

ISSN: 2319-0124

REDES NEURAI CONVOLUCIONAIS NA IDENTIFICAÇÃO DE PRODUTOS ALIMENTÍCIOS COMO FERRAMENTA AUXILIAR PARA DEFICIENTES VISUAIS

Matheus A. SIMÕES¹; Hugo RESENDE²

RESUMO

Diversas limitações observadas em humanos podem ser contornadas por meio do uso de técnicas computacionais. Nesse sentido, esse trabalho tem como objetivo, apresentar uma proposta de utilização de Redes Neurais Artificiais a fim de auxiliar deficientes visuais na manipulação de produtos alimentícios. Tais indivíduos, por possuírem uma limitação sensorial, recorrem a utilização de Tecnologias Assistivas para contornar este problema da falta de informações sensoriais. As tecnologias das RNCs possuem um grande potencial quando utilizadas em cenários de computação visual, ao analisar e classificar imagens e reconhecer padrões. Os resultados oriundos das RNCs serão reportados mediante uma leitura por voz sintetizada.

Palavras-chave:

Acessibilidade; Inteligência Artificial; Tecnologia Assistiva; Autonomia.

1. INTRODUÇÃO

A existência de indivíduos com algum tipo de deficiência visual é algo que vem sendo observado ao longo da história mundial, nas últimas décadas, fez-se ainda mais notório o número de pessoas diagnosticadas com tais deficiências (DA SILVA, 1987). Tais indivíduos que apresentam algum tipo de deficiência, necessitam de uma maior atenção médica em muitas das vezes. As limitações enfrentadas por este grupo podem ser contornadas com uso de ferramentas auxiliares, como os óculos, as lupas e/ou as bengalas.

Com o grande avanço da tecnologia, abordagens mais atuais podem ser adotadas, como Tecnologias Assistivas (TAs) que podem ser consideradas como um conjunto de serviços e ferramentas que auxiliam e proporcionam a ampliação das habilidades de pessoas com deficiência, de modo que as tais possam ser independentes em suas tarefas (BERSCH; TONOLLI, 2006). Um uso recorrente para essas TAs tem sido a tecnologia de Inteligência Artificial, de forma que se realize uma classificação/automação de processos previamente treinados nesta tecnologia. Uma das melhores abordagens neste assunto tem sido a de Redes Neurais Convolucionais (RNCs) as quais têm a capacidade de extrair as principais informações e características de uma imagem.

Desse modo, este documento demonstra uma proposta de utilização de uma RNC a fim de se obter uma TA construída em linguagem Python que se utiliza dos frameworks Tensorflow e Keras, de modo que possa classificar e identificar produtos alimentícios, a fim de auxiliar no dia a dia de deficientes visuais.

2. FUNDAMENTAÇÃO TEÓRICA

A inclusão, além de ser um processo que tem se tornado cada vez mais evidente, é também uma necessidade para todos os indivíduos independentemente da sua limitação. Embora seja

¹Orientando, IFSULDEMINAS – Campus Passos. E-mail: matheus.simoes@alunos.ifsuldeminas.edu.br.

²Orientador, IFSULDEMINAS – Campus Passos. E-mail: hugo.resende@ifsuldeminas.edu.br.

possível observar a existência dessas novas limitações aos deficientes visuais, a academia e a indústria têm buscado meios de propiciar, por meio da proposição de recursos auxiliares, a inclusão dessas pessoas na sociedade. Tais recursos incluem ferramentas/mecanismos de acessibilidade, tais como, programas sociais e aplicações de softwares, de modo que essas pessoas possuidoras de cegueira ou baixa visão consigam se restabelecer e retornar ao exercício das suas atividades habituais (ORRICO; CANEJO; FOGLI, 2007).

Sendo assim, a utilização de recursos tecnológicos que visem auxiliar esse grupo de indivíduos, tendo em vista as suas necessidades, se torna muito pertinente. Particularmente, a utilização de TAs auxiliam e incentivam o indivíduo na realização de suas tarefas, pois estas lhe passam o sentimento de autonomia e independência (FREITAS, 2015).

A fim de compreender o funcionamento de uma Inteligência Artificial, faz-se necessária a compreensão de um neurônio artificial, o qual é um dos principais componentes de qualquer Rede Neural Artificial (RNA). Tanto as RNAs quanto as RNCs são compostas por vários neurônios artificiais agrupados em camadas e interconectados entre si por meio dessas camadas. Em uma RNA, os neurônios ficam, via de regra, agrupados em camadas dispostas mediante alguma arquitetura. A principal arquitetura de RNAs encontrada na literatura computacional é a arquitetura feedforward de múltiplas camadas, também conhecidas como Multi-Layer Perceptron (MLP). Nessa arquitetura, o sinal de entrada é propagado da esquerda para a direita. Tal arquitetura também permite a contenção de uma camada intermediária, ou também várias outras camadas escondidas entre a camada de entrada e a camada de saída (DA SILVA, SPATTI, FLAUZINO, 2010).

De modo a classificar dados eficientemente, as redes Perceptrons de Múltiplas Camadas (MLP) se utilizam de algoritmos de treinamento, dentre os quais se destaca o algoritmo de backpropagation. Nesse algoritmo, o objetivo central é ajustar os pesos sinápticos da rede (gerados, inicialmente, de forma aleatória), de modo que a RNA possa realizar a classificação de dados similares, com a máxima acurácia possível. De forma resumida, os ajustes de pesos são baseados nos erros observados da rede a partir de comparações entre os resultados obtidos e os resultados esperados, ou seja, o aprendizado é construído por meio de uma supervisão algorítmica.

As RNCs são uma variação das redes de MLP, inspiradas nos processamentos de dados visuais. Sendo assim, uma RNC é capaz de realizar a aplicação de filtros nas imagens, mantendo uma relação de vizinhança entre os pixels durante o processamento da rede. Um filtro pode ser entendido como uma submatriz de pixels que possui como objetivo fornecer a capacidade de encontrar padrões similares a seu valor onde o filtro for utilizado. Em muitos trabalhos encontra-se a utilização das RNCs na tarefa de se realizar o processo de classificação, detecção e reconhecimento em imagens e vídeos (VARGAS; PAES; VASCONCELOS, 2016). Um exemplo de RNC generalista pode ser observado por meio da Figura 1.

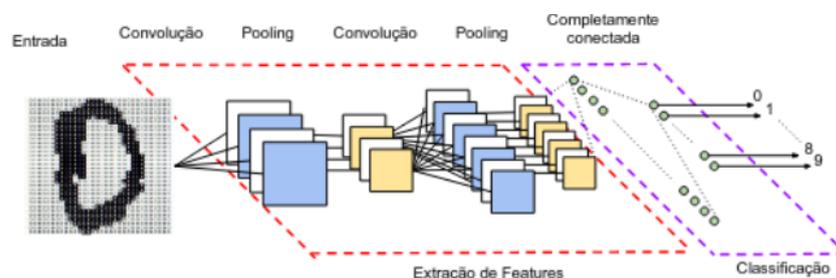


Figura 1. Modelo uma RNC. Extraído e adaptado de Vargas, Paes e Vasconcelos (2016).

Conforme se observa na Figura 1, há, constitui o processamento da RNC as camadas de convolução, pooling e completamente conectada. Tal modelo serve como ponto de partida para os demais modelos de RNCs utilizados para os diversos fins.

3. MATERIAL E MÉTODOS

O presente projeto foi desenvolvido na plataforma Google Colab, uma ferramenta gratuita que possibilita a criação de projetos em linguagem Python, de forma que é possível realizar operações num servidor remoto. Para a implementação do projeto, será utilizada a linguagem Python, a qual possui uma vasta gama de bibliotecas à disposição, tais como a TensorFlow e a Keras. Essas bibliotecas fornecem recursos especializados que permitem trabalhar de forma eficiente com projetos que envolvem redes convolucionais, técnicas objeto de implementação neste trabalho.

No que tange aos dados que serão classificados pelas RNCs propostas no trabalho, far-se-á o uso do dataset Freiburg Groceries Dataset (JUND et al., 2016). Inicialmente, o objetivo era trabalhar com alguma espécie de dataset que contivesse dados de produtos regionais e que fosse originalmente brasileiro, no entanto, não foi encontrado um banco de imagens que possuía tais requisitos. Tal motivação está relacionada ao interesse na propositura de uma solução que eventualmente ajudaria deficientes visuais brasileiros em suas tarefas cotidianas. O dataset freiburg groceries dataset possui imagens com resolução de 256x256 pixels, as quais são categorizadas em 25 classes de alimentos, tais como café, milho, farinhas, açúcar, arroz, leite, macarrão, entre outros. Na Figura 2 observa-se algumas imagens contidas no supramencionado dataset.



Figura 2. Exemplo de conteúdo do dataset Freiburg Groceries Jund *et al.* (2016).

Para a elaboração deste trabalho, serão realizadas as averiguações de modo que se encontre o melhor modelo que consiga obter uma acurácia aceitável a dado cenário. O dataset utilizado possui, nativamente, um conjunto para treinamento e um outro conjunto para testes. Neste trabalho, serão utilizados tais conjuntos sem modificações manuais, juntamente com as métricas de validação holdout, k-fold e leave-one-out. Pretende-se, dessa forma, excluir os possíveis vícios do modelo devido a uma má distribuição manual para o treinamento e validação.

4. RESULTADOS ESPERADOS

Uma vez que o trabalho estiver concluído, pretende-se obter como resultados, primeiramente, um modelo de RNC capaz de classificar a imagem capturada pelo usuário, e retornar a classe encontrada ao usuário por meio da leitura por voz sintetizada. Além disso, é esperado que esta aplicação possa gerar um impacto no meio acadêmico e tecnológico, a fim de promover mais interesse no desenvolvimento de novas TAs. Por fim, pretende-se desenvolver uma TA que possa assistir de forma eficaz aos deficientes visuais durante as suas atividades de compras em ambiente comerciais, para que estes a possam realizar de forma mais prática e segura, além de promover o sentimento individual de independência neste usuário.

5. CONCLUSÕES

Nos últimos anos tem-se observado que o avanço tecnológico propiciou uma melhora significativa na qualidade de vida de diversas pessoas que possuem algum tipo de deficiência ou limitação física. Nesse sentido, no presente projeto foi apresentada uma proposta de desenvolvimento de uma TA dotada de uma técnica da IA conhecida como Redes Neurais Convolucionais. Uma vez concluída a pesquisa, espera-se gerar uma gama considerável de contribuições. Dentre elas, pode-se apresentar o próprio modelo construído neste trabalho, pois não foi encontrada uma variedade de trabalhos correlatos no mesmo âmbito. Devido a isso, a elaboração deste modelo convolucional se mostra ser muito importante, podendo ser utilizado em futuros trabalhos, visto que existe uma carência de preocupação nesta área de aplicação. Por fim, destaca-se como a mais importante das contribuições a utilização da aplicação pela comunidade de deficientes visuais, de modo que a tecnologia os agregue tanto quanto possível. Dessa forma, a aplicação cumprirá o seu objetivo inicial, que consiste essencialmente em promover a independência do indivíduo visualmente debilitado.

REFERÊNCIAS

BERSCH, Rita; TONOLLI, José C. Introdução ao conceito de Tecnologia Assistiva e modelos de abordagem da deficiência. Bengala Legal, p. 25, 2006.

DA SILVA, Ivan Nunes; SPATTI, Danilo Hernane; FLAUZINO, Rogério Andrade. Redes neurais artificiais para engenharia e ciências aplicadas-curso prático. São Paulo: Artliber, 2010.

DA SILVA, Otto Marques. A epopéia ignorada: a pessoa deficiente na história do mundo de ontem e de hoje. Centro São Camilo de Desenvolvimento em Administração da Saúde, 1987.

FREITAS, Flaviane Pelloso Molina et al. A inclusão na percepção dos alunos deficientes visuais: um desafio a toda equipe escolar. Revista ESPACIOS| Vol. 36 (Nº 15) Ano 2015, 2015.

JUND, Philipp et al. The freiburg groceries dataset. arXiv preprint arXiv:1611.05799, 2016.

ORRICO, Helio; CANEJO, Elizabeth; FOGLI, Bianca. Uma reflexão sobre o cotidiano escolar de alunos com deficiência visual em classes regulares. 2007.

VARGAS, Ana Caroline Gomes; PAES, Aline; VASCONCELOS, Cristina Nader. Um estudo sobre redes neurais convolucionais e sua aplicação em detecção de pedestres. In: Proceedings of the xxix conference on graphics, patterns and images. sn, 2016.