

ADVIR: Ferramenta para Auxílio de Deficientes Visuais na Identificação de Remédios

Daniel Ferreira Costa
Universidade do Estado do Rio Grande do Norte
Departamento de Computação
Natal/RN, Brasil
Email: daniel.ferreira.costa@hotmail.com

Anderson Souza
Universidade do Estado do Rio Grande do Norte
Departamento de Computação
Natal/RN, Brasil
Email: anderson@uern.br

Resumo—Hoje em dia, a autonomia pessoal e social das pessoas tornou-se um problema da sociedade. É importante fornecer ferramentas para ajudar as pessoas a fim de permitir que elas sejam independentes, especialmente para as atividades diárias. O ADVIR é uma ferramenta desenvolvida para ajudar deficientes visuais na identificação de medicamentos. A idéia principal é apontar um dispositivo de câmera (como uma câmera de smartphone) para a uma caixa de medicamento, a fim de identificá-lo e saber, por sinais sonoros, informações sobre sua composição e seu uso. Para isso, o algoritmo de visão computacional ORB foi utilizado para identificar recursos na caixa de medicamentos e para realizar a correspondência com padrões previamente armazenados. Experiências preliminares foram realizadas para comprovar a viabilidade do sistema.

Abstract—Nowadays, the personal and social autonomy of people have become a problem of society. It is important to provide tools for assisting people in order to enable them to be independent, especially for daily activities. ADVIR is a tool developed for helping visually impaired people in drug identification. The main idea is to point a camera device (such as a smartphone camera) to the box of a drug in order to identify it and know, by sound signals, information about its composition and its use. For this, computational vision algorithms were used to identify features ORB in the drug box and to perform matching with previously stored patterns. Preliminary experiments were performed to prove the feasibility of the system.

I. INTRODUÇÃO

Do total da população brasileira, 3,5% declaram ter algum tipo de deficiência visual. É a deficiência mais comum do país [1], em seguida, ficaram problemas motores (2,3%), intelectuais (1,4%) e auditivos (1,1%). Segundo a Organização Mundial da Saúde [2], as principais causas de cegueira no Brasil são: catarata, glaucoma, retinopatia diabética, cegueira infantil e degeneração macular. Algumas das doenças que podem levar a baixa visão, estão relacionadas a velhice, associando-se à perda da autonomia e independência. Dessa forma, outra informação importante, considerando os objetivos deste trabalho, é salientar o número igualmente expressivo de pessoas idosas no Brasil, que segundo o IBGE [3] em 2017 chega a 30,2 milhões. A baixa visão é relatada como o terceiro maior problema crônico do idoso, após as artrites e as cardiopatias [4].

Neste grupo de pessoas existem usuários de medicamentos e, assim como todos, precisam acessar as informações contidas em suas embalagens, as quais, muitas vezes são ilegíveis

para quem possui restrições visuais mais severas. Apesar de a Associação Nacional de Vigilância Sanitária (ANVISA) ter publicado, no dia 23 de dezembro de 2009, em seu Diário Oficial [5], uma norma na qual diz que os fabricantes de medicamentos são obrigados a colocar nas caixas o nome do medicamento em braile, grande parcela da população não é familiarizada com o método braile.

A visão computacional é o conjunto de métodos e técnicas, através dos quais, sistemas computacionais são capazes de interpretar imagens, transformando-as em dados digitais [6]. Diante disso, neste trabalho é proposto um sistema que utiliza técnicas de visão computacional para identificar caixas de medicamentos e informar, através de sinais sonoros, a composição e o uso desse medicamento. O algoritmo de detecção e descrição de características (*features*) ORB (*Oriented FAST and Rorared BRIEF*) [7] é utilizado para localizar os pontos mais notáveis na embalagem dos medicamentos, enquanto que, em uma etapa posterior, a biblioteca FLANN (*Fast library for Approximate Nearest Neighbors*) [8] é utilizada para realização do matching entre os padrões identificados e os previamente armazenados.

Dessa forma, uma pessoa com deficiência visual poderá usar este sistema para obter essas informações sobre o medicamento, bem como identificar o nome da medicação e suas indicações, ajudando-a a tomar seus medicamentos de forma correta.

Este trabalho segue com a Seção II, que traz trabalhos que se relacionam com este. Seção III apresenta a fundamentação teórica para se entender os algoritmos estudados; a Seção IV mostra o sistema implementado; a Seção V mostra os experimentos e as análises feitas com os dados obtidos; e a Seção VI traz os aspectos conclusivos deste estudo realizado e trabalhos futuros.

II. TRABALHOS RELACIONADOS

O uso de algoritmo de detecção e descrição de características (*features*) para identificação de objetos é bem consolidado na comunidade de visão computacional. Amato, Galchi e Bolettieri [9], por exemplo, relatam a eficiência de se utilizar o algoritmo SURF (*Speed-Up Robust Features*) para identificar monumentos, estatuas, edifícios, entre outros. Dentre suas conclusões, é o bom desempenho que essas técnicas obtiveram

no reconhecimento dos marcos selecionados para o teste que as fazem ser bem utilizadas. Apesar do contexto diferente (não se tratar de medicamentos), a ideia tem como base a mesma, utilizar uma câmera para identificar algo.

Quando se trata de sistema de identificação visual de caixas de medicamentos, Xiankleber e colegas [10] desenvolveram sua ferramenta baseada no algoritmo SURF. O trabalho consistiu em utilizar uma câmera (de dispositivos populares) para identificar caixas de medicamentos através de imagens e áudios, mostrando as deficientes informações sobre anamnése, tais como: posologia, indicação e contra-indicações de medicação. De acordo com a análise feita pelos autores, o sistema foi útil ou muito útil a 93% dos usuários para identificar os medicamentos pela caixa. O trabalho é semelhante ao proposto neste artigo, porém, a nossa proposta se baseia no uso do ORB para detecção e descrição de características. O SURF é um algoritmo mais custoso computacionalmente e, além disso, não é open source, sendo necessária uma licença para poder ser utilizado.

Além disso, no trabalho citado uma etapa de redução de resolução é aplicada nas imagens analisadas, ficando o foco principal em apenas uma parte da imagem, com uma técnica de foveamento [10]. Isso foi feito para reduzir o custo computacional de execução do sistema. No nosso caso, por utilizarmos o ORB [7], um algoritmo mais rápido que o SURF, não alteramos a resolução das imagens, entendendo que, com isso, não haverá perda de informações por conta de baixas resoluções.

III. FUNDAMENTAÇÃO TEÓRICA

A. Detecção de Características

As features são características que se destacam na imagem, por possuírem variações constantes em uma determinada região. Se referem a uma parte de uma imagem com alguma propriedade especial, por exemplo, um círculo, uma linha, ou uma região textura, pontos, curvas e superfícies [6]. Existem vários métodos para detecção de características, tais como: o filtro de Roberts e o filtro de Sobel, que são usados para detectar bordas e cantos [11]. Essas características podem ser utilizadas para diferentes propósitos: reconhecimento de objetos, odometria visual, mapeamento, detecção de movimentos, entre outros.

Porém, são necessários algoritmos de detecção para que essas características possam ser extraídas e utilizadas. Esses algoritmos produzem como saídas um conjunto de descritores (*features descriptors*), os quais especificam a posição e outras propriedades essenciais das características encontradas na imagem. Esses algoritmos suportam variações tanto de escala quanto de rotação, ou seja, mesmo que uma imagem seja redimensionada, ou mesmo que seja rotacionada, esses algoritmos são capazes de detectar os pontos notáveis. Essa é uma das principais vantagens de se utilizá-los nas aplicações relacionadas à identificação de objetos.

Para se obter as características da imagem, existem algumas etapas importantes, são elas: a aquisição da imagem, que consiste em obter a imagem em si; o pré-processamento, que

são pequenos ajustes como correções de ruídos e contraste; a extração de características, ou seja, extrair algum ponto representativo da imagem (*feature*) com seu respectivo elemento descritor. Após isso, a característica pode ser utilizada para o processamento de alto nível, por exemplo, detectar um objeto utilizando as *features*.

Neste trabalho, foi utilizado o algoritmo ORB [7] por ter sua eficácia comprovada na identificação de objetos [12].

B. ORB - Oriented FAST and Rotated BRIEF

O ORB é um algoritmo que surgiu como uma opção alternativa ao SIFT e ao SURF, propondo melhor desempenho computacional [7]. Basicamente, o ORB é uma fusão do detector de pontos (*keypoints*) FAST (*Features from Accelerated Segment Test*) com o descritor BRIEF (*Binary Robust Independent Elementary Features*).

Para que haja a detecção de características em uma imagem, o ORB, detecta a posição do pixel, que será definido como uma característica, e cria um descritor que funciona como o identificador do pixel.

C. FLANN - Fast Library for Approximate Nearest Neighbors

Contém uma coleção de algoritmos otimizados para pesquisa rápida de vizinhos mais próximos em grandes conjuntos de dados e para recursos dimensionais elevados [13].

A biblioteca FLANN analisa as correspondências, os algoritmos de busca se baseiam em busca em árvores k-d, que são estruturas de dados úteis em aplicações que trabalham com busca de dados multidimensionais [8]. Ele compara todos os descritores (atribuídos às *features*) de um primeiro conjunto, com todos os descritores de um segundo conjunto, usando algum cálculo de distância, e o mais próximo é retornado, isto é, o ponto correspondente [8], conforme pode ser visto na Fig. 1.



Figura 1. Reconhecimento entre duas imagens [13].

IV. SISTEMA IMPLEMENTADO

Este trabalho tem como motivação principal ajudar as pessoas com deficiência visual a identificar informações de medicamentos, que muitas vezes são ilegíveis para quem possui restrições visuais, por meio da análise de imagens e descrição sonora. Para realizar este trabalho, foram utilizados algoritmos para a identificação das caixas de medicação, através das *features*, usando a câmera instalada em um dispositivo (como computador ou smartphone, por exemplo), e apontando para a caixa de remédio. Após a identificação, é possível escutar, por meio de áudio, o nome e a indicação do medicamento.

Para o desenvolvimento das etapas do sistema foi utilizada a biblioteca OpenCV (*Open Source Computer Vision Library*), originalmente desenvolvida pela Intel em 2000. Essa biblioteca é multiplataforma e totalmente livre para uso acadêmico e comercial. A OpenCV possui módulos de Processamento de Imagens e Video I/O, além de mais de 350 algoritmos de Visão Computacional como: filtros de imagem, calibração de câmera, reconhecimento de objetos, análise estrutural e outros [14].

A. Funcionamento do sistema

É importante mencionar que imagens de medicamentos precisam ser previamente cadastradas e suas informações descritas em um tipo de banco de dados.

Para identificar as características das caixas dos remédios, devemos apontar a câmera para a caixa, conforme a Fig. 2. Com a captura da imagem, segue a etapa de identificação os pontos (*features*) na imagem pelo algoritmo ORB.



Figura 2. Detecção da caixa de medicamento usando uma Câmera.

As características detectadas nas imagens capturadas pela câmera, no momento do uso do sistema pelo usuário, são, portanto, comparadas com as características detectadas nas imagens padrão armazenadas no banco de dados. Essa comparação (*matching*) é feita por meio da biblioteca FLANN. Caso um determinado número de características com suas respectivas correspondentes seja alcançado, logo estabelece-se que o medicamento foi reconhecido. Caso contrário, o medicamento não é reconhecido por falha ou por não ter sido cadastrado no banco de dados. A Fig. 4 mostra um exemplo de detecção de características e *matching*.

Acontecendo o reconhecimento do medicamento, o sistema emite um áudio com as informações referentes ao remédio. O diagrama da Fig. 3 mostra o funcionamento geral do sistema. O primeiro passo é executar o programa deixando a câmera pronta; em seguida o usuário deve apontar o remédio para a câmera, até ser identificado. Em seguida, o áudio com a identificação do remédio é iniciado, mencionando o nome do medicamento e suas indicações. Por fim, caso a caixa que esteja na mão do usuário não seja a que ele busca, pode-se apontar uma outra caixa para a câmera até se encontrar o medicamento correto cadastrado.

B. Módulo de Controle de Qualidade

A caixa do medicamento é melhor identificada estando numa posição plana, porém podem ocorrer identificações falsas. Esses falsos positivos podem levar a erros indesejáveis,

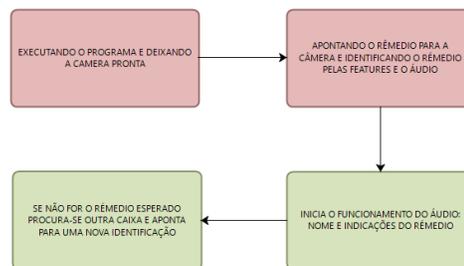


Figura 3. Diagrama do Sistema.

uma vez que o sistema não deve, sob nenhuma circunstância, indicar o remédio errado. Desta forma, o áudio não é reproduzido na primeira vez que o remédio é detectado. É feita uma contagem de quantas vezes o medicamento foi corretamente identificado. Se atingir um número satisfatório, então o áudio é reproduzido com as informações do medicamento. Dessa forma, pode ser necessário um tempo maior de posicionamento da caixa do medicamento em frente a câmera.



Figura 4. Detecção do medicamento.

Normalmente, cada caixa de medicamento é reconhecida pelo seu lado frontal, que é plano. A parte traseira pode também ser usada, mas pode não ser tão eficaz quanto ao número de acertos como o lado da frente da caixa. No entanto, pode-se optar por registrar ambos os lados, a fim de que o utilizador possa apresentar a caixa à câmera, usando os dois lados.

O processamento da detecção é uma questão importante, porque o sistema deve fornecer um feedback ao usuário sobre a detecção do objeto em tempo real (exemplo Fig. 4. Dentre tantos algoritmos de extração de características sugeridas pela literatura, optamos por utilizar as do ORB [7] por ser, atualmente, um dos mais rápidos, em termos de processamento, por sua eficiência em extrair boas características e por ser open source.

V. EXPERIMENTOS

Como parte final do processo de desenvolvimento, foram realizadas experiências com oito caixas de medicamentos, com o intuito de avaliar a eficácia no reconhecimento dos remédios.

Foram realizadas 20 tentativas de reconhecimento para cada caixa, a taxa de sucesso de detecção de cada medicamento pode ser vista na Tabela I. Note que o número 6 da tabela tem taxa de sucesso de 90%, e o número 7 de 95%, enquanto as demais tiveram 100% de taxa de sucesso. Isto significa que as caixas que não obtiveram 100% de sucesso na detecção, tem menos recursos (menos características detectadas) do

Tabela I
TAXA DE SUCESSO NA DETECÇÃO DAS CAIXAS

Caixa de Medicamentos	Taxa de Sucesso
#1	100%
#2	100%
#3	100%
#4	100%
#5	100%
#6	90%
#7	95%
#8	100%

que as outras. Na verdade, o rótulo pequeno e o logotipo podem ser confundidos com outra caixa de medicamento. Foi possível notar que a taxa de sucesso é proporcional ao número de características. Na Fig. 5 é possível observar as caixas utilizadas no experimento.



Figura 5. Caixas utilizadas no experimento.

Outro aspecto avaliado foi o tempo de processamento. O hardware utilizado foi um laptop Core i7 2.00GHz com 16GBRAM e uma webcam com resolução de imagem de 640 x 480. O tempo de processamento e detecção do medicamento pode ser visto na Tabela II. Note que a medicação que mais demorou para ser reconhecida foi a caixa de número 3, ainda assim, demorou menos de 50ms, que pode ser considerado um resultado satisfatório em relação ao tempo de processamento. Todavia, não foram feitos testes de comparação de tempo com o algoritmo SURF, por não ser código aberto.

Tabela II
TEMPO DE PROCESSAMENTO NA DETECÇÃO DAS CAIXAS

Caixa de Medicamentos	Processamento de tempo(ms)
#1	39.43ms
#2	21.61ms
#3	48.02ms
#4	28.50ms
#5	23.74ms
#6	19.06ms
#7	14.15ms
#8	41.03ms

VI. CONCLUSÃO E TRABALHOS FUTUROS

Foi proposto um sistema de visão computacional que objetiva ajudar pessoas com deficiência visual, detectando e fornecendo informações sobre medicamentos. Foi observado

que, apesar da maioria das caixas conterem o nome em Braille, não são todas as pessoas que sabem essa linguagem, principalmente aqueles que se tornaram deficientes visuais durante a vida. Além disso, as indicações do medicamento não estão disponíveis em Braille.

O reconhecimento dos medicamentos foi satisfatório, todavia, rótulos comuns podem dificultar a detecção do objeto. Utilizar um banco maior de imagens pré-processadas pode ajudar a melhorar a detecção. A utilização do algoritmo ORB e da biblioteca FLANN faz com que o tempo de processamento seja bem curto, fazendo com que todo o processamento seja realizado em tempo real.

Para trabalhos futuros, pretende-se testar o sistema com um grupo de pessoas com deficiência visual e obter um retroativo dos mesmos, e posteriormente, aplicar um questionário com objetivo de saber as dificuldades encontradas pelo usuário. Além disso, o sistema será implementado em uma versão para dispositivos móveis com uma interface amigável e adequada aos deficientes visuais.

AGRADECIMENTO

A elaboração deste trabalho não teria sido possível sem a colaboração, estímulo e empenho de diversas pessoas. Muitos contribuíram, de diferentes formas, para que este trabalho se realizasse. Por este facto, de expressar toda a minha gratidão e apreço a todos aqueles que contribuíram para que esta tarefa se tornasse uma realidade: ao meu professor orientador desse trabalho e a SIBGRAPI 2018, pela oportunidade de submeter este trabalho a um evento deste porte.

REFERÊNCIAS

- [1] Fundação Dorina Nowill para Cegos. fundacaodorina.org.br.
- [2] Organização Mundial de Saúde (OMS). who.int.
- [3] Instituto Brasileiro de Geografia e Estatística (IBGE). ibge.gov.br.
- [4] Gasparetto, M.E.R.F; Nobre, M.I.R.S. Avaliação do funcionamento da visão residual: educação e reabilitação. São Paulo: fcm.unicamp.br, 2007.
- [5] Ascom (Acessoria de Imprensa da Anvisa) Rótulos de medicamentos vão mudar. portal.anvisa.gov.br, 23 de dezembro de 2009.
- [6] Trucco, E.; Verri, A. *Introductory Techniques for 3-D Computer Vision*. Upper Saddle River, NJ, USA: Prentice Hall PTR, 1998.
- [7] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "Orb: an efficient alternative to sift or surf," in *In Proc. of IEEE International Conference on Computer Vision (ICCV)*, 2011, p. 2564–2571.
- [8] M. Muja and D. Lowe, "Fast approximate nearest neighbors with automatic algorithm configuration," in *International Conference on Computer Vision Theory and Application*, 2009, pp. 331–340.
- [9] Amato, G.; Galchi, F; Bolettieri, P. Recognizing landmarks using automated classification techniques: Evaluation of various visual features. Atenas: In: *Advances in Multimedia (MMEDIA) Second International Conferences*, 2010.
- [10] X. C. Benjamim, R. B. Gomes, A. F. Burlamaqui, and L. M. G. Gonçalves, "Visual identification of medicine boxes using features matching," in *2012 IEEE International Conference on Virtual Environments Human-Computer Interfaces and Measurement Systems (VECIMS) Proceedings, July 2012*, pp. 43–47.
- [11] Gonzalez, R. C.; Woods, R. E. *Digital Image Processing Third Edition*. Pearson Prentice Hall, 1992.
- [12] B. A. S. Santana, R. S. Maia, W. Blanco, and A. Souza, "Desempenho de algoritmos detectores de keypoints para um sistema de navegação visual de robôs baseados em smartphones," in *XII Simpósio Brasileiro de Automação Inteligente (SBAI 2015)*, Natal, Brasil, 2015.
- [13] OpenCV – Documentation FLANN. docs.opencv.org.
- [14] Bradsky, G. R. et al. *Learning opencv, Computer Vision with the OpenCV Library*. Springer, 2006.