantidades finitas e discretas, se torna então uma imagem digital. A imagem é formada por diversos elementos cada um com sua localização e valor que são chamados de elementos pictóricos, elementos de imagem, *pixels*, porém o termo mais utilizado para representar tais elementos é Pixel (Gonzales; Woods, 2010). Resumidamente, a imagem digital é um conjunto de *pixels* que são os elementos da imagem em um plano cartesiano definida por uma função $f(x, y)$.

Gonzales e Woods (2010) consideram três tipos de processos computacionais que são os processos de níveis baixo, médio e alto. Os processos de nível baixo usam operações consideradas primitivas como o pré-processamento de imagens, o realce de contraste e o aguçamento, e é caracterizado principalmente pelo fato da entrada e saída serem imagens. Os de nível médio usam tarefas como a segmentação, a descrição de objetos e classificação que é o reconhecimento de objetos, e é caracterizado pelo fato de suas entradas no geral serem de imagens, porém as saídas são atributos dessas imagens. E os processos de nível alto se resumem a reagrupar este conjunto de objetos reconhecidos, realizando as funções cognitivas normalmente associadas à visão.

2.1 O QUE É TEXTURA

A textura apesar de ser um termo muito utilizado, não é algo precisamente definida, porém existem diversos autores que a explicam de variadas maneiras. De acordo com Haralick (1979, p. 786):

Uma imagem com textura pode ser descrita pelo número e tipos de suas primitivas e pela organização espacial ou layout de suas primitivas. A organização espacial pode ser aleatória ou pode obedecer a um padrão relativamente uniforme entre uma primitiva e suas primitivas vizinhas. Esta dependência pode ser estrutural, probabilística ou funcional.

Desta forma, textura é uma propriedade de um pixel com relação aos seus *pixels* vizinhos. Sendo assim, pode-se interpretar texturas com distintos graus de aspereza, granulosidade, suavidade, entre outras características.

Para Abbadeni, Ziou e Wang (2000), a textura é uma das mais significantes características da imagem extremamente usada na classificação de imagens, reconhecimento e na recuperação de problemas e que tem uma função importante na percepção visual humana. Abbadeni, Ziou e Wang (2000) consideram a textura como sendo a distribuição espacial dos níveis cinzas em uma imagem, que podem ser uma repetição determinística ou aleatória de uma ou várias primitivas em uma imagem.

Bhiwani, Khan e Agrawal (2010) consideram que a textura pode ser observada em diversas imagens artificiais a partir de imagens como de satélites multiespectrais ou imagens microscópicas. A textura é uma característica inata de todas as superfícies, como por exemplo, um chão de madeira, a pele humana, um tecido de uma roupa etc. Cada textura da imagem aparenta conter uma informação que pode ser utilizada para analisar a característica daquela imagem.

Gonzales e Woods (2010) também concordam que não há uma definição concreta para textura. A textura contém três abordagens principais as quais se usam no processamento de imagens para descrever a textura de alguma região, as quais são:

- a) Abordagens estatísticas, que produzem caracterizações das texturas como suave, rugosa, granulada;
- b) Abordagens estruturais, que lidam com arranjos de primitivas de imagens, como a descrição de textura baseada em linhas paralelas espaçadas regularmente;
- c) Abordagens espectrais, que estão baseadas em propriedades do espectro de Fourier e são usadas principalmente para detectar a periodicidade global em uma imagem pela identificação de picos de alta energia no espectro.

3 ABORDAGENS PARA DESCRIÇÃO DE TEXTURA

Visto que a textura ainda é um termo muito subjetivo, diversos autores usam várias abordagens para assumir uma definição. Como já citado anteriormente, existem três principais abordagens que são usadas no processamento de imagens que são as estatísticas, estruturais e espectrais (GONZALES; WOODS, 2010). No entanto, Nixon e Aguado (2002) classificam os métodos de análise de textura em estrutural, estatística e abordagem combinatória.

As características das texturas podem ser categorizadas em quatro famílias segundo Xie e Mirmehdi (2008), as quais são: características estatísticas, estruturais, características baseadas em filtros e características baseada em modelos.

A seguir iremos detalhar melhor as abordagens estatísticas e estruturais, uma vez que foram utilizadas neste trabalho.

3.1 ABORDAGEM ESTATÍSTICA

Na abordagem estatística, é medida a distribuição espacial dos valores dos *pixels*, sendo que as características da textura são calculadas baseadas na distribuição estatística da intensidade da imagem especificada em relação a posição dos *pixels*. (XIE; MIRMEHDI, 2008).

Segundo Ojala e Pietikäinen (2013), este método estatístico analisa a distribuição espacial dos valores de cinza e calcula as características locais em cada ponto da imagem, derivando um conjunto de estatísticas das distribuições das características locais. Os métodos também são classificados pela sua quantidade de *pixels*, como primeira ordem (um *pixel*), segunda ordem (dois *pixels*) e ordem superior (três *pixels* ou mais). A principal diferença entre estas estatísticas é que a de primeira ordem estima a propriedade e o valor de cada *pixel* individualmente, ignorando a interação entre os outros *pixels*, já as de segunda ordem e ordem superiores estimam estes valores de dois ou mais *pixels* em determinada região.

3.2 ABORDAGEM ESTRUTURAL

Do ponto de vista estrutural, a textura é caracterizada por primitivas de textura ou os elementos da textura e a organização espacial dessas primitivas. (VILNROTTER *et al.*, 1986).

Com base neste contexto, Xie e Mirmehdi (2008) enfatizam que os principais objetivos das abordagens estruturais antes de tudo devem extrair as primitivas de textura e depois modelar ou difundir as regras de colocação espacial. A textura primitiva pode ser tão simples como os *pixels* individuais, uma região com os níveis de cinza uniformes ou com segmentos de linha. As regras de colocação podem ser obtidas através da modelagem de relações geométricas entre primitivas ou aprendendo suas propriedades estatísticas.

Julesz (1981) propôs o conceito de textons como uma estrutura fundamental da imagem. Os textons foram considerados como átomos de percepção visual humana pré-atenciosa.

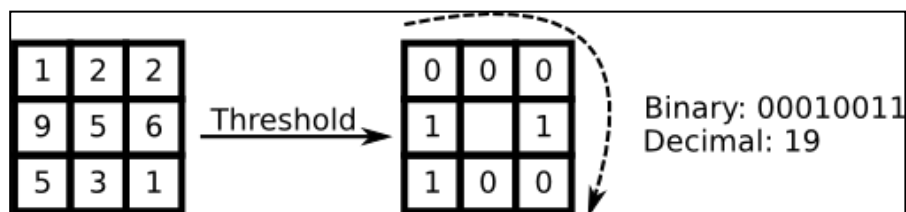
As características estruturais geralmente são mais estáveis nas mudanças de iluminação geral do que as características estatísticas, mas dependem fortemente da detecção primitiva. Depois que as primitivas são encontradas, pode-se calcular as propriedades estatísticas das primitivas, construindo então regras de colocação para elas, ou utilizar os textons encontrados. Muitas vezes, os métodos baseados em estruturas são justificados usando estudos psicofísicos da percepção de textura: um humano poderia discriminar texturas com diferentes elementos textons (JULESZ, 1981).

4 METODOLOGIA PADRÃO BINÁRIO LOCAL

A Metodologia Padrão Binário Local do inglês *Local Binary Pattern* (LBP) foi a metodologia escolhida para servir de base para o algoritmo deste trabalho. De acordo com Ojala, Pietikäinen e Harwood (1996), o LBP é o descritor de textura invariante simples e poderoso das escalas de cinza. O operador LBP combina as características de análise de textura estatística e estrutural e com isto descreve a textura com micro primitivas e suas regras de colocação estatística. O LBP original opera em uma base de *pixels* e descreve os oito *pixels* vizinhos no código binário e resume todos os códigos em um histograma que serve como recurso de textura. Este método produz 256 padrões de textura para 3x3 *pixels* vizinhos.

Neste método, o pixel central é usado como uma entrada quando comparado com seus vizinhos. Se o valor desse pixel central for maior do que o valor do vizinho, definimos o valor do vizinho como "1", caso contrário, ele será definido como "0". Conforme pode ser observado na Figura 1 temos 8 vizinhos em um bloco 3x3, o LBP produz um número binário de 8 dígitos. Esse número será convertido em um sistema decimal numérico.

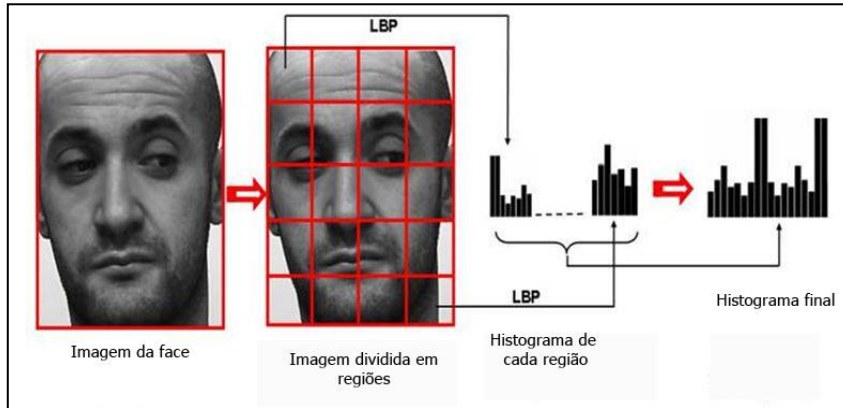
Figura 1 - Extração de características



Fonte: OpenCV (2020).

Como proposto por Ahonen *et al.* (2006), a metodologia que o LBP usa para descrever uma face dá-se pela divisão das regiões locais e então o descritor de textura do LBP são extraídos de cada região. Os descritores são concatenados para formar uma descrição única do rosto, conforme ilustra a Figura 2.

Figura 2 - Descrevendo a face com LBP

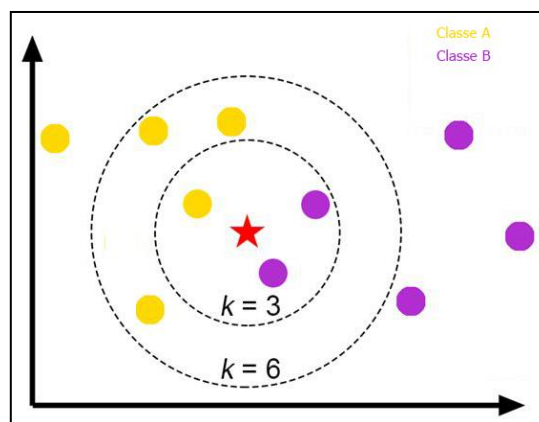


Fonte: ADVANCEDSOURCECODE (2020).

5 K-NEAREST NEIGHBORS (KNN)

De acordo com Parsian (2015), *K-Nearest Neighbors* (KNN), ou em para português, *K-Vizinhos Próximos* é um algoritmo importante de aprendizado de máquina, onde k é um número inteiro maior que 0. O kNN encontra os k pontos de dados mais próximos em um conjunto de dados para um determinado dado de consulta ponto. Esta operação pode ser representada como: dados dois conjuntos de dados X e Y , queremos encontrar o vizinho mais próximo de Y para cada objeto em X . Portanto, o KNN analisa o seu vizinho ou *pixel* mais próximo que classifica os objetos n -dimensionais com base na sua semelhança com outros objetos n dimensionais. A Figura 3 apresenta este contexto.

Figura 3 - Exemplificando o KNN



Fonte: Elaborada pelos autores.

No exemplo apresentado na Figura 3, são aferidas as distâncias de uma nova amostra, representada por uma estrela, às demais amostras de treinamento, representadas pelos círculos roxas e amarelas. A variável k representa a quantidade de vizinhos mais próximos que serão agrupados, portanto a Classe A receberá o valor $k=6$ e a Classe B receberá o valor $k=3$.

6 OUTRAS APLICAÇÕES

Não só o reconhecimento facial, mas toda a computação visual, tem diversas aplicações, como mostrado por Szelisk (2011) tais como:

- a) vigilância ou monitoramento;
- b) inspeção automática;
- c) reconhecimento ótico de caracteres;
- d) construção de Modelos 3D;
- e) segurança automotiva;
- f) biometria.

Especificamente, no que diz respeito ao reconhecimento facial, Parmar *et al.* (2013) destacam as seguintes aplicações:

- a) identificação Facial: identifica as pessoas pela sua face, para autorizar a presença de uma pessoa em um determinado local, ao invés de se usar o RG ou outros meios;
- b) controle de acesso: considerando um funcionário em um ambiente de trabalho, o reconhecimento facial pode determinar se aquele funcionário tem permissão para usar tal equipamento na empresa, por exemplo um computador;
- c) segurança: a segurança nos ambientes é de grande preocupação nos aeroportos. Diversos aeroportos já adotaram o reconhecimento facial como método de segurança. A tecnologia de reconhecimento facial é designada para localizar com base nas faces, um suposto terrorista que pode adentrar nos aeroportos, por exemplo;
- d) investigações da base de dados da imagem: através dessa base de dados através das imagens pode se localizar, crianças desaparecidas, imigrantes, policiais e motoristas licenciados;
- e) vigilância: o reconhecimento facial pode ser também utilizado para vigilância, porém a tecnologia ainda não traz total satisfação para o usuário pois, as condições de iluminação, a orientação do rosto, e outros fatores continuam sendo uma tarefa desafiadora para a vigilância em larga escala.

7 MATERIAIS E MÉTODOS

A pesquisa em questão é do tipo laboratorial executada através do programa Matlab R2016, o qual foi programado um algoritmo para efetuar o reconhecimento de faces. Para efetuar

os testes, utilizamos uma máquina com Windows 10, processador i5 8250U, 8GB de RAM e armazenamento SSD.

A base de dados de imagens utilizada na pesquisa possui 40 faces de pessoas distintas no formato PGM, sendo o tamanho de cada imagem de 92x112 *pixels*, com 256 níveis de cinza por *pixel*. Cada face contém um total de 10 imagens que chamamos de amostra, cada uma variando o ângulo, a expressão do rosto (olhos abertos ou fechado, sorrindo ou não) e demais características faciais, portanto, foi utilizado um total de 400 faces. Todas as imagens foram tiradas em um fundo escuro homogêneo. A base de dados de imagens pertence a AT&T Laboratories Cambridge e pode ser encontrada no site: <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>. Por algum motivo o site originário da base não está mais disponível. A base de dados também pode ser encontrada no link: <https://www.kaggle.com/kasikrit/att-database-of-faces>.

Não foi necessário efetuar etapas para remoção de ruídos, sendo assim, após a binarização fez-se a segmentação afim de criar um vetor de características com base nas imagens (faces) utilizando o LBP e, após, este vetor foi classificado utilizando o kNN.

8 RESULTADOS OBTIDOS

Após a leitura da imagem, foi utilizado o LBP para efetuar a extração de informações da imagem e criar o vetor de característica. Isso foi submetido ao algoritmo *KNN* com vizinhança 1 (padrão) que fez a classificação das imagens. Para o treinamento do algoritmo, foi utilizado 50% dos dados obtidos através do LBP. O algoritmo usa a primeira face como um treinamento, as próximas faces serão suas amostras calculado então o percentual de ser o rosto da mesma pessoa. A Figura 4 ilustra a divisão dos dados.

Face 1	Face 1.1	Face 2	Face 2.1
Treinamento	Amostra	Treinamento	Amostra

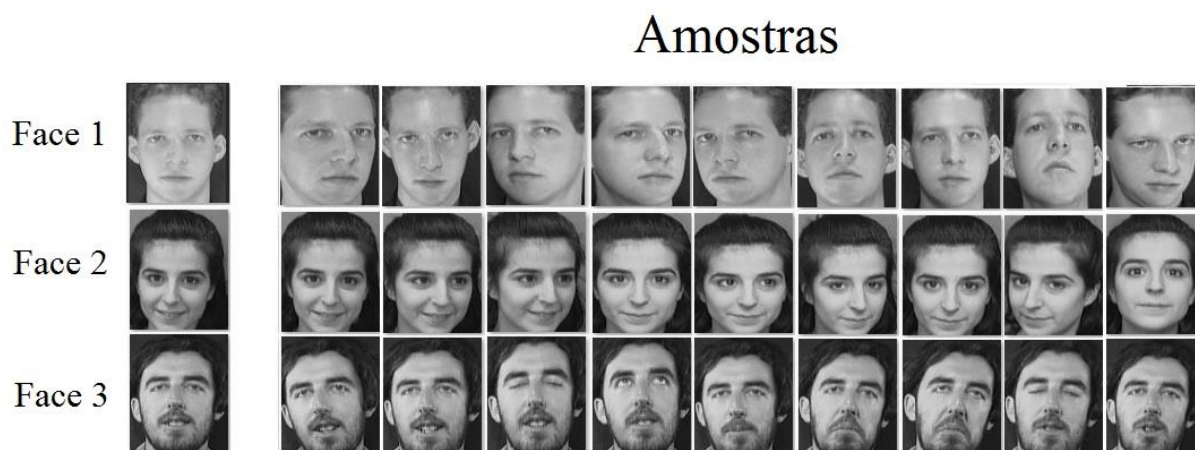
Figura 4 - Divisão dos dados

Fonte: Elaborada pelos autores.

No exemplo apresentado na Figura 4 tem-se dois faces (Face 1 e Face 2), onde cada modelo possui duas representações (Face 1 e Face 1.1; Face 2 e Face 2.1). Onde foram divididos em treinamento e amostra.

A Figura 5 representa algumas faces extraídas da base de dados. Ela demonstra as três diferentes faces de três pessoas distintas, sendo que cada pessoa tem dez faces de um ângulo e expressões diferentes. Como apresentado na Figura 5, a primeira face é o treinamento, as seguintes são as amostras.

Figura 5 - Exemplo de faces e suas respectivas amostragens



Fonte: Elaborada pelos autores.

No total foram feitos quatro experimentos. Para o primeiro experimento utilizou-se trinta faces com duas amostras para cada. Sendo assim temos o total de sessenta faces no primeiro experimento. Para o segundo experimento utilizou-se quarenta faces com duas amostras para cada. Sendo assim temos o total de oitenta faces no segundo experimento. Para o terceiro experimento utilizou-se quarenta faces com seis amostras para cada, portanto, utilizou-se toda base de dados com o total de duzentos e quarenta faces no terceiro experimento. Para o quarto experimento utilizou-se quarenta faces com dez amostras para cada, portanto, utilizou-se toda base de dados com o total de quatrocentos faces. O Quadro 1 exhibe os resultados (acurácia) obtidos em cada experimento. Observa-se que com o aumento no número de faces houve uma queda no valor da acurácia. Com até 80 faces o algoritmo foi capaz de manter 100% na classificação correta. Os piores resultados foram obtidos quando se considerou um maior número de amostras o que sugere que o algoritmo empregado apresenta um limite máximo de faces que devem ser consideradas para obter um desempenho satisfatório.

Quadro 1 - Resultados de Acurácia

Experimento	Número de faces utilizadas	Acurácia (%)
Primeiro	60	100,00
Segundo	80	100,00
Terceiro	240	33,3333
Quarto	400	20,00

Fonte: Elaborada pelos autores.

9 CONSIDERAÇÕES FINAIS

A identificação de faces utilizando LBP obteve bons resultados com pequenos grupos de faces. A queda drástica na acurácia para o terceiro teste pode ser resultante de amostras de imagens mais diferentes da que está sendo usada na comparação, por exemplo, entre outros casos.

Os testes sugerem que o LBP é bom apenas em situações que utilizam poucas imagens, pois no terceiro e quarto experimento ocorre uma queda considerável na acurácia não passa de 30%. Não é possível afirmar que o problema está no LBP, uma vez que, talvez, com algoritmos de classificação mais complexos, como *Fuzzy* ou redes neurais, seja possível obter melhores resultados em experimentos com número maior de faces. Seria interessante também o uso de outras curvas para mensurar o desempenho do algoritmo, como as de precisão e revocação.

REFERÊNCIAS

- ABBADENI, N.; ZIOU, D.; WANG, S. **Perceptual Textures Features Corresponding to Human Visual Perception**. Université de Sherbrooke. Canada. 2000.
- ADVANCEDSOURCECODE. Disponível em: <http://www.advancedsourcecode.com>. Acesso em: 27 jul. 2020.
- AHONEN, T.; ABDENOUR, H.; PIETIKÄINEN, M. Face Description with Local Binary Patterns: Application to Face Recognition. **IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE**. v. 28, n. 12, Dez. 2006.
- BHIWANI, R.J.; KHAN, M.A.; AGRAWAL, S.M. Texture Based Pattern Classification. **International Journal of Computer Applications** (0975 – 8887). V. 1. N. 1. 2010.
- GONZALES, R. C.; WOODS, R.E. **Processamento Digital de Imagens**. 3. ed. São Paulo. Pearson. 2010.
- HARALICK, R. M. Statistical and structural approaches to texture. **Proceedings of the IEEE**, 67(5):786-804, 1979.
- HARALICK, R.M.; SHANMUGAM, K. DINSTEIN, I. Textural Features For Image Classification. **Institute of Electrical and Electronics Engineers**. v. 3. n. 3. 1973.
- JULESZ, B. Textons, the elements of texture perception, and their interactions. **Nature**. v. 209, p. 91-97. 1981.
- NIXON, M.S.; AGUADO, A.S. **Feature Extraction and Image Processing**. 1. ed. Newnes. 2002.
- OJALA, T.; PIETIKÄINEN, M. **Texture classification**, 2013. Disponível em: http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/OJALA1/texclas.htm. Acesso em: 27 jul. 2020.
- OJALA, T.; PIETIKÄINEN, M. HARWOOD, D. A comparative study of texture measures with classification based on distribution. **Pattern Recognition**, v. 29, n.1, p.51-59, 1996.
- OPENCV. Disponível em: https://docs.opencv.org/2.4/modules/contrib/doc/facerec/facerec_tutorial.html. Acesso em: 27 jul. 2020.
- PARMAR, D. N.; MEHTA, B. B. Face Recognition Methods & Applications. **Int. J. Computer Technology & Applications**. v. 4. 2013.

PARSIAN, M.; **Data Algorithms**. 1. ed. United States of America. O'Reilly Media. 2015.

SZELISK, R. **Computer Vision: Algorithms and Applications**. Texts in Computer Science, Springer, 2011.

VILNROTTER F.; NEVATIA, R.; PRICE, K. Structural analysis of natural textures, **IEEE. Transactions on Pattern Analysis and Machine Intelligence**. 8, 76–89, (1986).

XIE, X.; MIRMEHDI, M. A Galaxy of Texture Features. Department of Computer Science, University of Bristol. 2008.

ZHAO et al.; **Face Recognition: A Literature Survey**. v. 35, n. 4, pp. 399–458, Dez. 2003.