

Uma Abordagem para Reconhecimento de Placas de Perigo para Pessoas com Deficiência Visual

Jailson Lucas Panizzon
Universidade Tecnológica Federal do Paraná
Dois Vizinhos, PR, Brasil
jailsonpanizzon@alunos.utfpr.edu.br

Alinne Cristinne Correa Souza
Universidade Tecnológica Federal do Paraná
Dois Vizinhos, PR, Brazil
alinnesouza@utfpr.edu.br

Gabriel Lopes Guilherme
Universidade Tecnológica Federal do Paraná
Dois Vizinhos, PR, Brasil
gabrielguilherme@alunos.utfpr.edu.br

Francisco Carlos Monteiro Souza
Universidade Tecnológica Federal do Paraná
Dois Vizinhos, PR, Brasil
franciscosouza@utfpr.edu.br

ABSTRACT

According to the Brazilian Institute of Geography and Statistics, there are about 6.5 million visually impaired in Brazil. A frequent issue is the insertion of them in hazardous environments signaled by signs only. Solutions must be produced to improve the life quality of visually impaired whether at work or in public places. In this sense, the present paper presents an application of danger sign recognition aimed at people with visual impairment through a supervised learning classifier called Haar Cascade. Furthermore, the paper reports preliminary results regarding effectiveness of the proposed system applied in a set of images. Overall, for the majority of the samples, the proposed system achieved a promising score and this can be a good alternative for studies that aim visually impaired inclusion in dangerous environments.

KEYWORDS

Visão Computacional, Reconhecimento de Padrões, Aprendizado de Máquina

1 INTRODUÇÃO

De acordo com o Instituto Brasileiro de Geografia e Estatística [1], existem mais de 6,5 milhões de deficientes visuais no Brasil, destes, 582 mil são pessoas completamente cegas e seis milhões possuem baixa visão.

O Brasil, também possui uma lei que exige a contratação de pessoas com deficiência em empresas com mais de 100 funcionários, como diz o texto: A empresa com 100 ou mais empregados é obrigada a ter de 2% a 5% dos seus cargos preenchidos por pessoas portadoras de deficiência, de acordo com o Art. 93 da Lei nº 8.213, de 24 de julho de 1991 [2].

Por ser inevitável a existência de lugares amplos, a localização torna-se um problema. Diferentes formas de sinalização visam orientar as pessoas, repassando informações e sinais importantes, para que seja facilitada a mobilidade no local.

Neste contexto, uma situação preocupante em um ambiente laboral com vários riscos sinalizados somente por placas de advertência. Além de outros ambientes a qual está presente no cotidiano que só utilizam de advertências visuais de perigos existentes.

Para minimizar esses problemas, uma possível solução seria o uso de uma aplicação que realize a análise do ambiente e informe à pessoa sobre os riscos presentes. Segundo Alexandria [3], várias

aplicações como: detecção, reconhecimento e classificação de objetos; estimativa de parâmetro como tamanho, posição, orientação e velocidade utilizam sistemas de Visão Computacional para auxiliar na solução desse tipo de problema.

Este artigo visa o desenvolvimento de um sistema para o reconhecimento de perigo a partir de imagens obtidas com uma câmera e indicação da presença da placa para o usuário por meio de um alerta utilizando alto-falante. Para avaliar o sistema proposto foi realizado um experimento com 643 imagens diferentes utilizando as métricas de precisão, revocação e F1-Escore.

No geral, as contribuições do presente trabalho podem ser resumidas nos seguintes pontos:

- (i) um sistema para auxiliar as pessoas com deficiência visual na identificação de possíveis riscos apresentados por meio das placas perigosas;
- (ii) criação de um banco de imagens normalizadas e mapeadas;
- (iii) avaliação empírica da eficácia do sistema proposto.

Este artigo está organizado da seguinte forma: na Seção 2 são apresentados os principais conceitos relacionados ao estudo. Na Seção 3 são apresentados os trabalhos relacionados. Na Seção 4 é detalhado o sistema de reconhecimento de placas para deficientes visuais. Na Seção 5 são apresentadas a definição, o design e o processo de execução do experimento. Na Seção 6 são apresentados e discutidos os resultados alcançados. Por fim, as considerações finais e trabalhos futuros são apresentados na Seção 7.

2 REFERENCIAL TEÓRICO

Esta seção apresenta os principais conceitos relacionados ao reconhecimento de padrões e processamento de imagens.

2.1 Reconhecimento de padrões

Reconhecimento de padrões é uma área da ciência que visa identificar características que são comuns em grupos de elementos. Na área de visão computacional Gonzales e Wood [4] define característica como um arranjo de descritores. As características obtidas são usadas como parâmetro para identificar a pertinência de um novo elemento ao grupo. Para pertencer ao um grupo o elemento deve preencher requisitos mínimos de similaridade de características com o seu grupo. Dependendo da metodologia existem regras diferentes para os descritores onde um elemento pode ser encaixado em

um grupo ou o mesmo pode receber um grau de pertinência em vários grupos.

A grande parte das metodologias de reconhecimento de padrões tem grande base matemática e estatística. Um conceito importante quando se discute reconhecimento de padrões é capacidade do modelo aprender com os dados fornecidos [4], tendo em vista essas necessidades é comum os sistemas de reconhecimento de padrões serem apoiados por técnicas de Aprendizado de Máquina (AM).

2.2 Aprendizado de máquina

AM é uma área da inteligência artificial que visa simular em uma máquina a capacidade humana de aprender. Dada uma classe de performance mensurável, um programa aprende com uma experiência, se a performance da experiência está dentro da classe de performance mensurável definida, e o aprendizado está sendo melhorado pela experiência [5]. Uma técnica de aprendizado de máquina pode ser dada de três maneiras:

- **aprendizado supervisionado:** o sistema possui as entradas para o treinamento com rótulos que definem a qual classe pertence aquele dado, e baseado nos rótulos o sistema pode identificar descritores para cada dados e agrupar esses descritores formando vetores de características, que serão vinculados a classe do rótulo fornecido, como uma característica extraída;
- **aprendizado não-supervisionado:** o sistema possui as entradas para o treinamento porém sem rótulos. O sistema por si só é responsável por identificar descritores e estabelecer ligações entre os vetores de características obtidas e estabelecer as classes. Na maioria das vezes o sistema irá estabelecer graus de pertinência de um objeto dado as classes estabelecidas;
- **aprendizado semi-supervisionado ou por reforço:** é uma junção dos itens anteriores onde o sistema pré-treinado com entradas iniciais rotuladas, ao interagir com o ambiente irá obter *feedback* para usar como parâmetro para reavaliar seus vetores de características.

2.3 Processamento de imagem digital

Uma imagem pode ser definida como uma função bidimensional, em que x e y são as coordenadas espaciais e a amplitude de f em qualquer par de coordenadas (x, y) é denominada intensidade ou nível de cinza da imagem nesse ponto. Quando x , y e os valores de intensidade de f são quantidades finitas e discretas, chamamos a imagem de imagem digital. Com o processamento de imagens digitais por meio de um computador digital, tem-se o processamento de imagem digital.[4]

A área de processamento digital de imagens traça uma linha de níveis de processo que vai de operações mais primitivas, como pré-processamento de imagens, até operações mais complexas de análise de imagens. O campo da visão computacional está situado no extremo dessa linha com operações de alto nível que fazem a análise das informações contidas na imagem visando dar sentido as informações fornecidas [4].

2.3.1 Correção de histograma. A segmentação de imagens baseada no histograma da imagem é chamada de correção do histograma. O histograma de uma imagem é uma função discreta que define os

intervalos de intensidade de cor na imagem. A nitidez da imagem está relacionada diretamente com o nível dos intervalos, quanto maior a diferença, maior será a nitidez e as cores diferentes serão diferenciadas com uma maior percepção. Então, a correção de histograma se baseia em analisar, selecionar e agrupar os intervalos, para que haja uma diferença maior entre seus vizinhos [4].

2.4 Classificadores

No aprendizado de máquina, os classificadores são uma abordagem de aprendizado supervisionado, na qual o programa aprende com a entrada de dados fornecida e, em seguida, usa esse aprendizado para classificar novas observações [6]. Dentre os diferentes tipos de classificadores é importante destacar o *Haar cascade*.

O *Haar Cascade* é um classificador que possui uma cascata de classificadores impulsivos trabalhando com *haar-like features*, que são recursos de imagem digital usados no reconhecimento de objetos. É treinado com centenas de imagens de um objeto específico, chamados de imagens positivas, e imagens negativas que são imagens arbitrárias do mesmo tamanho.

O processo do *Haar Cascade* consiste em um conjunto de etapas em que o algoritmo reconhece um objeto por meio de uma janela em movimento sobre a imagem, aplicando vários classificadores fracos em cascata na região formada pela janela. Cada classificador fraco representa uma característica extraída no processo de treinamento. O processo de classificação utilizado pelo *Haar Cascade* é demonstrado na figura 1. Em cada etapa, os classificadores inserem rótulos de positivo ou negativo em uma determinada região definida pelo local atual da janela, se um classificador retornar positivo o próximo classificador é aplicado na janela, se um classificador retornar negativo a região é dada como negativa e a janela é movida para o próximo ponto. Para uma região ser dada como positiva todos os classificadores devem retornar positivo para aquela região.

Para um objeto ser identificado é preciso se atingir um mínimo de janelas positivas vizinhas sobre a região do objeto.

Após o classificador ser treinado, o mesmo pode ser aplicado a uma imagem de entrada do mesmo tamanho usado no treinamento. Após a imagem de entrada passar pelo classificador, ele retornará "1" se a imagem de entrada for reconhecida utilizando a base de dados do treinamento, e retornará "0" se a imagem não for reconhecida [7].

De acordo com Souza [8] um classificador pode ser avaliado em termos da sua capacidade de generalização, ou seja, presumir corretamente a classe de novos dados que não foram ensinados ao modelo. Quando o modelo treinado é capaz de classificar corretamente dados que não foram apresentados a ele, isso significa que a técnica de AM conseguiu aprender corretamente, porém se ele apresenta uma baixa taxa de acerto, ou os exemplos passado a ele pode não ter sido suficientemente bons ou a técnica pode ter aprendido fortemente o conjunto de treinamento e não consegue generalizar quando é confrontado com novos dados.

3 TRABALHOS RELACIONADOS

A motivação deste artigo se baseia no trabalho de Sousa e Marengoni [9], que desenvolveram um sistema para o reconhecimento de faixa de pedestres para deficientes visuais, com o uso de dispositivos móveis. Nesse reconhecimento, o usuário precisa fotografar o local,

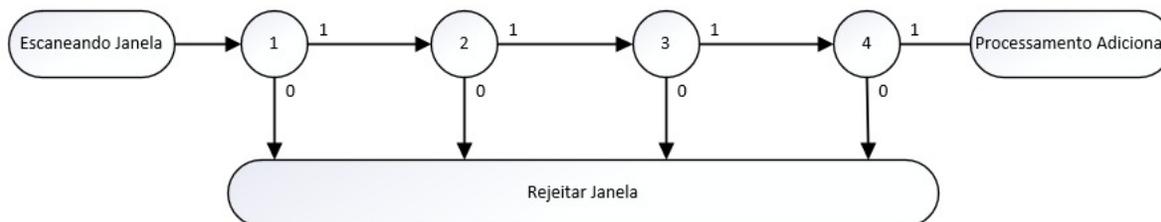


Figure 1: Visão geral do funcionamento do classificador *Haar Cascade*

seja com o toque ou por comando de voz, e depois do local ser fotografado, é dado um retorno de voz para o usuário se no local fotografado existe uma faixa de pedestre.

A maior contribuição vinda desse trabalho de Sousa e Marengoni [9], é a ajuda social para os deficientes visuais que não conseguem perceber os perigos do cotidiano, podendo utilizar de inteligência artificial para auxiliá-los em sua vida.

Outro trabalho a ser destacado é o de Gomes, Rebouças e Pedrosa Filho [10], que desenvolveu um sistema de visão computacional para a detecção e o reconhecimento de placas de regulamentação de trânsito. Este trabalho trouxe a esse projeto uma inspiração mais prática de como alcançar o objetivo proposto. Duas etapas foram fundamentais, a primeira é a segmentação, onde o sistema detecta a existência de uma placa de regulamentação e a outra etapa é a de reconhecimento de caracteres, onde o sistema reconhece os caracteres da placa.

Todos os trabalhos têm a função de reconhecimento de imagem para alertar um usuário. No entanto, nenhum conseguiu alcançar um desempenho maior que 80%, um deixando uma recomendação para utilizar um sistema mais eficiente de detecção e o outro para a continuação do mesmo projeto.

4 SISTEMA DE RECONHECIMENTO DE PLACAS DE PERIGO

Nesta seção serão delineados os detalhes do sistema de reconhecimento de padrões que identifica placas de perigo para ajudar pessoas com algum tipo de deficiência visual. Conforme pode ser visto na Figura 2, o sistema é composto por três fases: (i) obtenção das imagens; (ii) identificação e processamento da placa; e (iii) identificação e leitura do texto da placa.

Na primeira fase, "Obtenção da imagem", as placas de perigo são captadas por meio de uma câmera que grava o percurso de uma pessoa com deficiência visual. Na segunda fase ("Identificação e processamento da placa"), as placas de perigo são identificadas e processadas utilizando o classificador *Haar Cascade* da biblioteca *OpenCV 3.2* [7]. Por fim, na terceira fase ("Identificação e leitura do texto da placa"), após a identificação da placa o sistema emite um alerta, por meio de voz, avisando da existência de uma placa de perigo e lendo o seu conteúdo para a pessoa com deficiência visual. Nas subseções a seguir será detalhada cada fase do sistema.

4.1 Obtenção da imagem

As imagens são obtidas na horizontal por uma câmera que converte as imagens analógicas para um padrão digital em tempo real. O dispositivo deve ter um dispositivo de memória cache para armazenar os dados por alguns segundos enquanto aguardam para ser processados. As imagens somente serão armazenadas em tempo de execução.

4.2 Identificação e processamento das placas

As placas identificadas e processadas a partir do arquivo de vídeo são a entrada para a identificação e leitura do texto da placa. Para este processo foi utilizado o classificador *Haar cascade*. Esse classificador é baseado em processo *Adaboost* que usa vários classificadores e aplica-os na forma de cascata da mesma forma que o processo descrito por Viola e Jones [11].

Para treinamento do classificador *Haar Cascade*, por se tratar de um classificador de aprendizado supervisionado, foram necessários dois conjuntos de imagens: (i) positivas; e (ii) negativas. Um exemplo de imagem positiva e negativa são apresentadas na Figura 3a e Figura 3b, respectivamente.

As imagens positivas podem auxiliar o processo de aprendizado do algoritmo, então foi realizado mapeamento das imagens positivas com a quantidade e a posição dos objetos positivos na imagem. Essas informações são adicionadas em um documento juntamente com o diretório das imagens que será passado posteriormente para o classificador usar como base de imagens positivas.

As imagens positivas foram obtidas a partir de mecanismos de busca e bancos de imagens usando a strings "*Placas de perigo*" traduzida nas línguas, francês, espanhol, chinês simplificado, alemão e inglês. O processo de busca das imagens foi realizado de forma manual. Foram obtidas imagens em diferentes posições, contextos e estados de conservação e também imagens contendo mais de um objeto positivo. Após as imagens serem obtidas, as mesmas foram normalizadas, redimensionadas para 690x720 e convertidas para o formato de imagens *Bitmap*.

As imagens negativas devem ser fornecidas de forma a simular situações reais onde o objeto positivo pode estar inserido. As imagens negativas têm a função de servirem como contraparte às imagens positivas fornecidas, uma vez que se uma característica candidata estiver presente em uma imagem negativa então a mesma deve ser reavaliada.

Uma base de dados com 555 imagens positivas e 1245 imagens negativas foi criada para o treinamento do classificador. Para ser considerada um imagem positiva a mesma deveria possuir uma

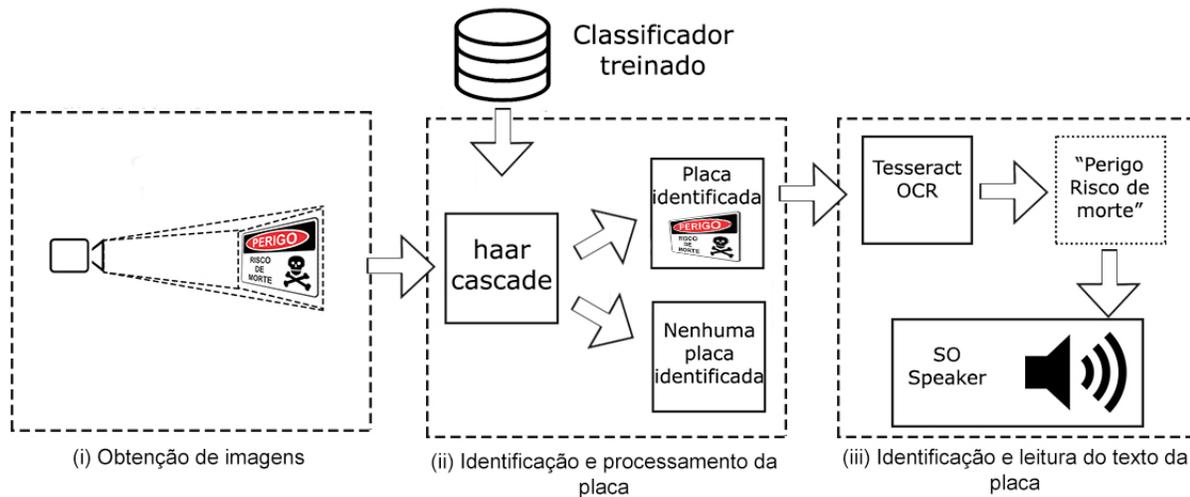


Figure 2: Visão geral do sistema de reconhecimento de placas perigosas para deficientes visuais

placa de perigo que fosse possível de ser identificada por uma pessoa e que segue um padrão mínimos de formato e posicionamento do texto como mostrado na Figura 3a.

É importante destacar que a língua em que o texto escrito na placa de perigo não foi considerado, pois a intenção do classificador é identificar o padrão que normalmente é utilizado em placas de perigo. Dentre as imagens positivas algumas imagens foram invertidas como no trabalho de Bruno et. al [12] para auxiliar no processo de treinamento evitando que o classificador adquira vícios de posição de imagens e, conseqüentemente, ajudando a melhorar a base de dados. O processo de mapeamento das imagens positivas foi realizado com auxílio do *software* gratuito chamado *Dasar Haartrain* [13].

As imagens negativas foram selecionadas considerando dois aspectos: (i) contexto em que o classificador irá atuar como indústrias e ruas; e (ii) não possuir placa de perigo. O classificador *Haar Cascade* utiliza amostras no formato de vetores em tamanhos específicos de acordo com o número máximo de imagens negativas que devem ser criadas antes do treinamento.

Após a criação da base de dados, é iniciado o treinamento passando os seguintes parâmetros: (i) a pasta para salvar os arquivos de treinamento; (ii) amostras criadas; (iii) arquivo com os diretórios das imagens negativas; (iv) número de imagens positivas; (v) número de imagens negativas; (vi) número de estágios; (vii) tamanho das amostras; e (viii) modo de operação. Ao final do processo dentro da pasta informada para salvar os dados haverá um arquivo XML que possui os parâmetros extraídos.

Após a etapa de treinamento foi aplicada uma função da biblioteca *OpenCV* utilizando um *arquivo.xml* treinado para classificar uma imagem. Essa função é executada para cada *frame* do vídeo que terá como retorno as coordenadas caso a placa tenha sido identificada. Antes de passar a imagem para o classificador, foi realizado o pré-processamento fazendo a equalização de histograma de imagem para aumentar a nitidez da imagem e diminuir o contraste de áreas muito escuras ou muito claras.

4.3 Identificação e leitura do texto da placa

A identificação e a leitura do texto da placa foram executada utilizando a biblioteca *Pytesseract* [14]. Essa biblioteca realiza o reconhecimento óptico de caracteres para a linguagem *Python* do módulo de código aberto Tesseract da Google[15].

A *Pytesseract* traz bases de dados já treinadas para reconhecimento de caracteres, nesse projeto foi utilizado a base de dados inglês fornecido pelo próprio módulo Tesseract.

Uma vez que a placa é detectada, a partir das coordenadas fornecidas pelo classificador a imagem da placa é recortada com bordas superiores à 40% da largura do objeto e inferiores à 40% da altura do objeto (para facilitar a leitura e evitar textos cortados). Após a placa ser recortada, então é passada para a função de leitura da biblioteca *Pytesseract*.

É importante destacar que se o tamanho do texto retornado for maior que 4 caracteres o narrador do sistema operacional irá ler o texto retornado, senão será apresentado um aviso que existe uma placa de perigo, mas o conteúdo não poderá ser lido. Para narrar o texto retornado foi utilizado um comando no terminal a partir do *script python* que usa o narrador do sistema operacional. Neste projeto foi usado o *speaker* para Ubuntu.

5 AVALIAÇÃO EXPERIMENTAL

Neste estudo foi conduzido um experimento a fim de avaliar a eficácia do sistema. Para a condução do experimento foram utilizados as diretrizes propostas por Wholin et al. [16].

5.1 Definição do Experimento

Para auxiliar a definição dos objetivos deste experimento foi utilizado o modelo *Goal-Question-Metric (GQM)* proposto por Basili e Weiss [17], conforme pode ser visto a seguir:

"Analisar o sistema com o propósito de avaliar com respeito à eficácia do ponto de vista de um pesquisador no contexto de 643 imagens diferentes."



Figure 3: Tipos de imagens. a) Superior - exemplo de imagem positiva. b) Inferior - exemplo de imagem negativa.

Com o objetivo definido, a seguinte Questão de Pesquisa (QP) foi investigada:

QP₁: O quão eficaz é o sistema proposto para reconhecer placas de perigo para deficientes visuais?

Para responder a QP₁, a eficácia foi avaliada utilizando a precisão, revocação e F1-Escore do sistema no reconhecimento de placas de perigo para deficientes visuais. Para essa questão de pesquisa as seguintes hipóteses foram definidas:

H₁₀: O sistema não é eficaz no reconhecimento de placas perigosas para pessoas com deficiência visual.

H₁₁: O sistema é eficaz no reconhecimento de placas perigosas para pessoas com deficiência visual.

5.2 Design do Experimento

Para este estudo foi conduzido um experimento (e_1) no qual foram analisadas o reconhecimento das placas perigosas como variável independente e duas variáveis dependentes:

- **minNeighbors (M)**: representa o número de vizinhos que cada retângulo candidato deve ter para retê-lo. Foram utilizados quatro parâmetros diferentes ($M = 10, 15, 20$ e 25) para cada grupo de imagens.
- **Regras de classificação (R)**: representa a forma como as placas perigosas foram identificadas nas imagens. Foram utilizados quatro parâmetros diferentes ($R = VP, FP, VN, FN$).
 - **Verdadeiro Positivo (VP)**: o sistema ao ser alimentado com o grupo de imagens positivas o mesmo devia salvá-las na pasta de positivos, sendo assim o número total de imagens salvas na pasta de positivas é caracterizado pelo número de verdadeiros positivos;
 - **Falso Positivo (FP)**: são as imagens que não foram salvas na pasta de positivos quando o sistema foi alimentado com o grupo de imagens positivas;
 - **Verdadeiro Negativo (VN)**: são as imagens que não foram salvas na pasta de positivas quando o sistema foi alimentado com o grupo de imagens negativas;
 - **Falso Negativo (FN)**: quando o sistema foi alimentado com imagens negativas o total de imagens salvas na pasta rotulada para imagens positivas caracteriza o número de imagens falsas positivas.

O experimento foi realizado com 643 imagens utilizando um notebook com processador Intel Core i3-2375M 1.5 GHz, memória 4GB e o sistema operacional Ubuntu 18.10.

A eficácia do sistema para o reconhecimento de placas perigosas foi avaliada utilizando as métricas de precisão, revocação e F1-Escore. Precisão é a razão entre as observações positivas previstas corretamente e o total de observações positivas previstas. Assim, Precisão (P) é definida como o número de verdadeiros positivos (VP) sobre o número de verdadeiros positivos mais o número de falsos positivos (FP), conforme pode ser visto na equação 1.

$$P = \frac{VP}{VP + FP} \quad (1)$$

Revocação é a razão entre observações positivas previstas corretamente e todas as observações na classe real. Assim, Revocação (R) é definida como o número de verdadeiros positivos (VP) sobre o número de verdadeiros positivos mais o número de falsos negativos (FN), conforme pode ser visto na equação 2.

$$R = \frac{VP}{VP + FN} \quad (2)$$

F1-Escore combina precisão e revocação em relação a uma classe positiva específica. F1-Escore (F1) pode ser interpretada como uma média ponderada da precisão (P) e revocação (R), onde o F1-Escore atinge seu melhor valor em 1 e pior em 0, conforme pode ser visto na equação 3.

5.3 Configuração do Experimento

O procedimento realizado durante a execução do experimento consistiu em seis etapas:

- (1) **criação da base de dados**: a base de dados criada para avaliação foi dividida entre 243 imagens positivas e 400 imagens negativas;
- (2) **configuração do sistema**: a saída do sistema foi configurada para salvar as imagens identificadas em uma pasta de

acordo com o grupo de entrada. Por exemplo, se a imagem continha a placa de perigo o sistema salvou a mesma em uma pasta rotulada como positiva;

- (3) **pré-processamento:** nesta etapa as imagens foram normalizadas de duas formas: (i) automática e (ii) manual. De forma automática foi utilizado um algoritmo em *Python* visando identificar imagens parecidas. De forma manual foi realizado uma normalização para remover imagens corrompidas. Além disso, foi aplicada a equalização de histograma da imagem para aumentar a nitidez e diminuir o contraste de áreas muito escuras ou muito claras;
- (4) **treinamento:** o classificador Haar Cascade foi treinado utilizando as imagens positivas e negativas;
- (5) **classificação das imagens:** as imagens foram classificadas utilizando uma função da biblioteca *Opencv*. Essa função é executada para cada *frame* do vídeo que terá como retorno as coordenadas caso a placa tenha sido identificada. No caso do experimento de avaliação, a função executava para cada imagem e retornava as coordenadas caso uma placa tenha sido identificada;
- (6) **validação:** foi computada a precisão, revocação e F1-Escore do reconhecimento das placas perigosas a partir do grupo de imagens positivas e negativas.

Para a construção da base de dados de avaliação as imagens positivas foram obtidas a partir de mecanismos de busca e bancos de imagens (diferentes dos sites visitados na criação da base de treinamento) usando as strings "Danger Sign", "segno di pericolo", "gevaar teken" e "Gefahrenzeichen" totalizando 313 imagens.

Esse conjunto de imagens foi normalizado com ajuda de um algoritmo em *Python* visando identificar imagens parecidas. Também foi realizado uma normalização manual para remover imagens corrompidas. Após a normalização a base de dados totalizou 243 imagens positivas.

As imagens negativas da base de avaliação foram obtidas por meio do mecanismo de busca de imagem utilizando a string "imagens de industrias" totalizando 400 imagens negativas.

6 RESULTADOS

Os resultados do experimento são apresentados e discutidos nesta seção.

$$F1 = 2 * \frac{(P * R)}{(P + R)} \quad (3)$$

Na Tabela 1 é apresentada a matriz de confusão dos resultados obtidos para cada regra de classificação. Na primeira coluna são apresentadas as regras de classificação. Para cada parâmetro tem-se o número de imagens e o percentual computado.

Avaliando o contexto de aplicação a caracterização de cada uma das regras de classificação se dá da seguinte forma:

- Verdadeiro positivo: caracterizado por uma imagem de entrada que possui o objeto de interesse para o contexto do sistema (placa de perigo) e durante a execução do experimento o sistema foi capaz de identificar o objeto de interesse contido na imagem, caracterizando com um asserto;
- Falso positivo: ocorre quando o sistema recebe como entrada um imagem que não contém uma placa de perigo, e a saída

do sistema foi que a imagem continha o objeto de interesse, caracterizando um erro;

- Verdadeiro negativo: ocorre quando o sistema é alimentado com uma imagem negativa e o mesmo não identifica nenhuma placa de perigo na imagem, caracterizando um asserto;
- Falso negativo: caracterizado pela entrada ser uma imagem que continha um objeto de interesse, porém o sistema não foi capaz de identificar a placa de perigo dessa imagem, caracterizando um erro.

Com a finalidade de responder a QP_1 , os resultados alcançados relacionados à precisão, revocação e F1-Escore são apresentados nas Figuras 4 e 5, respectivamente.

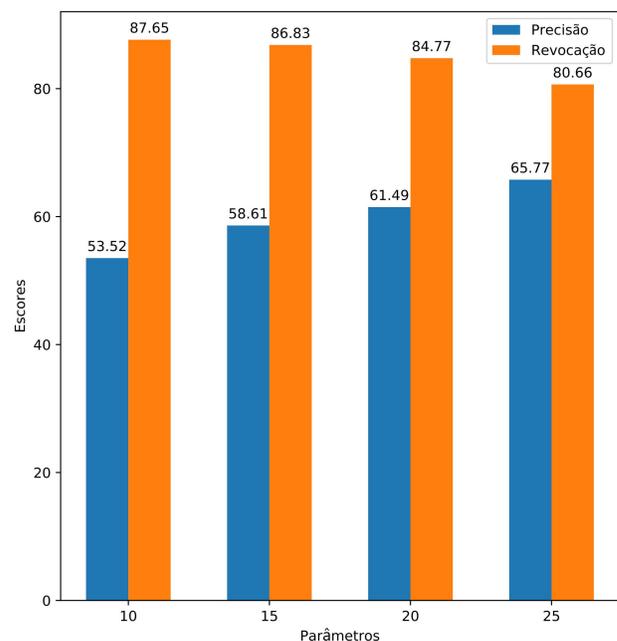


Figure 4: Precisão e revocação do reconhecimento de placas perigosas considerando os parâmetros minNeighbors

Conforme pode ser visto na Figura 4, dentre as métricas de avaliação a revocação se caracteriza por ser o melhor indicador. Nesta métrica os falsos negativos são considerados mais prejudiciais que os falsos positivos. Portanto, para o sistema, se uma placa de perigo não for identificada (falso positivo) a motivação principal do artigo está sendo ferida pois, o sistema pode estar deixando o usuário à mercê dos perigos sinalizados na placa. Sendo assim o resultado da métrica de revocação demonstra que o sistema tem um bom desempenho na resolução da problemática.

Ainda é possível constatar que a revocação apresentou resultados superiores a 80% para todos os parâmetros minNeighbors. O F1-Escore também alcançou bons resultados variando de 65% à 70%. Os resultados para a métrica precisão foram regulares variando de 53% à 65%, conforme pode ser visto na Figura 5.

Em comparações com trabalhos similares como por exemplo, o trabalho proposto por Souza e Marengoni [9], no qual o sistema proposto identificou 31 cenários de 41, ou seja, cerca de 75% de

Table 1: Matriz de confusão para cada regra de classificação considerando os parâmetros minNeighbors

	Parâmetros minNeighbors							
	10		15		20		25	
	n°.	%	n°.	%	n°.	%	n°.	%
Verdadeiro positivo	213	87,65%	211	86,83%	206	84,77%	196	80,66%
Falso positivo	185	46,25%	149	37,25%	129	15,23%	102	19,34%
Verdadeiro negativo	215	53,75%	251	62,75%	271	67,75%	298	74,50%
Falso negativo	30	12,35%	32	13,17%	37	15,23%	47	19,34%

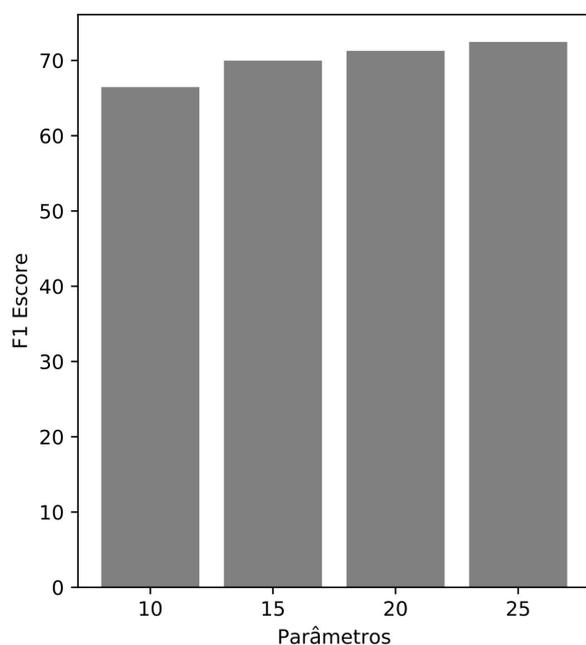


Figure 5: F1-Escore do reconhecimento de placas perigosas considerando os parâmetros minNeighbors

assertividade em regra de proporção simples, os resultados de verdadeiro positivo alcançados neste artigo são superiores, uma vez que foram acima de 80%.

7 CONSIDERAÇÕES FINAIS

O sistema proposto neste artigo visa auxiliar pessoas com algum tipo de deficiência visual na sua inserção no mercado de trabalho ou no cotidiano. Portanto a contribuição social foi um grande impulsionador para o seu desenvolvimento.

A principal métrica de avaliação para esse contexto se mostrou satisfatória alcançando em média 84,97% com a revocação. A preocupação em realizar avaliações de estresse teve o objetivo de revelar o máximo de problemas possíveis e mapeá-los de forma sistemática. A proposta tem caráter de sistema crítico, por ser responsável em manter a segurança de pessoas com deficiência visual que utilizarem o sistema.

Os problemas ocorridos na detecção e reconhecimento das placas têm como principais fatores: (i) iluminação do ambiente; (ii) posicionamento da placa; (iii) condições climáticas; e (iv) condições de conservação da placa. Estes fatores estão diretamente relacionados às dificuldades em criar uma base de imagens para treinamento do problema proposto. Portanto, outra importante contribuição deste trabalho trata-se também do banco de imagens normalizadas e mapeadas que poderá ser utilizado em outros treinamentos e avaliações.

Os trabalhos futuros são direcionados aos seguintes tópicos: (i) inserção de mais imagens na base de dados de treinamento; (ii) melhorias no âmbito de reconhecimento do texto das placas e no desempenho desse classificador; (iii) melhorar a voz do narrador tornando-a mais amigável; e (iv) realizar experimentos em dispositivos móveis com a finalidade de permitir a interação do usuário com o sistema.

REFERENCES

- [1] IBGE. Censo 2010: número de católicos cai e aumenta o de evangélicos, espíritas e sem religião, 2010. URL <https://censo2010.ibge.gov.br/noticias-censo?id=3&idnoticia=2170&view=noticia>.
- [2] BRASIL. Lei nº 8.213, de 24 de julho, jul 1991. URL <https://www.jusbrasil.com.br/busca?q=Art.+93+da+Lei+8213%2F91>.
- [3] Auzuir R. Alexandria. Sistema de reconhecimento óptico de algarismos para medidores convencionais de energia. 9, 2015.
- [4] Rafael C. Gonzales and Richard E. Wood. *Digital Image Processing*. Prentice-Hall, New Jersey, 2 edition, Nov. 2002.
- [5] Tom M. Mitchell. *Machine Learning*. McGraw-Hill Science/Engineering/Math, March 1997.
- [6] Mandy Sidana. Types of classification algorithms in machine learning, feb 2017. URL <https://medium.com/@Mandsidana/machine-learning-types-of-classification-9497bd4f2e14>.
- [7] OpenCV. Opencv 2.4.13.7 documentation, mar 2019. URL https://docs.opencv.org/2.4/modules/objdetect/doc/cascade_classification.html.
- [8] Francisco Carlos Monteiro Souza. *Técnicas de buscas heurísticas para otimização de parâmetros de máquinas de vetores suporte*. PhD thesis, Universidade Federal de Pernambuco, 2011.
- [9] Souza, K; Marengoni, M. Uso de Visão Computacional em Dispositivos Móveis para o Reconhecimento de Faixa de Pedestres. *Universidade Presbiteriana Mackenzie*, 2012.
- [10] Samuel Gomes, Elizângela Rebouças, and Pedro Pedrosa Filho. Reconhecimento optico de caracteres para reconhecimento das sinalizações verticais das vias de trânsito. *Soluções para Desenvolvimento do País*, 9:9, 05 2014.
- [11] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, volume 1, pages 1–I, 2001.
- [12] D. R. Bruno, D. O. Sales, J. Amaro, and F. S. Osório. Analysis and fusion of 2d and 3d images applied for detection and recognition of traffic signs using a new method of features extraction in conjunction with deep learning. In *2018 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8, 2018.
- [13] Dasar. Dasar haartrain, 2016. URL <http://www.mediafire.com/file/1aq02tpidk105fv/dasahaartrain.rar/file>.

- [14] Python-Software-Fundation. Pytesseract - pypl, jun 2019. URL <https://pypi.org/project/pytesseract/>.
- [15] Google-Open-source. Tesseract ocr, jun 2019. URL <https://opensource.google.com/projects/tesseract>.
- [16] C. Wohlin, P. Runeson, M. Höst, M. C. Ohlsson, B. Regnell, and A. Wesslén. *Experimentation in Software Engineering: An Introduction*. Springer-Verlag Berlin Heidelberg, 1st edition, 2012.
- [17] Victor R. Basili and David M. Weiss. A methodology for collecting valid software engineering data. *IEEE Transactions on Software Engineering*, 10(6):728–738, 1986.