

A Proposal for Recognition of Digital Facial Images: An Applied Case Study to Artificial Neural Networks. Uma Proposta para o Reconhecimento de Imagens Faciais Digitais: Um Estudo de Caso Aplicado à Redes Neurais Artificiais

ABSTRACT: Human traffic control systems are used in cities that have a large flow of people, especially places such as airports, squares, shopping malls and traffic. Therefore, they are increasingly used in biometric authentication applications where facial image is used to identify people previously registered. In some cases, facial image can also locate suspects in a crowd or can even assist in the search for missing persons. In order to do so, softwares are used for filtering the image that use Image Processing Techniques in order to define standards and simplify the search. Therefore, processing is streamlined. In many cases, machine learning algorithms are used for automating detection such as Artificial Neural Networks. Solutions like these perform real time comparisons that disregard simple changes, for instance aging, shaving, tanning and wearing glasses or hats. This work proposes methodology for facial recognition. After performing the image treatment and using filters such as Gaussian and Histogram Equalization, the Canny edges algorithm is applied for transforming the resulting image into a binary vector. This vector is used as input in a Neural Network to extract face patterns and perform facial detection. The proposal is promising with hits above 80% with the Gaussian filter and reaching 90% with the Histogram Equalization filter that allow the proposed methodology is used regardless of the filter applied to the image.

Keywords: Artificial Neural Networks, Facial Recognition, Canny Algorithms.

RESUMO: Os sistemas de controle de tráfego humano são utilizados em cidades que possuem um grande fluxo de pessoas, principalmente em locais como aeroportos, praças, shoppings e trânsito, sendo, por isso, cada vez mais comum o uso em aplicações de autenticação biométrica onde a imagem facial é usada para identificar pessoas previamente cadastradas, em alguns casos, localizando suspeitos em meio a uma multidão ou até mesmo, auxiliando na busca de pessoas desaparecidas. Para isso, são utilizados softwares que filtram a imagem usando técnicas de Processamento de Imagem afim de definir padrões e simplificar a busca, agilizando dessa forma, o processamento. Em muitos casos, com o objetivo de automatizar a detecção, são utilizados algoritmos de aprendizado de máquina, como as Redes Neurais Artificiais. Soluções como estas, realizam comparações em tempo real, desconsiderando mudanças simples, como envelhecimento, barba, bronzeamento e acessórios, como o uso de óculos ou chapéus. Neste trabalho, é proposta uma metodologia para reconhecimento facial que, após realizar o tratamento da imagem através de filtros como o Gaussiano e Equalização de Histograma, aplica-se o algoritmo bordas Canny, e transforma a imagem resultante, em um vetor binário o qual é utilizado como entrada em uma Rede Neural para extrair padrões da face e realizar a detecção facial. A proposta se mostrou promissora com acertos superiores a 80%, com o filtro Gaussiano, e chegando a 90% com o filtro Equalização de Histograma, permitindo, dessa forma, que a metodologia proposta seja utilizada, independente do filtro aplicado na imagem.

Palavras-chave: Redes Neurais Artificiais, Reconhecimento Facial, Algoritmo Canny.

1. INTRODUÇÃO

O avanço da tecnologia trouxe inovações na forma de se interagir com as máquinas, e com ela, inovações na forma de se acessar informações, equipamentos e até mesmo lugares. O uso de cartões de acesso e senhas digitadas, que podem ser facilmente clonados, adulterados ou extraviados, foram substituídos por algo mais pessoal e único para cada indivíduo, como por exemplo, a impressão digital, o reconhecimento de voz, da íris e/ou da face (Kolb, 2016).

O reconhecimento facial disponibiliza uma forma segura e conveniente de controle de acesso devido a sua natureza não intrusiva (Roychowdhury, 2016). Por conta disso, tem se tornado uma importante área de pesquisa para a comunidade em visão computacional, pois além de proporcionar a identificação de cada indivíduo, como único, através das suas características pessoais, possibilita, também, avaliar os seus estados emocionais, o que dá a capacidade crítica para realizar interações sociais bem-sucedidas (Geissnbuhler, 2020).

Embora os sistemas de reconhecimento facial estejam atingindo desempenhos similares ao humano, através do uso da Inteligência Artificial (IA), a maioria dos sistemas de reconhecimento facial ainda precisa superar algumas particularidades, como por exemplo: a vulnerabilidade e a captação das imagens.

A vulnerabilidade está ligada a ataques de apresentação, também conhecidos como ataques de falsificação. Simplesmente, apresentar uma foto impressa a um sistema de reconhecimento de rosto desprotegido pode ser suficiente para enganá-lo. A vulnerabilidade a ataques de apresentação, limita a implantação confiável de tais sistemas, para aplicativos, em condições não supervisionadas (Geissnbuhler, 2020).

Já a captação das imagens, em sistemas reais, pode ter variações de escala, iluminação e pose, ocultações e expressões faciais, dependendo do equipamento e do momento da captação. É difícil compensar essas questões no processo de reconhecimento facial. Para isso, diferentes técnicas são aplicadas nos mais variados campos, tentando inovar o processo, e alcançar, dessa forma, o melhor desempenho e precisão possíveis (Metzger, 2018).

Com o objetivo de contornar essas dificuldades, este trabalho apresenta uma metodologia para reconhecimento facial, onde as imagens são tratadas, no sentido de contornar algumas dessas falhas na captação. Após o tratamento, são aplicados filtros, como o Gaussiano e a Equalização de Histograma, sendo em seguida, utilizado o algoritmo de bordas Canny para captar as linhas da face. A imagem resultante é transformada em uma matriz bidimensional através da escala RGB (*Red-Green-Blue*) e, de posse desta, é gerado um vetor binário, que é utilizado como entrada na Rede Neural para realizar o reconhecimento facial.

2. TRABALHOS RELACIONADOS

Em (Nascimento, 2018) foi aplicada a técnica heurística de interpolação linear com a finalidade de corrigir a estimativa de pontos que contornam o rosto. Após essa etapa, utiliza-se o PCA (Análise de Componentes Principais) para criar um modelo parametrizado, capaz de descrever as formas do conjunto de característica comuns. Em seguida, é aplicado o algoritmo de árvores de regressão, onde são construídas árvores com uma fraca estimativa para realizar a predição, que serão gradativamente ajustadas. E isso é realizado até convergir à padrões mais precisos. Nesta etapa, são usadas distâncias euclidianas para calcular o posicionamento dos pontos vizinhos que formaram o contorno da face.

O trabalho de (Singh, 2016) realiza a identificação facial, após fazer o tratamento das imagens, através dos pontos de iluminação entre os olhos, já que os mesmos refletem um feixe luminoso sempre que recebem uma quantidade de luz adequada. Utilizam para isso

filtros e modelos como: RGB (*Red-Green-Blue*), HSV (*Hue, Saturation and Value*) e NTSC (*National Television System Committee*), para processar as imagens com operações morfológicas. A partir da localização das posições dos olhos, é calculada a mediana entre os mesmos para, em seguida, o centro da face ser localizado. Para separar a região do plano de fundo da imagem do rosto, foi aplicado o filtro Sobel, dando mais ênfase as bordas do rosto, olhos, nariz e boca e, logo em seguida, é aplicado o filtro de dilatação morfológica para se obter uma versão binária e se distinguir os olhos da face.

No trabalho de (Shanmugavadivu, 2016) são aplicados os procedimentos para analisar a cor da pele, que são considerados com baixo custo computacional, e para isso, foram determinadas faixas nos modelos RGB, HSV, Lab e YCbCr. As características, cromáticas e de luminosidade, são obtidas através dos padrões de cor no Lab e no HSV, respectivamente.

Após esse processo, são utilizados os modelos RGB, HSV e YCbCr para definir a tonalidade aceita como cores de pele e, posteriormente, são feitas operações morfológicas para remover as imperfeições na detecção. Com a região do rosto localizada, inicia-se o processo de classificação do nariz, onde o algoritmo considera o centro da face, da boca, considerada como uma barra horizontal, a região dos olhos, que são consideradas as regiões mais escuras do rosto. Após a etapa de localização dos órgãos da face é feita uma segmentação para detectar as bordas, e todos os pontos são armazenados na forma vetorial para realizar a classificação das imagens.

O trabalho em (Dang, 2016) realiza uma análise de vários modelos de detecção de face, mais especificamente entre os algoritmos Viola-Jones, SMQT *Features* e SWON *Classifier*, *Neural Network-Based Face Detection* e *Support Vector Machines-Based Face Detection*. O algoritmo Viola-Jones, por exemplo, foi criado em 2001 e é conhecido por retornar uma baixa taxa de falsos-positivos, pois utiliza *Haar Feature Selection*, responsável por calcular as características comuns de rosto humano, como distância dos olhos e tamanho do nariz. Um outro algoritmo também utilizado é o *Adaboost Training*, o qual é utilizado para construir classificadores e treiná-los, ajudando a encontrar pequenos detalhes críticos nas imagens.

Apesar de alguns dos métodos descritos utilizarem técnicas de contorno da face, bem como de vetores, a diferença entre os métodos descritos e o proposto neste trabalho, é a utilização de filtros para tratar a imagem captada, com o objetivo de retirar algumas falhas durante o processo de captação, permitindo, dessa forma, a utilização apenas do contorno da face.

Neste trabalho, as imagens captadas são convertidas para escala de cinza, e em seguida, passam por filtros, neste caso, o Gaussiano e Equalização de Histograma, com o objetivo de reduzir os detalhes do fundo através da segmentação. Após esse processo, é aplicado o algoritmo de bordas Canny, que tem como finalidade detectar o contorno do rosto e dos elementos que o compõe, como olhos, nariz e boca, por exemplo. Ao final desse processo, tem-se a proposta principal deste trabalho, pois a imagem resultante de todo esse processo terá apenas duas cores, o contorno do rosto (em branco) e o fundo da imagem (em preto), utiliza-se, então, o modelo RGB para transformar a imagem em uma matriz, que em seguida, é convertida em um vetor binário, utilizado como entrada na Rede Neural.

3. METODOLOGIA PROPOSTA

O tratamento de imagem vem sendo utilizado nas mais diversas áreas da vida cotidiana, seja para visualizar doenças, na área médica; na identificação de produtos, em anúncios de imagens editoriais; reconhecimento de faces, na área da segurança; na agricultura, para identificar áreas desmatadas, entre outros (Macarini, 2019).

O objetivo, em se utilizar algum processo para tratamento de imagens, é melhorar o seu aspecto visual para o seu posterior processamento. Neste trabalho, foi utilizada uma estrutura composta de cinco etapas, que vai da captação (aquisição) da imagem, até a transformação desta em um vetor binário, o qual é utilizado como entrada da Rede Neural, para o reconhecimento de faces. As etapas da metodologia proposta para o tratamento da imagem, podem ser observadas na Figura 1 e são descritas nas subseções seguintes.

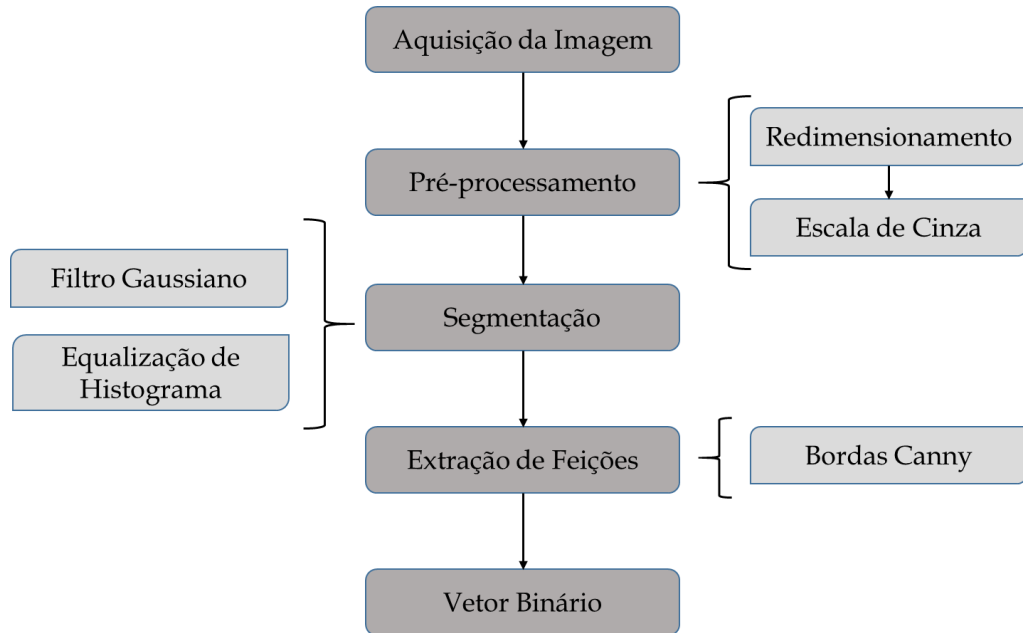


Figura 1 – Etapas da Metodologia Proposta

Fonte: Elaboração própria (2020).

3.1. Aquisição.

O processo de aquisição da imagem é realizado através da captura da face de um indivíduo, por algum dispositivo eletrônico, como por exemplo, celulares, câmeras de monitoramento, máquinas fotográficas, entre outros, como mostrado na Figura 2 (a). Essa imagem será armazenada e processada pelas etapas posteriores, com o objetivo de se realizar a sua identificação em uma base pré-armazenada.

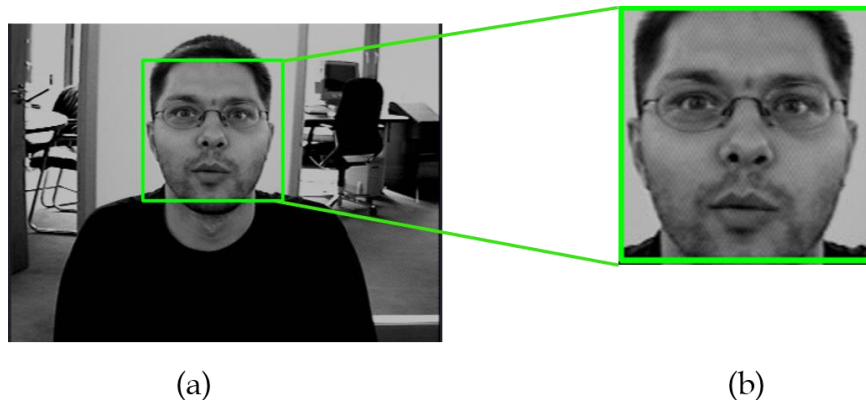


Figura 2 – Etapa 1 do Processamento da Imagem: Em (a), imagem original captada; em (b), Captura e redimensionamento apenas no rosto do indivíduo.

Fonte: Adaptado de *The BioID Face Database* (2019).

3.2. Pré-processamento.

O pré-processamento tem como objetivo corrigir possíveis defeitos e imperfeições que possam ter sido ocorridos durante a etapa de aquisição, como por exemplo, a filtragem de ruídos, o aumento do contraste, além de operações de rotação, ou redimensionamento. Nesta etapa são realizados dois procedimentos: o redimensionamento da imagem e a aplicação de um filtro para a transformação em escala de cinza.

Como as imagens, no geral, são obtidas por diferentes equipamentos, isso pode resultar em diferentes tamanhos. Neste trabalho, após a etapa de aquisição, realiza-se a padronização no formato das imagens (ou redimensionamento das imagens) que serão processadas, pois isso facilita e agiliza o processamento das etapas seguintes.

Para realizar o redimensionamento das imagens foi utilizado o OpenCV (Minichino, 2015), o qual é uma biblioteca de programação, com funções de divisão computacional em tempo real. As imagens utilizadas, independentemente do tamanho que são adquiridas, são redimensionadas para o tamanho de 40x30 pixels, como representado na Figura 2 (b).

Após o redimensionamento da imagem, é aplicado um filtro para transformá-la em escala de cinza, caso a imagem já esteja em escala de cinza, essa etapa é passada. Essa transformação é necessária, pois a técnica de detecção de bordas, depende de uma imagem em escala de cinza (Singh, 2016) para determinar os contornos da face humana. E tem como objetivo, também, separar o primeiro plano (face humana), do fundo.

3.3. Segmentação

Após a realização das etapas anteriores, o próximo passo, é aplicar o processo de segmentação, que consiste em particionar a imagem em regiões, ou objetos distintos. Este processo é geralmente guiado por características do objeto ou região, como por exemplo, cor ou proximidade. O nível de detalhamento em um processo de segmentação depende da tarefa a ser executada e da resolução da imagem que se tem. Existem diversas técnicas que podem ser utilizadas para a segmentação, neste trabalho, foram utilizadas a segmentação com filtro Gaussiano e com filtro Equalização de Histograma (Marengoni, 2009).

3.3.1. Filtro Gaussiano.

É um filtro que tem os valores da máscara determinados a partir de uma função bidimensional gaussiana discreta, com média igual a zero e desvio padrão σ , como mostrado na equação (1):

$$Gauss(x, y) = \frac{1}{2\pi\sigma^2} \exp \frac{-(x^2+y^2)}{2\sigma^2}, \quad (1)$$

onde x e y são as posições da máscara e Gauss (x,y) dá o valor a ser colocado na posição (x,y) da máscara. Os filtros gaussianos são filtros de média e são utilizados para suavizar as imagens de forma ponderada e simétrica (Marengoni, 2009), como pode ser visualizado na Figura 3 (a).

3.3.2. Equalização de Histograma.

É possível realizar cálculos sobre a distribuição de pixels para aumentar o contraste da imagem. A intenção neste caso é distribuir de forma mais uniforme as intensidades dos pixels sobre a imagem. A análise do histograma de uma imagem permite realizar considerações a respeito da distribuição de níveis de cinza da imagem, tornando possível identificar o quanto uma imagem é escura, clara, ou de níveis intermediários (Souza, 2007).

Contudo, conforme pode ser visto na Figura 3 (b), é possível que ocorram distorções e alterações nas cores da imagem equalizada, portanto, nem sempre a imagem mantém suas

características. Quando uma imagem apresenta uma distribuição insuficiente de níveis de cinza o processo de equalização de histograma, que visa o aumento da uniformidade da distribuição de níveis de cinza de uma imagem, produz bons resultados, permitindo dessa forma, realçar diferenças de tonalidade e resultando no aumento do nível de detalhes perceptíveis (Souza, 2007).

A expressão que apresenta um histograma equalizado é apresentado na Equação (2):

$$Hist_Equal = \frac{(L-1)}{MN} \sum_{j=0}^k n_j, \quad (2)$$

onde k é a intensidade no histograma equalizado, L é o valor máximo de intensidade na imagem, M e N são as dimensões da imagem e n_j é o número de pixel na imagem com valor de intensidade igual a j (Marengoni, 2009).

3.4. Extração de Feições

O propósito desta etapa é reduzir a quantidade de dados através da observação de certas feições, ou propriedades, que distinguem os padrões de entrada. Na extração de feições, transforma-se uma matriz de observação (foto original) em uma matriz de feições (matriz na escala RGB), empregando alguma função ortogonal ou não-ortogonal, de modo a se obter um espaço de características não correlacionado.

Variações nos níveis de intensidade dos pixels podem ser determinadas pelas derivadas primeira e/ou derivada segunda. Alguns métodos de determinação de borda utilizam estas técnicas. O processo consiste em se definir máscaras, que caracterizem estas variações, e em seguida, é realizada a convolução da imagem pela máscara.



Figura 3 – Etapa 2 do Processamento da Imagem: Em (a), imagem com o filtro Gaussiano; em (b), imagem com Filtro Equalização de Histograma; e em (c) extração das feições através das Bordas Canny.

Fonte: Adaptado de *The BioID Face Database* (2019).

3.4.1. Algoritmo Bordas Canny

Uma borda em uma imagem é caracterizada por uma mudança, normalmente abrupta, no nível de intensidade dos pixels. Os detectores de borda são definidos para encontrar este tipo de variação e, quando estes estão muito próximos, eles podem ser conectados formando uma borda ou um contorno, e assim definindo uma região ou objeto.

A técnica de detecção de borda desenvolvida por Canny (Canny, 1986), embora seja a mais complexa, é também a de melhor desempenho. O detector de bordas Canny possui três objetivos básicos: baixa taxa de erro (todas as bordas devem ser encontradas), os pontos

da borda devem ser bem localizados (as bordas localizadas devem ser próximas das bordas reais), e resposta única para os pontos de uma borda (o operador deve retornar apenas um valor para cada ponto sobre a borda) (Marengoni, 2009).

Este filtro, tem por finalidade embaçar a imagem para a remoção de linhas e detalhes desnecessários, pois o objetivo é fazer a identificação das bordas pelo formato original, como pode ser observado na Figura 3 (c).

3.5. Vetor Binário

Após a aplicação das bordas Canny, a imagem passa a conter apenas duas cores básicas: o branco, que na escala RGB decimal, é representado pelo valor (255, 255, 255), e o preto, que na mesma escala é representado por (0, 0, 0). Dessa forma, é possível transformar a Figura 3 (c), na matriz bidimensional, representada pela Equação (3).

$$M = \begin{pmatrix} 0 & 0 & 255 & \dots & 0 \\ 0 & 255 & 0 & \dots & 255 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 255 & 0 & \dots & 0 \end{pmatrix} \quad (3)$$

Esse procedimento é possível, pois acontece uma leitura na imagem, na qual se percorre pixel a pixel. O trajeto começa no primeiro pixel do canto superior esquerdo e segue em linha reta, da esquerda para a direita, até o final da primeira linha. Finalizada a leitura da primeira linha, a leitura recomeça no início da segunda linha, e o processo segue até chegar ao final da imagem. O objetivo dessa leitura é transformar cada pixel da imagem em um par $[x,y]$ na matriz.

Após a definição da matriz bidimensional (Equação 3), é necessário transformá-la em um vetor binário, pois é este vetor que será utilizado como entrada na Rede Neural, para o processo de reconhecimento de faces. Para realizar a transformação da matriz em vetor é utilizado o Algoritmo 1.

Algorithm 1 Obtenção do vetor binário

```
1: for  $x \leftarrow 1, 40$  do  
2:   for  $y \leftarrow 1, 30$  do  
3:     if  $[x, y] = 255$  then  
4:       print"1"  
5:     else  
6:       print"0"  
7:     end if  
8:   end for  
9: end for
```

Fonte: Elaborada pelo autor (2020).

Este algoritmo, irá substituir os valores (255) encontrados, pelo valor (1), e o que for diferente de (255), será substituído pelo valor (0). Ao final do processo, o que antes era uma imagem representada por uma matriz bidimensional (Equação 3), agora será um vetor binário de 1200 posições, já que a imagem é 40x30 pixels, (Equação 4), que será utilizado na camada de entrada da Rede Neural para realizar a identificação das faces.

$$V = [0 \ 1 \ 1 \ 0 \ 0 \ 1 \ 1 \ 1 \ 0 \ \dots \ 1 \ 0 \ 1] \quad (4)$$

4. REDES NEURAIS ARTIFICIAIS (RNA)

Uma Rede Neural Artificial é uma rede em multicamadas, que consiste em uma camada de entrada, uma (ou mais) camadas ocultas e uma (ou mais) camadas de saída e são comumente utilizadas na resolução de questões complexas, onde os modos das variáveis não são rigorosamente conhecidos. Uma de suas principais características é a capacidade de aprender através de exemplos e de generalizar a informação aprendida, gerando um modelo não linear e, tornando sua aplicação na análise espacial bastante eficiente (Qi, 2017).

Há ainda outros elementos que podem ser configurados ou predeterminados na arquitetura de uma rede, como: a quantidade de camadas ocultas e a quantidade de neurônios de cada camada; a função de transferência, o treinamento, o método de aprendizagem e medição de desempenho (Souza, 2019). Definidas estas configurações, uma RNA está preparada para ser treinada com base em uma amostra, composta por entrada e saída. A Figura 4 representa a RNA que foi utilizada neste trabalho, onde W_i , representa os pesos de entrada, W_o , os pesos de saída.

Uma RNA conta com inúmeros algoritmos para reconhecimento de padrões. Neste trabalho foi utilizado o *Backpropagation*, sendo a sua principal vantagem é que ele trabalha com multicamadas e resolve problemas “não-linearmente separáveis”.

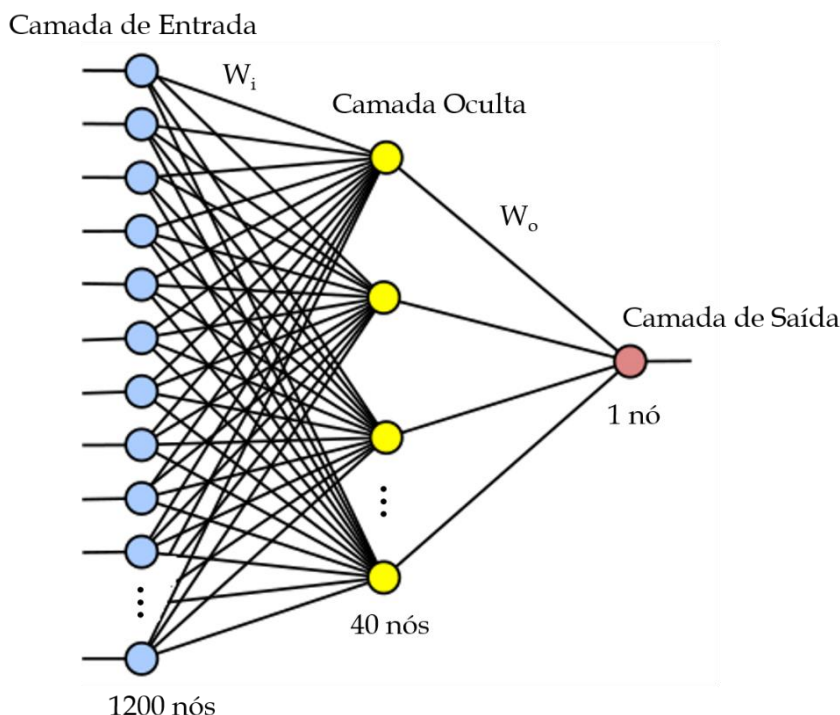


Figura 4 – Rede Neural Artificial (RNA) utilizada no trabalho.

Fonte: Elaborada pelo autor (2020).

4.1. Algoritmo *Backpropagation*

Sua principal característica é a transmissão direta, ou seja, os dados são inseridos pela camada de entrada, e processados pela camada oculta, chegando à camada de saída. Se a

camada de saída não obtiver o valor esperado, o algoritmo faz o processo inverso, para realizar o nivelamento dos pesos (Qi, 2017).

Os dados de erro retornam ao longo das camadas e os pesos e limites de cada camada são ajustados, sucessivamente, até que a saída se aproxime do valor esperado, ou alcance os critérios estabelecidos para a parada.

O número de nós na camada oculta tem uma grande influência na precisão do reconhecimento. Caso o número de nós seja muito pequeno, o tempo de treinamento será curto, o que resultará em uma baixa precisão; no entanto, caso o número de nós seja muito grande, o tempo de treinamento aumentará, o que levará ao ajuste excessivo da rede (Yang, 2019). Como o objetivo deste trabalho é mostrar o uso do algoritmo de bordas Canny, como uma ferramenta que pode ser utilizada para o reconhecimento facial, foi utilizada uma RNA simples, apenas como estudo de caso, não sendo, portanto, foco principal deste.

4.2. Base de Dados: Treino, Validação e Teste da Rede Neural.

A metodologia proposta na seção 3 foi aplicada sobre a base de dados *BioID Face Database* (BioID, 2019). Essa base foi gravada e publicada para oferecer à pesquisadores que trabalham na área de detecção de faces, a possibilidade de comparar a qualidade de seus algoritmos de detecção. Pode ser utilizada para esses fins sem permissão adicional. A base consiste em 1521 imagens, já em escala de cinza, com resolução 384x286 pixels.

Na base há imagens de diferentes indivíduos com visão frontal. No entanto, apesar da grande quantidade de imagens, a base não apresenta quantidades expressivas de imagens de um mesmo indivíduo, o que acaba por reduzir os testes a algumas dezenas de imagens de uma mesma pessoa.

Para validar a proposta apresentada neste trabalho, foram utilizadas 85 imagens padrões, divididas em dois conjuntos, o de treinamento e o de testes, com proporções de 75% e 25%, respectivamente. Dos 75% do conjunto de treinamento, 25% foram utilizadas como conjunto de validação.

Foram utilizados, como dados de entrada na Rede Neural, os vetores binários, de 1200 posições (Equação 4), resultantes do processo de tratamento da imagem descrito na Seção 3. Para limitar o número de buscas por melhores resultados, foram estipuladas 50 execuções como critério de parada. Esse procedimento foi adotado, pois caso não se encontre um resultado satisfatório, este critério, fará com que o algoritmo não fique sendo executado infinitamente.

4.3. Métricas de Avaliação

Os resultados obtidos foram avaliados utilizando as métricas de avaliação características das tarefas de classificação. Foram calculadas as métricas acurácia, *recall*, precisão e *F1-score* (Martins, 2019).

A acurácia é definida como sendo a razão entre o número de previsões corretas e o número total de amostras utilizadas. Já a precisão é calculada como o número de resultados positivos corretos dividido pelo número de resultados positivos previstos pelo classificador. O *Recall* é definido como sendo o número de resultados positivos corretos dividido pelo número de todas as amostras relevantes (todas as amostras que deveriam ter sido identificadas como positivas). Tanto para a precisão quanto para a *recall*, quanto mais próximo de (um) 1, melhor é a representação do modelo.

Já o *F1-score* é a média harmônica entre a precisão e o *recall*. O intervalo previsto para o F1 é [0,1]. Essa métrica informa a precisão do classificador (quantas amostras ele classifica corretamente), bem como, o quanto ele é robusto. Quanto maior a pontuação de *F1-score*, melhor é o desempenho do classificador (Martins, 2019).

5. RESULTADOS

Para validar os resultados obtidos com a utilização do algoritmo bordas Canny, no reconhecimento facial, utilizando como estudo de caso, uma rede neural, foram realizados testes, com o objetivo de se definir a quantidade de neurônios que deveriam ser utilizados, já que essa quantidade pode influenciar no desempenho da RNA. Para isso, foram realizados testes com uma RNA onde o número de neurônios na camada oculta foi variado de 30 a 100, alternando de 10 em 10. Para validar a proposta foi utilizado o filtro Gaussiano, onde foram analisadas as variáveis como, número de neurônios, validação (o esperado é que fique acima dos 90%) e base de teste (variação do número de imagens do alvo na referida base), para se definir as características da Rede Neural, e a Equalização de Histograma, para analisar o comportamento da Rede utilizando um filtro diferente do avaliado.

5.1. Filtro Gaussiano.

Os testes iniciais foram realizados no intuito de validar a RNA que estava sendo utilizada para a identificação. Nos resultados apresentados na Figura 5, pode-se perceber que a Rede Neural possui bons percentuais de acertos, ficando acima dos 90%, chegando em 100% em alguns casos. Apenas em duas situações a RNA não apresentou bons resultados utilizando 30 e 90 neurônios, pois a validação ficou inferior a 70%.

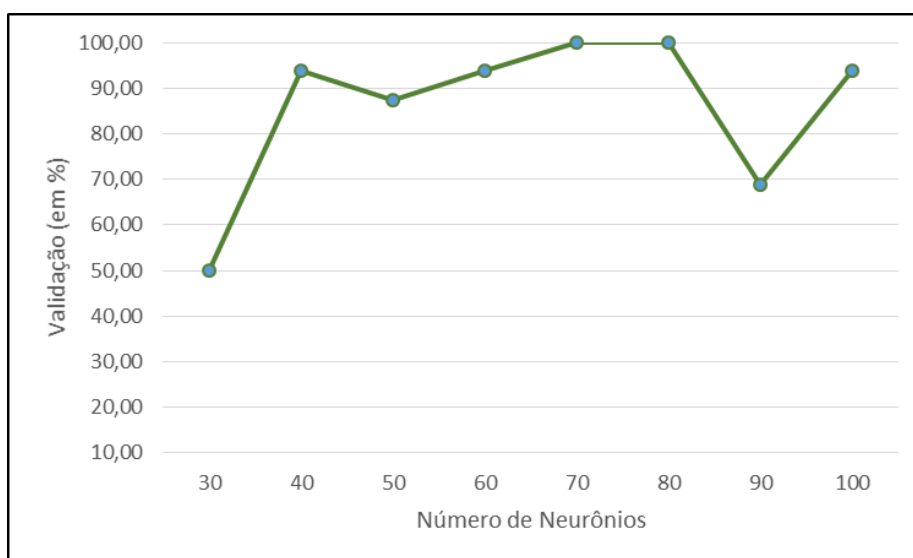


Figura 5 – Validação da Rede Neural Artificial (RNA) com o Filtro Gaussiano.

Fonte: Elaborada pelo autor (2020).

O segundo teste realizado foi para se definir a quantidade de neurônios que seriam utilizados na camada oculta. Dessa forma, foram avaliados a quantidade de acertos (os chamados, verdadeiros-positivos), bem como, a quantidade de erros (os chamados, falsos-positivos) retornados. Neste caso, o objetivo é maximizar os acertos e minimizar os erros, com isso, não basta a RNA encontrar todas as imagens do alvo utilizadas, precisa-se, também, minimizar os falsos-positivos para que tenha uma taxa de acerto de sucesso.

O gráfico da Figura 6 apresenta os resultados encontrados após a etapa de teste para o conjunto de neurônios especificados. Pode-se perceber que, para alguns casos, o número de acertos (verdadeiro-positivo) ficou acima de 80%, como é o caso de 40, 70 e 80 neurônios na camada oculta. No entanto, como citado anteriormente, não basta a rede ter uma alta taxa

de acertos, é aconselhável, também, que a mesma tenha uma baixa taxa de erros, ou seja, retorne um reduzido número de falsos-positivos.

Dessa forma, quando a camada oculta possui 70 neurônios, o índice de falsos-positivos fica acima de 40%, passando de 80%, quando a rede possui 80 neurônios. Percebe-se, pela Figura 6, que quando a rede possui 40 neurônios na camada oculta, a taxa de acerto é similar quando a mesma é treinada com 70 ou 80 neurônios. A principal vantagem é que com 40 neurônios ela retorna uma menor taxa de arquivos falsos-positivos, próximo a 10%, além de ser executada em um menor tempo, também.

O terceiro teste realizado foi na base de teste. A ideia foi variar a quantidade de imagens do alvo na base de teste com o objetivo de verificar a quantidade de acertos (os chamados, verdadeiros-positivos), assim como, a quantidade de erros (os chamados, falsos-positivos). A variação se deu em torno de 30% até 60% de imagens do alvo na base de teste. O resultado pode ser visualizado na Figura 7.

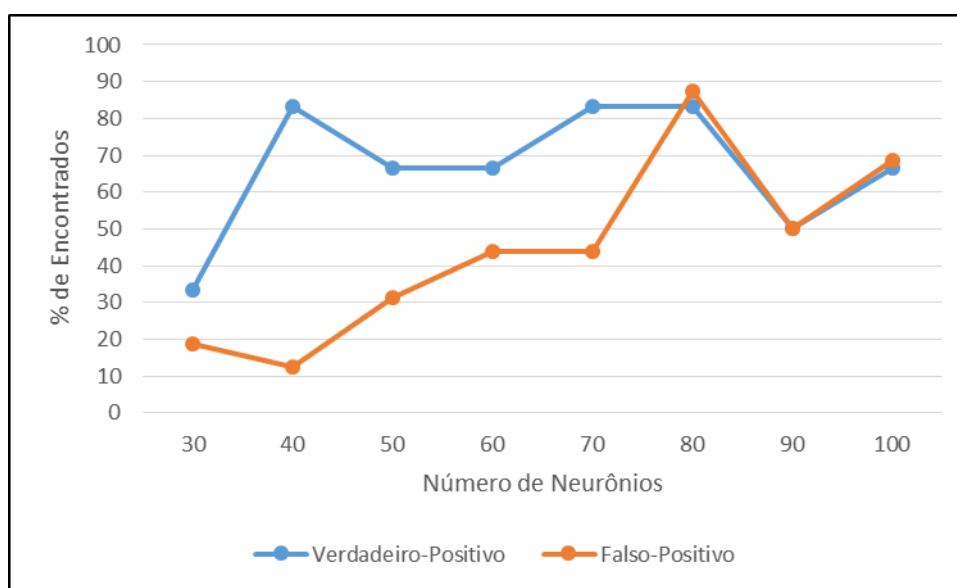


Figura 6 – Testes da Rede Neural Artificial (RNA) com o Filtro Gaussiano.

Fonte: Elaborada pelo autor (2020).

Na Figura 7, pode-se perceber que há um aumento no número de imagens corretas encontradas (verdadeiros-positivos) conforme se aumenta o número de imagens do alvo na base de dados, por outro lado, perceber, também, um leve aumento no número de falsos-positivos até próximo a equidade da base de testes (50% para imagens do alvo e 50% de imagens diferentes do alvo), onde se percebe uma queda do número de falsos-positivos. Isso pode ser justificado pois, passou-se a ter mais imagens do alvo, na base de teste, do que imagens diferentes do alvo.

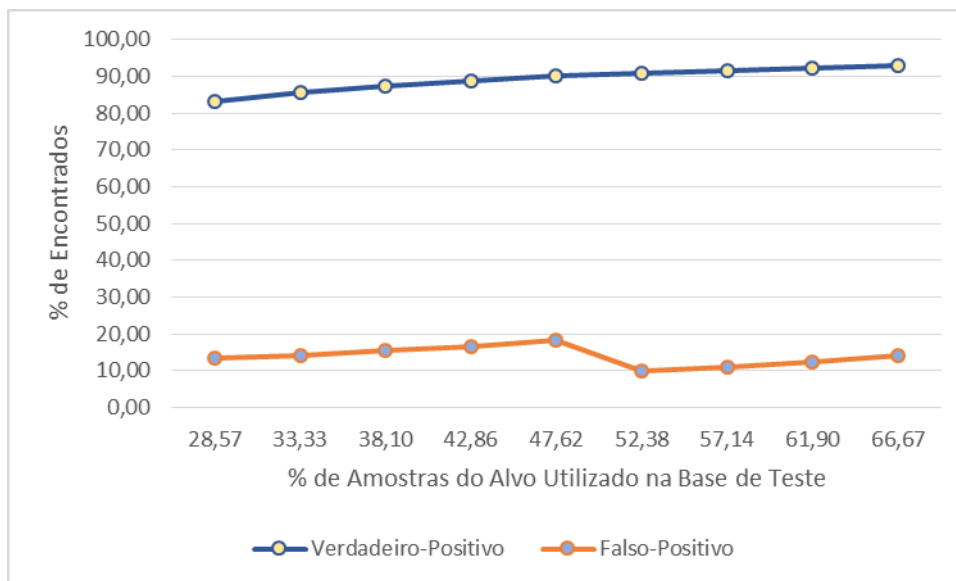


Figura 7 – Variação do Número de Imagens do Alvo na Base de Teste.

Fonte: Elaborada pelo autor (2020).

Diante dos resultados apresentados nas Figuras 5 e 6, foi definida uma Rede Neural com uma camada oculta de 40 neurônios para a validação do algoritmo bordas Canny, na proposta de reconhecimento facial utilizado o Filtro Gaussiano. Para este resultado foram utilizadas, em média, 28% de imagens do alvo na base de teste. Os resultados podem ser melhor detalhados na matriz confusão apresentada na Tabela 1.

Tabela 1 – Matriz Confusão (em porcentagens) para o Filtro Gaussiano com 28% de Imagens do Alvo na Base de Teste.

		Referência	
		É o alvo	Não é o alvo
Classificação	É o alvo	83,33%	12,50%
	Não é o alvo	16,67%	87,50%

Pela Tabela 1, pode-se perceber que a taxa de acerto está acima de 80%, ou seja, a rede conseguiu identificar corretamente, tanto as imagens que eram do alvo, como as que eram diferentes. No entanto, caso o número de imagens do alvo passe de 28% para 47%, a taxa de acertos fica em torno de 90%, como pode ser observado na Tabela 2.

Tabela 2 – Matriz Confusão (em porcentagens) para o Filtro Gaussiano com 47% de Imagens do Alvo na Base de Teste.

		Referência	
		É o alvo	Não é o alvo
Classificação	É o alvo	90,00%	18,18%
	Não é o alvo	10,00%	81,81%

5.2. Filtro Equalização de Histograma.

A Rede Neural utilizada neste filtro foi a mesma definida para o filtro Gaussiano, onde a validação seguiu o que foi definido na Figura 5. Após a validação, definiu-se o uso de 40 neurônios na camada oculta, por essa configuração alcançar uma taxa superior a 90% de acertos.

Pela Figura 8, percebe-se que o percentual de retorno de falsos-positivos teve uma queda, apresentou índice de 90%, com o aumento do percentual de imagens do alvo na base de teste (passando de 28,57 para 47,62%), no entanto, o percentual de encontrados de forma correta se manteve nos 100%, ou seja, a RNA conseguiu identificar todas as imagens do alvo durante os testes, independente do percentual de imagens do alvo utilizadas nesta fase. Porém, um ponto merece destaque, o número de falsos-positivos identificado. E isso pode ser explicado pelas próprias características do filtro utilizado.

A equalização faz com que a distribuição das intensidades dos pixels de 0 a 255 seja o mais homogêneo possível. Dessa forma, tem-se a mesma quantidade de pixels com valores na faixa de 0 a 10 (pixels muito escuros) e na faixa de 245 a 255 (pixels muito claros). Isso faz com que o reconhecimento facial obtenha uma excelente taxa de acerto (verdadeiro-positivo), mas também, uma grande quantidade de erro (falsos-positivos). Na Figura 8, pode-se visualizar os resultados encontrados variando o número de imagens do alvo na base de teste, para o filtro Equalização de Histograma.

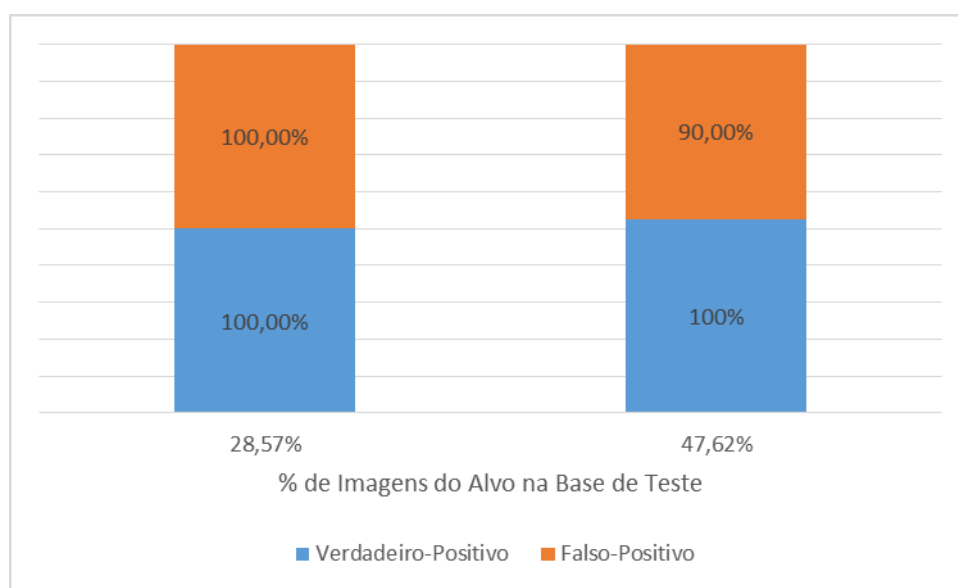


Figura 8 – Variação do Número de Imagens do Alvo na Base de Teste para o Filtro Equalização de Histograma.

Fonte: Elaborada pelo autor (2020).

Percebe-se, pela Figura 9, que a Rede Neural, utilizando o filtro Equalização de Histograma, conseguiu localizar um percentual significativo de imagens do alvo, ou seja, chegou a localizar 100% das imagens utilizadas na base de teste. Já utilizando o filtro Gaussiano, a taxa de acerto passou de 83% para 90%. Isso demonstra que a rede consegue identificar as imagens independente do filtro utilizado. No entanto, o percentual de acertos irá depender das características de cada filtro.

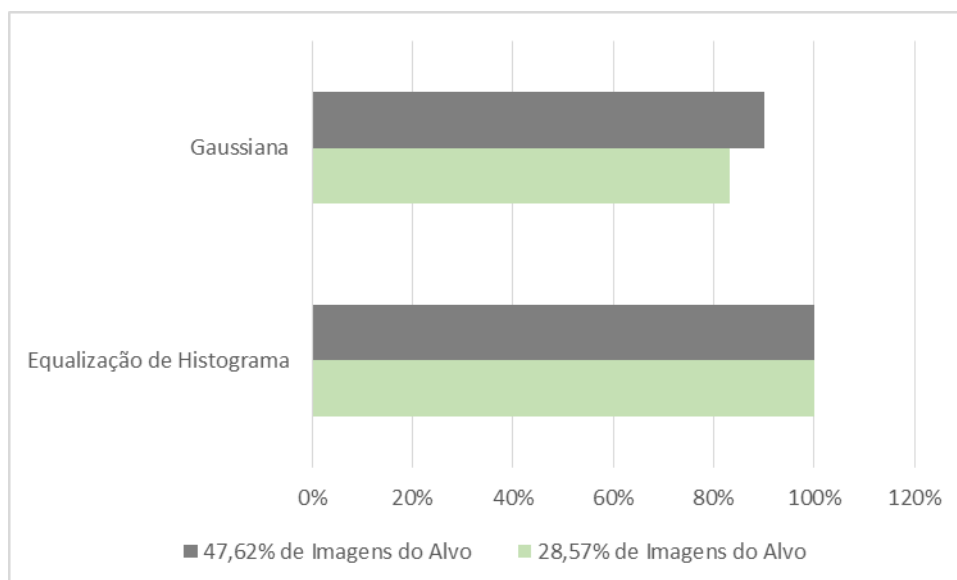


Figura 9 – Filtro Gaussiano \times Filtro Equalização de Histograma. Porcentagem de verdadeiros-positivos encontrados.

Fonte: Elaborada pelo autor (2020).

A Tabela 1 apresenta os resultados das métricas utilizadas comparando os dois filtros utilizados, com 28% de imagens do alvo na base de teste. Nota-se que o filtro Gaussiano apresenta os melhores resultados em termos de todas as métricas utilizadas, com exceção do *recall*. Quando o número de imagens do alvo, na base de teste, é alterado para 47%, os resultados podem ser observados na Tabela 2.

Tabela 1 – Métricas de Avaliação da Classificação Utilizando 28% de Imagens do Alvo na Base de Testes.

Métrica	Gassiano	Equalização
Acurária (%)	85,71	28,571
Precisão	0,714	0,286
Recall	0,833	1,000
F1-score	0,769	0,444

Percebe-se que, mesmo aumentando o número de imagens do alvo na base de teste, o filtro Gaussiano ainda apresenta os melhores resultados em todas as métricas utilizadas, com exceção, também, do *recall* e da acurácia, que não há alteração. Como citado anteriormente, o *recall* quanto mais próximo de 1, melhor é o modelo. A Equalização de Histograma apresenta os melhores resultados, no *recall*, devido realizar a localização total das imagens na base de teste, no entanto, acaba por retornar um grande número de falsos-positivos, o que o faz ter um desempenho menos satisfatório nas outras métricas do que o filtro Gaussiano.

Tabela 2 – Métricas de Avaliação da Classificação Utilizando 47% de Imagens do Alvo na Base de Testes.

Métrica	Gassiano	Equalização
Acurária (%)	85,71	52,381
Precisão	0,818	0,500
Recall	0,900	1,000
F1-score	0,857	0,667

Como o principal objetivo do trabalho foi apresentar uma proposta para identificação de imagens faciais, tendo como base, o algoritmo das bordas Canny e o uso comparativo de dois filtros, com a finalidade de demonstrar a eficiência da RNA, mesmo com diferentes tipos de imagens, as quantidades retornadas, por filtro, foram suficientes para validar a proposta com uma taxa de acerto entre 83% e 90% para o filtro Gaussiano, e chegando a 100% no caso do Equalização de Histograma.

5.3. Discussão dos Resultados

Apesar da limitação da quantidade de imagens disponibilizadas, por indivíduo, na base de teste, os resultados mostram que o algoritmo de Canny pode ser utilizado para realizar reconhecimento facial. Nesta primeira avaliação, o principal objetivo é demonstrar a usabilidade do algoritmo, e isso foi possível, através do estudo de caso nas Redes Neurais. No entanto, fica explícito que, para uma melhora significativa dos resultados, neste mesmo cenário (Redes Neurais), há a necessidade de se melhorar a quantidade de imagens para poder realizar um estudo mais aprofundado e, conseqüentemente, propor melhorias para a estrutura da Rede Neural.

As métricas mostraram que, em dois dos cenários apresentados, a precisão chegou a ser superior a 83%. Isso implica concluir que o algoritmo conseguiu reconhecer o alvo, mesmo diante das limitações da quantidade de imagens utilizadas nos testes. Pode-se avaliar também, que entre os filtros utilizados o Gaussiano se mostrou mais eficiente quando comparado ao filtro Equalização de Histograma, como demonstrado nas métricas apresentadas nas Tabelas 1 e 2. No entanto, é relevante citar, que o objetivo deste trabalho, foi demonstrar que a metodologia proposta, pode ser utilizado para o reconhecimento de imagens, e que o uso de Redes Neurais, neste processo, pode ser, também, uma solução, mas que precisa ser ajustada e melhorada, para este fim, o que não foi o foco deste trabalho.

6. CONCLUSÃO

O uso de filtros para realizar o tratamento de imagens, passou de uma necessidade comercial ou científica, para um emprego no dia-a-dia, principalmente, em se tratando de redes sociais, onde as pessoas os aplicam para melhorar as imagens postadas. E a combinação desses tratamentos de imagens, com técnicas como, a Inteligência Artificial, pode resultar em ferramentas, robustas e úteis, para as mais variadas finalidades.

Neste trabalho, foi apresentado o uso de filtros, escala de cinza e segmentação, e o algoritmo de Canny, juntamente com uma Rede Neural, para realizar o reconhecimento facial. Nos resultados apresentados, pode-se perceber que a proposta foi bem-sucedida, pois houve a detecção do alvo em todos os cenários propostos. No entanto, para se obter melhores resultados, como por exemplo, a redução de falsos positivos e uma melhor taxa de

verdadeiros positivos, pode-se ajustar a Rede Neural aplicada à este trabalho, ou utilizar outras técnicas no reconhecimento de imagens.

Vale ressaltar que, a metodologia proposta neste trabalho, pode ser utilizada não apenas para o reconhecimento de faces, mas também para o reconhecimento de imagens em geral, como por exemplo, para reconhecer tipos e espécies de flores ou folhas, grafia de letras, além de outros tipos de imagens.

Como trabalhos futuros, pode-se realizar melhorias, tanto na execução do algoritmo, quanto na Rede Neural utilizada, que não era o principal foco deste trabalho. Pode-se, também, melhorar a precisão no algoritmo, além de realizar testes utilizando outros filtros, com o objetivo de se realizar um estudo comparativo entre os mesmos, bem como, utilizar outras imagens para analisar o comportamento da Rede.

Na Rede Neural, pode-se aumentar o número de camadas ocultas, para se avaliar o treinamento e com isso realizar possíveis ajustes nos pesos. Esse aperfeiçoamento na rede, pode possibilitar o uso da metodologia em uma base de teste maior, aumentando possivelmente a precisão no reconhecimento facial. Além disso, o uso de outras técnicas de Inteligência Artificial pode ser válido, como Algoritmo Genético e Máquina de Vetores de Suporte, por exemplo, de forma a definir qual delas apresenta melhor desempenho na metodologia proposta.

REFERÊNCIAS BIBLIOGRÁFICAS

BioID FACE DATABASE (2019). Disponível em: <https://www.bioid.com/facedb/>. Acessado em junho de 2019.

CANNY, J. F., (1986). A Computational Approach for Edge Detection. *IEEE Trans. Pattern Analysis and Machine Intelligence* 08, 679–698.

DANG, K. and SHARMA, S. (2017). Review and comparison of face detection algorithms. 7th International Conference on Cloud Computing, Data Science Engineering – Confluence. 12-13 Jan. Noide, Índia. <https://doi.org/10.1109/CONFLUENCE.2017.7943228>

GEISSENBUHLER, D., NIKISINS, O., ANJOS, A., GEORGE, A., MOSTAANI, Z. and MARCEL, S. (2020). Biometric Face Presentation Attack Detection With Multi-Channel Convolutional Neural Network. *IEEE Transactions on Information Forensics and Security* 15, 42–55.

KOLB, A., STEINER, H., SPORRER, S. and JUNG, N. (2016). Design of an active multispectral SWIR camera system for skin detection and face verification. *Journal of Sensors*, 1–16. <https://doi.org/10.1155/2016/9682453>

MACARINI, L. A. and SOBIERANSKI, A. C., (2019). Using Convolutional Neural Network to Detect and Count Individuals on Eucalyptus Plantation. *X Computer on the Beach*, pp. 599–608. Florianópolis - SC.

MARENGONI, M., and STRINGHINI, S. (2009). Tutorial: Introdução à Visão Computacional usando OpenCV. *Revista de Informática Teórica e Aplicada XVI*, Número 1, pp. 125–160.

MARTINS, L. R., MURAMATSU Jr., M. e SERAPIÃO, A. B. S., (2019). Image classification of Kuzushiji ideograms with Convolutional Neural Networks. XVI Encontro Nacional de Inteligência Artificial e Computacional (ENIAC). pp. 309-320. Salvador – BA.

METZGER, G., PLICHOSKI, G. F., and CHIDAMBARAM, C. (2018). A Supervised Face Recognition in Still Images using Interest Points. IX Computer on the Beach, pp. 721–730.

MINICHINO, J. and HOWS, J. (2015). Learning OpenCV 3 Computer Vision with Python. Packt Publishing Ltd, Birmingham. 242 pages.

NASCIMENTO, A., COUTO, M., OLIVEIRA, D. and PAULA, JR. I. (2018). Facial Landmarks Detection on Faulty Datasets with Regression Trees and Principal Component Analysis Parametrization. XV Encontro Nacional de Inteligência Artificial e Computacional, pp. 343–352. <https://doi.org/10.5753/eniac.2018.4429>

QI, L. (2017). Research on Non-specific Speech Recognition Algorithm Based on Sigmoid Learning Rate BP Neural Network. Value Engineering, pp. 200–202.

ROYCHOWDHURY, A., ERIC, H. L, LEARNED-MILLER, HUANG, G. B. and HUA, G. (2016). Labeled faces in the wild: A survey. Advances in Face Detection and Facial Image Analysis. Cham, Switzerland: Springer, pp. 189–248.

SHANMUGAVADIVU, P. and KUMAR, A. (2016). Rapid face detection and annotation with loosely face geometry. 2nd International Conference on Contemporary Computing and Informatics, 14-17 Dec. Noida - Índia. <https://doi.org/10.1109/IC3I.2016.7918032>.

SINGH, M., SINGH, A. and SINGH, B. (2016). Face detection and Eyes extraction using Sobel Edge Detection and Morphological Operations. Conference on Advances in Signal Processing (CASP), 295–608.

SOUZA, D. A., SOUZA, E. T. E., SANTOS, F. J., MACEDO, M da S., LIBERA Jr., V. D. L. (2019). Redes Neurais Artificiais: Identificação de Danos em Elementos Estruturais. VI Jornada Interdisciplinar de Engenharia Civil. Portal de Periódicos Eletrônicos.

SOUZA, T. e CORREIA, S. (2007). Estudo de Técnicas de Realce de Imagens Digitais e suas Aplicações. II Congresso de Pesquisa e Inovação da Rede Norte Nordeste de Educação Tecnológica (CONNEPI). João Pessoa – PB.

STEINBACH, M., TAN, P., and KUMAR, V. (2009). Introdução ao Data Mining. Vol. 1ª Edição. Editora Ciência Moderna.

YANG, N. and SHI, F. (2019). Speech Emotion Recognition Based on Back Propagation Neural Network. Information Technology and Intelligent Transportation Systems, 216–223. <https://doi.org/10.3233/978-1-61499-939-3-216>.