

Auxílio a Deficientes Visuais utilizando Redes Neurais Convolucionais e Competição e Cooperação de Partículas

Jefferson Antonio Ribeiro Passerini,
Inst.Geociências e Ciências Exatas-IGCE
Universidade Estadual Paulista - UNESP
Rio Claro - SP
Jefferson.passerini@unesp.br

Fabricio Aparecido Breve,
Inst.Geociências e Ciências Exatas-IGCE
Universidade Estadual Paulista - UNESP
Rio Claro - SP
fabricio@rc.unesp.br

Resumo—Segundo a ONU existem 2,2 bilhões de pessoas com alguma deficiência visual. A visão computacional evoluiu na última década a partir da utilização de CNNs e constitui um recurso para auxílio a essas pessoas. Este trabalho tem por objetivo estabelecer um modelo baseado na extração de características através de CNNs e como classificador utilizar o modelo de Competição e Cooperação de Partículas (PCC) melhorando a acurácia apresentada em trabalhos anteriores. Atualmente o estudo está em fase de levantamento bibliográfico e determinou dificuldades como conjunto de imagens para testes.

Área: *Inteligência Computacional.*

I. INTRODUÇÃO

Pessoas com deficiências visuais ou totalmente cegas são um aspecto importante de nossa sociedade, o aumento de doenças relacionadas aos olhos e a redução da visão estão representando cada vez mais um desafio a instituições e governos.

Segundo a ONU [1] no mundo existem pelo menos 2,2 bilhões de pessoas com algum tipo de deficiência visual, no Brasil como levantado durante o censo demográfico em 2010 pelo IBGE [2] existiam 35 milhões de pessoas com algum acometimento visual, das quais 506.377 pessoas eram totalmente cegas e mais de 6 milhões possuíam grandes dificuldades para enxergar.

As tecnologias assistivas provêm a segurança, mobilidade e acesso às pessoas com deficiências [3], assim as pesquisas na área de reabilitação estão avançando e os dispositivos móveis e os aplicativos representam um dos vetores de melhoria da qualidade de vida dessas pessoas [4].

A área de visão computacional com foco no reconhecimento de objetos e detecção de obstáculos obtiveram grandes avanços na última década [5][6], muitos desses sistemas são caros, pesados ou ainda não estão amplamente disponíveis à população [7]. Existem ainda sistemas de acessibilidade que necessitam da utilização de redes de dados para acesso a servidores remotos [8], deste modo o dispositivo mais utilizando atualmente na detecção de obstáculos ainda é a bengala.

Esta pesquisa tem por objetivo desenvolver um modelo que permita realizar a detecção de obstáculos de forma autônoma, sem a necessidade de acesso a servidores remotos, através de imagens capturadas por um dispositivo móvel classificando-as em duas classes: caminho livre, ou com obstáculos.

II. CONCEITOS E TÉCNICAS

O advento das redes neurais convolucionais CNNs determinou um grande avanço na detecção de objetos e classificação de imagens [9][10]. Como qualquer modelo de aprendizado supervisionado, as CNNs necessitam de um grande volume de amostras rotuladas em sua fase de treinamento e o processo de preparar amostras rotuladas é muito custoso e demorado para ser realizado [11].

O treinamento é a etapa que exige o maior custo computacional na utilização de CNNs, uma vez treinadas é possível a utilização de modelos como VGG19 [12] ou o InceptionV3 [13] para determinar a existência de obstáculos a partir da captura de imagens em um *smartphone*.

A utilização de transferência de aprendizado (*transfer learning*) permite aplicar camadas que foram treinadas em grandes conjuntos de imagens otimizando uma CNN para aplicações onde os conjuntos de treinamento são escassos [11]. Essa técnica foi utilizada com sucesso em diversas situações [12] [13][14].

No cenário proposto, com limitações de recursos se houver a necessidade de um retorno do usuário ao sistema seria necessário realizar um novo treinamento da rede inviabilizando a sua utilização.

Um modelo de aprendizado semi-supervisionado como o modelo de Competição e Cooperação de Partículas - PCC [15] não necessita de treinamento como ocorre nos modelos supervisionados. O PCC [15] é um método de aprendizado baseado em grafos e inspirado na natureza, onde times de partículas caminham pela rede cooperando entre eles e competindo com outros times, tentando possuir o maior número de vértices possíveis.

Durante a preparação da rede complexa para o modelo de aprendizado é necessário a aplicação de métodos de redução de complexidade como PCA [16] ou UMAP [17].

III. METODOLOGIA DE DESENVOLVIMENTO

O modelo de classificação para determinar se o caminho a ser percorrido pelo deficiente visual está livre ou se possui obstáculos, será implementado na linguagem Python.

O modelo proposto terá algumas etapas como a extração de características, a redução de dimensionalidade dos dados extraídos, a geração de uma rede complexa (grafo) e a etapa de classificação.

Em uma imagem coletada por um aparelho celular será aplicado o processo de extração de características através de uma CNN (sem o seu bloco de classificação). Obtido os vetores de características será utilizado um método de redução de dimensionalidade antes da construção da rede complexa (Fig.1).

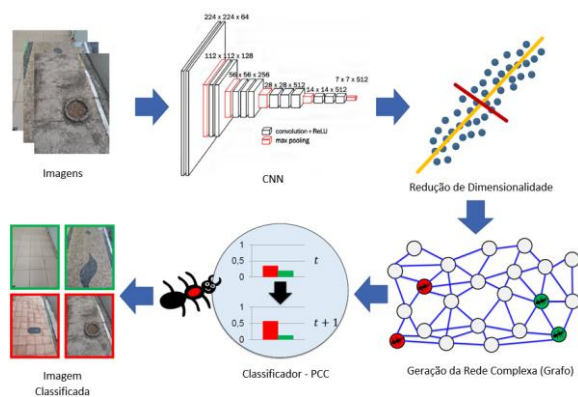


Fig. 1. Metodologia Proposta (Adaptado de Breve e Fischer [18]).

Com a rede complexa construída é aplicado o PCC para classificar o caminho como limpo ou com obstáculos.

Em trabalho anterior aplicando esta metodologia Breve e Fischer [18] obtiveram a acurácia média de 73,43% utilizando para extração de características as redes CNNs VGG16+VGG19, PCA como redutor de dimensionalidade e o algoritmo PCC como classificador.

O conjunto de imagens para teste foi proposto por eles sendo composto por 342 imagens, sendo 175 demonstrando caminho livre e 167 com obstáculos [18].

IV. RESULTADOS PRELIMINARES

O trabalho encontra-se em fase de levantamento de bibliografia e trabalhos correlatos. Observou-se que na abordagem de diversos trabalhos com detecção de obstáculos não existe um conjunto de imagens adequado que seja utilizado por diversos autores, mas sim verificou-se a adoção de conjuntos próprios nos diversos trabalhos.

Foi possível também conhecer modelos de CNNs e de redução de dimensionalidade que serão aplicados à metodologia proposta durante os testes.

V. CONSIDERAÇÕES FINAIS

Os conhecimentos adquiridos durante essa fase do estudo são importantes para estabelecer uma base de conceitos que será utilizada para o refinamento da metodologia e a aplicação de testes ao modelo proposto.

REFERÊNCIAS

- [1] ONU. World Report on vision. Disponível em: <https://www.who.int/publications/i/item/world-report-on-vision>. ONU, Genebra, Suíça, 2019.
- [2] IBGE, BRASIL. Censo Demográfico 2010. Instituto Brasileiro de Geografia e Estatística. Brasília-DF, 2010.
- [3] Manduchi, R; Coughlan, J. (Computer) vision without sight. *Comm. ACM*, vol. 55, nº 1, pags 96–104, 2012.
- [4] Terven, J.R; Salas, J; Raducanu, B. New Opportunities for Computer Vision-Based Assistive Technology Systems for the Visually Impaired. *Computer*, vol. 47, nº 4, pp. 52-58, 2014.
- [5] Rizzo, J.R.; Pan, Y.; Hudson, T.; Wong, E.K.; Fang, Y. Sensor fusion ecologically valid obstacle identification: Building a comprehensive assistive technology platform for the visually impaired. In 2017 7th ICMSAO. Pag. 1-5. IEEE, 2017.
- [6] Lakde, C.K.; Prasad, P.S. Review paper on navigation system for visually impaired people. *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 4, nº 1, 2015.
- [7] Hoang, V.N.; Nguyen, T.H.; Le, T.L.; Tran, T.H.; Vuong, T.P.; Vuillerme, N. Obstacle detection and warning system for visually impaired people on electrode matrix and mobile Kinect. *Vietnam Journal of Computer Science*, vol. 4, nº 2, pag. 71-83, 2017.
- [8] Jiang, B.; Yang, J.; Lv, Z.; Song, H. Wearable vision assistance system based on binocular sensors for visually impaired users. *IEEE Internet of Things Journal*, vol. 6, nº 2, pag. 1375-1383, 2019.
- [9] Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. *Mobilenets: Efficient convolutional neural networks for mobile vision applications*, 2017.
- [10] LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature*, vol. 521, nº 7553, pag. 436, 2015.
- [11] Oquab, M.; Bottou, L.; Laptev, I.; Sivic, J. Learning and transferring mid-level image representative using convolutional neural networks. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June, 2014.
- [12] Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *Computational and Biological Learning Society*, 2015, pag. 1-14.
- [13] Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the inception architecture for computer vision. In *The IEEE Conference on Computer Vision and Pattern Recognition (CPVR)*, Jun. 2016, pag.2818-2826, 2016.
- [14] Gopalakrishnam, K.; Khaitan, S.K.; Choudhary, A.; Agrawal, A. Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection. *Construction and Building Materials*, vol. 157, pag. 322-330, 2017.
- [15] Breve, F.; Zhao, L.; Quiles, M.; Pedrycz, W.; Liu, J. Particle competition and cooperation in networks for semi-supervised learning. *IEEE Transactions on Knowledge and Data Engineering*, vol. 24, nº 9, pag. 1686-1698, set. 2012.
- [16] Jolliffe, I. Springer-Verlag. *Principal Components Analysis*. Springer Series in Statistics. Springer, 2002.
- [17] McInnes, L.; Healy, J.; Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. Disponível em: arxiv.org/abs/1802.03426v3, 2020.
- [18] Breve, F.; Fischer, C.N. Visually Impaired Aid using Convolutional Neural Networks, Transfer Learning, and Particle Competition and Cooperation. In *International Joint Conference on Neural Networks (IJCNN)*, mar. 2020.