

Reconhecimento de Ambientes em Aplicativos Móveis

Guilherme Eduardo Rodrigues, Felipe Augusto Przysiada, Verônica Isabela Quandt, Caroline Mazetto Mendes
Grupo de Visão Computacional Aplicada, Universidade Positivo
Curitiba, Paraná, Brasil
E-mail: gvca@up.edu.br

Resumo—Um dos desafios enfrentados por indivíduos com deficiência visual é a localização em ambientes desconhecidos. Nesse contexto, tecnologias para reconhecimento de ambientes podem prover maior autonomia a esses indivíduos. Este trabalho apresenta o desenvolvimento de um sistema baseado em processamento de imagens e computação cognitiva para reconhecimento de ambientes em uma instituição de ensino. Por meio de um aplicativo móvel, usuários são informados sobre o ambiente ou local em que se encontram. A solução proposta realiza o reconhecimento de ambientes como salas de aula, sala de estudos, laboratórios, corredores, sanitários, entre outros. Como resultado, o sistema provê taxa de reconhecimento de 88,33% e 9,322 segundos de tempo de resposta médio.

Abstract—One of the challenges faced by individuals with visual impairment is the location in unfamiliar environments. In this context, technologies for recognizing environments can provide greater autonomy to these individuals. This work presents the development of a system based on image processing and cognitive computation for the recognition of environments in an educational institution. Through a mobile application, users are informed about the environment or location they are. The proposed solution carries the recognition of environment such as classrooms, meeting rooms, laboratories, corridors, toilets, among others. As a result, the system provides recognition rate of 88.33% and 9.347 seconds of average response time.

I. INTRODUÇÃO

Indivíduos com deficiência visual, como cegos ou pessoas com baixa visão, convivem diariamente com o desafio da localização e reconhecimento de ambientes. Além das dificuldades relacionadas à locomoção em ambientes desconhecidos, esses indivíduos comumente necessitam da ajuda de outras pessoas para conhecer as características do ambiente em que se encontram. Tal dificuldade está presente também no ambiente escolar ou nas instituições de ensino.

Tecnologias assistivas consistem em dispositivos, estratégias e práticas aplicadas para ampliar as habilidades dos indivíduos com deficiência [1]. Nesse contexto, aplicativos para dispositivos móveis podem ser utilizados para auxiliar deficientes visuais na tarefa de localização e reconhecimento de ambientes. Soluções baseadas em visão computacional fornecem informações sobre o ambiente, ampliando assim a capacidade de visão desses indivíduos.

Atualmente, existem variados aplicativos para dispositivos móveis que auxiliam deficientes visuais no reconhecimento de objetos contidos em um ambiente. BlindTool [2], TapTap See [3] e Aipoly Vision [4] são alguns exemplos. Em geral,

esses aplicativos fazem o reconhecimento de objetos a partir de uma imagem capturada pela câmera do dispositivo.

Aplicações baseadas em visão computacional têm sido propostas para auxiliar os deficientes visuais nas tarefas de reconhecimento [5]. Contudo, a maioria das soluções existentes para dispositivos móveis focam no reconhecimento de objetos, porém não realizam o reconhecimento de locais ou ambientes. Um indivíduo com deficiência visual poderia utilizar a câmera do celular também para descobrir se está dentro de uma sala, um corredor, um laboratório, entre outros locais que compõem um ambiente, sendo ele aberto ou fechado.

Este trabalho apresenta o desenvolvimento de um sistema que realiza o reconhecimento de ambientes comumente presentes em uma universidade. A solução utiliza recursos de processamento de imagens e computação cognitiva para realizar a tarefa de reconhecimento. Por meio de um aplicativo móvel são captadas imagens do ambiente. A fim de evitar falhas no reconhecimento, a qualidade da imagem é avaliada via aplicativo móvel. Após a identificação do ambiente, o aplicativo informa ao usuário o local utilizando mensagem de texto e voz.

II. TRABALHOS RELACIONADOS

Matusiak et al. [6] propuseram um aplicativo móvel de reconhecimento de objetos para pessoas com deficiência visual. Desenvolvido para a plataforma Android, o aplicativo reconhece objetos a partir da câmera do smartphone. O nome do objeto reconhecido é comunicado ao usuário por meio de mensagens verbais. Os autores destacam que o desempenho do sistema é dependente da qualidade da imagem capturada e condições de iluminação.

Juang et al. [7] realizaram o reconhecimento de objetos a partir de um dispositivo móvel. A solução proposta utiliza técnicas de processamento de imagens, considerando características de cor e textura dos objetos. Os resultados dos testes mostram que o aplicativo é capaz de reconhecer objetos em diferentes ângulos, tamanhos e direções.

Jadhav et al. [8] desenvolveram um aplicativo Android de reconhecimento de objetos para pessoas com deficiência visual. O usuário utiliza a câmera do celular para gravar um vídeo que contém o objeto de interesse. O objeto é então detectado na imagem e comparado a outros objetos armazenados no banco de dados da aplicação. Após o reconhecimento, o

aplicativo informa o nome do objeto utilizando um sintetizador de fala. Os autores concluem que o reconhecimento de objetos é útil para o dia a dia de usuários com deficiência visual.

Um aplicativo Android destinado a pessoas com deficiência visual foi desenvolvido por Kadam et al. [9]. O aplicativo é capaz de reconhecer objetos capturados com a câmera de um dispositivo móvel, como por exemplo um smartphone. A detecção dos objetos é realizada por meio de técnicas de processamento de imagens e um algoritmo de rede neural artificial. O nome do objeto é informado ao usuário por meio de uma mensagem de voz.

III. METODOLOGIA

O sistema de reconhecimento de ambientes desenvolvido é composto por um aplicativo Android responsável pela captura, pré-processamento das imagens e interação com o usuário.

O reconhecimento de ambientes foi implementado com a tecnologia de reconhecimento visual do IBM Watson, que utiliza algoritmos de Inteligência Artificial baseados em Computação Cognitiva para analisar imagens de cenários e objetos [10]. A Computação Cognitiva traz uma tecnologia capaz de processar informações e de aprender com elas de forma muito semelhante ao cérebro humano [11].

A Figura 1 mostra o fluxograma do sistema. Para realizar o reconhecimento de um ambiente, o usuário deve posicionar a câmera do celular no ponto de vista desejado e tocar em qualquer local da tela. A imagem capturada é avaliada quanto a luminosidade e depois enviada ao servidor para análise e reconhecimento do ambiente. O servidor, por sua vez, retorna ao aplicativo os resultados do reconhecimento. Por fim, o aplicativo realiza a conversão do texto do resultado em mensagem de voz, informando ao usuário em qual ambiente ele se encontra dentro da universidade.

A. Treinamento e Classificação

Para acesso à ferramenta de reconhecimento visual do IBM Watson foi criada uma conta na plataforma Bluemix¹. Nessa plataforma foram criadas 14 classes de treinamento com quantidade variável de imagens em cada uma, como mostra a Tabela I. Ao todo, foram utilizadas 463 imagens de diferentes ambientes, cada uma representando um local diferente dentro das instalações da universidade. As imagens foram capturadas utilizando as seguintes câmeras: câmera do smartphone Moto G(5S) e Canon EOS REBEL T5i. Por restrição da capacidade de armazenamento da conta Watson utilizada, que é de 250 MB, as imagens foram reduzidas para 96 dpi (horizontal e vertical).

Para realizar a classificação, utilizou-se a ferramenta de reconhecimento visual do IBM Watson. O método de classificação utilizado produz uma nota entre 0,0 e 1,0 para cada imagem de cada classe. Ou seja, o sistema retorna ao usuário o nome de todas as classes pertencentes ao projeto e a porcentagem da imagem pertencer a cada uma das classes.

¹<https://www.ibm.com/cloud-computing/bluemix/pt>

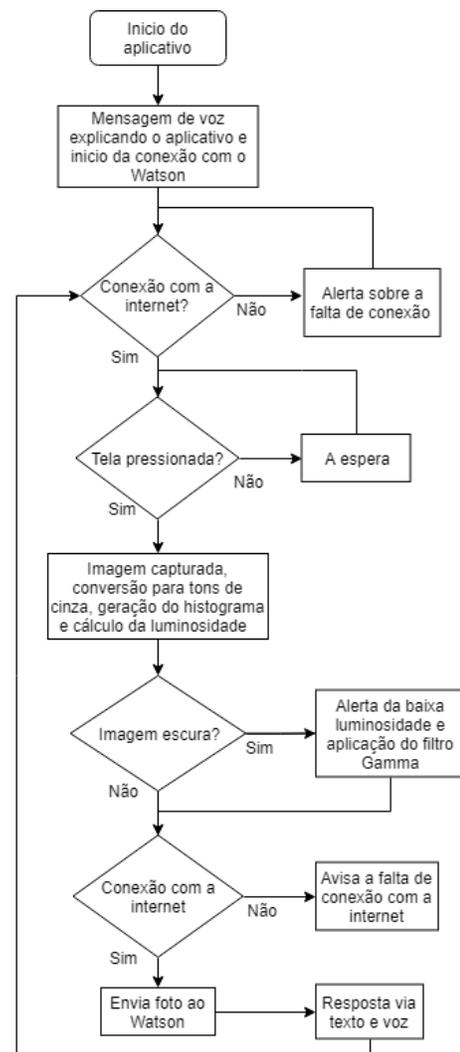


Figura 1. Fluxograma que exemplifica o funcionamento do aplicativo de reconhecimento de ambientes.

Tabela I
NOME DAS CLASSES E NÚMERO DE IMAGENS UTILIZADAS NO TREINAMENTO DE CADA CLASSE.

| Classe | Número de Imagens |
|------------------------------|-------------------|
| Hall | 39 |
| Laboratório | 35 |
| Sala de Aula | 65 |
| Banheiro | 21 |
| Corredores | 21 |
| Entrada principal | 16 |
| Entrada Auditório | 11 |
| Escada Externa | 44 |
| Escada Interna | 18 |
| Porta Entrada | 29 |
| Sala de estudos Biblioteca | 11 |
| Biblioteca | 43 |
| Bloco da Pós-Graduação | 22 |
| Área externa da Universidade | 50 |

B. Aplicativo Android

O aplicativo móvel foi desenvolvido para a plataforma Android, sendo utilizada a biblioteca OpenCV para realizar o pré-processamento das imagens. Para conversão do texto do resultado em mensagem de voz foi utilizada uma função do próprio Android Studio chamada "TextToSpeech".

A aquisição da imagem é feita no momento em que o usuário pressiona a tela do dispositivo. Para que a imagem não fique tremida ou borrada, foi habilitada a função de foco automático da câmera do smartphone. A orientação da tela foi definida como vertical. Assim, há uma configuração que rotaciona a foto para a orientação correta. Os parâmetros de configuração foram adicionados via código utilizando a Api Camera2 do Android, os quais geram ajustes na câmera dos dispositivos.

A imagem capturada é avaliada quanto a luminosidade para assegurar o reconhecimento do ambiente. Primeiro, a imagem é convertida para tons de cinza. Depois, calcula-se o histograma para verificar se a imagem é escura. Nesse caso, um filtro Gamma de valor 0,5 é aplicado, gerando uma cópia da imagem com a luminosidade ajustada.

O reconhecimento é realizado a partir da conexão com o servidor da IBM, para onde a imagem é enviada. O servidor então retorna o resultado do reconhecimento. Um ambiente é determinado pela classe com a maior taxa, desde que acima de 70%. Caso nenhuma classe atinja este valor, o aplicativo informa que o ambiente não foi reconhecido. Para garantir o correto funcionamento do aplicativo, é realizado o tratamento de exceções, por exemplo ocasionada por falha de comunicação com o servidor da IBM.

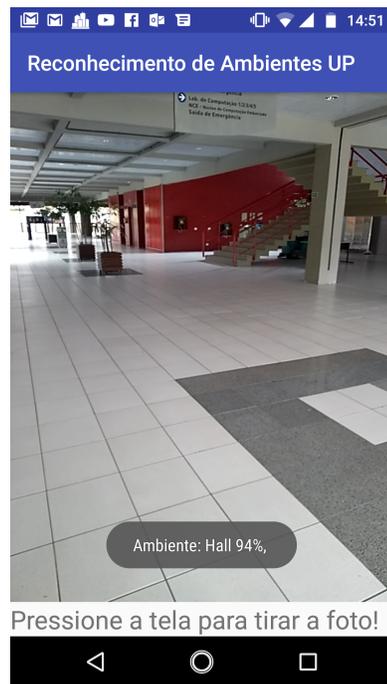
IV. RESULTADOS E DISCUSSÃO

O testes foram realizados nas instalações da Universidade Positivo, localizada em Curitiba, Paraná. Foi utilizado o smartphone Motorola Moto G5s com acesso à Internet wi-fi. Foram capturadas ao todo 439 imagens de diferentes pontos de vista dos ambientes treinados, possuindo 72 dpi cada.

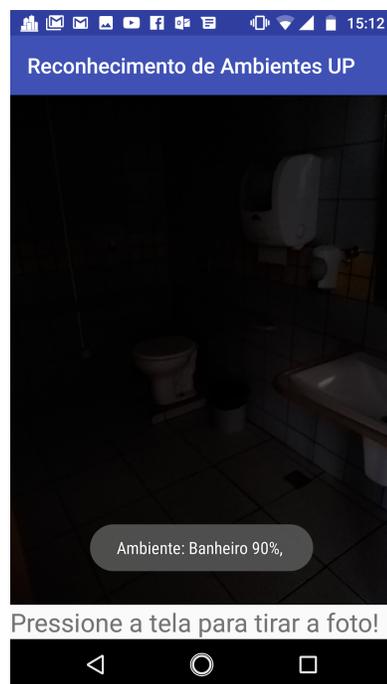
Como resultados, obteve-se uma taxa de reconhecimento média de 88,33% e tempo de resposta médio de 9,322 segundos. O tempo de resposta sofre influência da velocidade e intensidade de sinal da rede de Internet. Com uma rede wi-fi com grandes oscilações, houve retornos em pouco mais 5 segundos e em até 21 segundos.

A Figura 2a mostra o resultado para o ambiente Hall com taxa de reconhecimento de 94%. A Figura 2b mostra o ambiente Banheiro com taxa de reconhecimento de 90%. Observa-se que o sistema reconheceu o ambiente, apesar da baixa luminosidade.

Foram registrados 57 casos de falso positivo e 14 casos de falso negativo. O falso positivo consiste no reconhecimento incorreto de um ambiente. Já o falso negativo ocorre quando a classe com a maior taxa de reconhecimento corresponde ao ambiente correto porém essa taxa o não atingiu o limiar de 70%. Nesse caso, o aplicativo informa ao usuário que o ambiente não foi reconhecido.



(a)

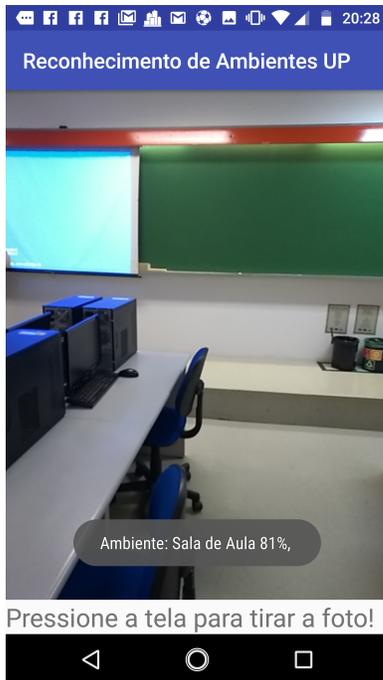


(b)

Figura 2. Reconhecimento de ambientes da universidade: (a) Hall e (b) Banheiro com baixa luminosidade.

A Figura 3 mostra um exemplo de falso positivo e falso negativo. O ambiente Laboratório foi reconhecido como sendo Sala de Aula (Figura 3a). Esses ambientes possuem em comum a presença de quadro-negro e tablado. Já o ambiente Entrada Principal (Figura 3b) não foi reconhecido devido a taxa de reconhecimento ter sido de 45%, menor que o limiar definido.

Na realização dos testes, observou-se que os erros de reconhecimento ocorrem devido a semelhança que algumas classes possuem entre si. Para minimizar tais casos, é necessário realizar novos treinamentos adicionando mais fotos para cobrir diferentes pontos de vista dos ambientes.



(a)



(b)

Figura 3. Exemplos de falso positivo e falso negativo: (a) Laboratório reconhecido como Sala de Aula e (b) Entrada Principal não reconhecida.

V. CONCLUSÃO

Este artigo apresentou o desenvolvimento de um sistema de reconhecimento de ambientes em aplicativos móveis. Por meio de uma aplicação Android, as imagens capturadas são avaliadas quanto a luminosidade e depois são enviadas ao servidor, que realiza o reconhecimento do ambiente. O resultado do reconhecimento é informado ao usuário por mensagem de texto e mensagem de voz.

Os resultados mostram que o sistema é capaz de reconhecer variados ambientes da universidade na qual os testes foram realizados, incluindo ambientes com baixa iluminação. Devido à limitação do espaço de armazenamento da conta IBM Watson, utilizou-se uma menor quantidade de fotos no treinamento das classes. Contudo, pode-se concluir que o IBM Watson contém recursos que podem ser utilizados na criação de aplicativos baseados em computação cognitiva e visão computacional.

Como trabalhos futuros, novos treinamentos serão realizados para minimizar os erros de reconhecimento. Uma nova versão do sistema será desenvolvida utilizando o TensorFlow com o objetivo de realizar uma comparação entre as abordagens. Por fim, serão realizados testes com alunos e colaboradores da universidade, incluindo pessoas com deficiência visual.

REFERÊNCIAS

- [1] K. Cook and L. Hussey, "Assistive technologies: Principles and practice," in *Assistive Technologies: Principles and Practice*, 1995.
- [2] "Blindtool," <https://play.google.com/store/apps/details?id=the.blindtool&hl=en>, acesso em: 27-08-2018.
- [3] "Taptapsee," <https://play.google.com/store/apps/details?id=com.msearcher.taptapsee.android&hl=en>, acesso em: 27-08-2018.
- [4] "Aipoly vision," https://play.google.com/store/apps/details?id=com.aipoly.vision&hl=pt_BR, acesso em: 27-08-2018.
- [5] M. Leo, G. Medioni, M. Trivedi, T. Kanade, and G. Farinella, "Computer vision for assistive technologies," *Computer Vision and Image Understanding*, vol. 154, no. C, pp. 1–15, Jan. 2017. [Online]. Available: <https://doi.org/10.1016/j.cviu.2016.09.001>
- [6] K. Matusiak, P. Skulimowski, and P. Strumiłło, "Object recognition in a mobile phone application for visually impaired users," in *2013 6th International Conference on Human System Interactions (HSI)*, June 2013, pp. 479–484.
- [7] L.-H. Juang, M.-N. Wu, and Z.-Z. Weng, "Object identification using mobile devices," *Measurement*, vol. 51, pp. 100–111, 2014.
- [8] S. Jadhav, S. Tomy, S. S. Jayswal, H. D. Dhaware, and A. R. Vijapure, "Object detection in android smartphone for visually impaired users," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 5, no. 11, pp. 332–334, 2016.
- [9] A. J. Kadam, S. Awate, S. Desai, R. Khese, and G. Patange, "Android application for visually impaired users," *Imperial Journal of Interdisciplinary Research (IJIR)*, vol. 3, no. 2, pp. 1674–1677, 2017.
- [10] "IBM Watson - getting started tutorial," <https://console.bluemix.net/docs/services/visual-recognition/getting-started.html#getting-started-tutorial>, acesso em: 15-08-2018.
- [11] Y. Chen, E. Argentinis, and G. Weber, "Ibm watson: How cognitive computing can be applied to big data challenges in life sciences research," *Clinical Therapeutics*, vol. 38, no. 4, pp. 690–701, 2016.