

IMPLEMENTAÇÃO E APLICAÇÃO DE MODELO PARA REDES NEURAIAS CONVOLUCIONAIS DE MÁSCARA

RESUMO: A área de processamento de imagens apresenta grandes avanços tecnológicos nas últimas décadas. Um exemplo disso é a crescente qualidade de produtos sendo comercializados que utilizam essa tecnologia em seus processos e a popularização dessas ferramentas na vida cotidiana, em computadores pessoais e outros aparelhos. Hoje já é comum um sistema reconhecer partes específicas do rosto de uma pessoa, definir a emoção que a face expressa, entre outros. Com isso, diversas arquiteturas de redes neurais são desenvolvidas e aprimoradas, visando implementações mais performáticas e eficazes. Recentemente, com o lançamento de novas arquiteturas de rede e tecnologias de desenvolvimento dessas ferramentas, novos modelos de redes foram implementados e disponíveis para o público utilizando de tecnologias disponíveis e diversos guias de passo a passo para sua implementação. O artigo visa estudar e aplicar uma implementação desta tecnologia com foco em acessibilidade, facilidade de desenvolvimento e rápidos resultados.

PALAVRAS-CHAVE: reconhecimento; detecção; redes; neurais; processamento; imagens.

IMPLEMENTATION AND APPLICATION OF MASK CONVOLUTIONAL NEURAL NETWORKS

ABSTRACT: Huge advances have been made in image processing in the last decade, with such examples as the crescent quality in the detection technology which products are being commercialized and the popularity these solutions have in day to day life, in personal computers and devices. Nowadays is common for a system to be able to recognize specific body parts in a face, detect which emotion the person is feeling, among other features. The field of artificial neural network architecture are being developed and upgraded, aiming for improvements in performance and efficiency. Recently, with new architectures for these networks being launched, some network models have been implemented and available for the public with documents and guides on how to implement them. This report aims to study and develop an implementation of this technology, focusing on ease, accessibility and speed to get results.

KEYWORDS: recognition; detection; neural; networks; processing; images.

INTRODUÇÃO

Com o rápido avanço da tecnologia de processamento de imagens e fácil acessibilidade à pesquisas neste campo surgiram inúmeras implementações em um esforço de tornar a tecnologia mais acessível e atrair maior público ao auxílio do desenvolvimento e evolução da visão computacional. Com isso, algumas ferramentas se destacam pela sua extensa documentação, fácil aprendizado e poderosas abstrações que reduzem o conhecimento necessário para se iniciar no reconhecimento de imagens, com guias passo a passo de implementações de sistemas para este fim até códigos públicos já implementados.

O artigo "Mask R-CNN" pelo grupo FAIR (*Facebook Artificial Intelligence Research*) introduzia uma estrutura flexível, compacta e simples para segmentação de instâncias de objeto. As Mask-RCNNs eram extensões do artigo escrito pelos mesmos autores, "Faster R-CNN", que introduziam essa capacidade de segmentação antes indisponível.

Nesse estudo escolhemos uma implementação das Mask R-CNNs para estudarmos sua aplicação e extração de resultados, avaliando a simplicidade e agilidade no desenvolvimento da solução. Procuramos no final do estudo ter uma implementação de uma rede neural convolucional que seja capaz de realizar a segmentação de instância conforme descrito.

MATERIAL E MÉTODOS

O conjunto de metodologias de investigação utilizadas neste projeto se fundamentarão em duas práticas de pesquisa: a descritiva e a experimental.

A descritiva será utilizada no levantamento do referencial teórico das técnicas de processamento digital de imagens e a implementação de redes neurais convolucionais destacadas. Serão buscadas técnicas e

ferramentas voltadas para o desenvolvimento de sistemas de visão computacional de uso livre, gratuito e de conteúdo educativo sobre sua implementação facilmente disponível.

A metodologia experimental será utilizada pela própria natureza do projeto, cujo objetivo geral propõe modelar e criar uma ferramenta de processamento computacional de imagens. Além disso, será desenvolvido um repositório de imagens para o treinamento da ferramenta com diversos exemplos positivos e negativos de modelos e suas medições reais com o intuito de formar uma melhor plataforma para o desenvolvimento da solução de aprendizado de máquina.

Após isso, treinar a ferramenta utilizando o repositório de imagens previamente desenvolvido e fazer a avaliação e testes da aplicação com o objetivo de avaliar o desempenho da ferramenta, o tempo investido em sua implementação e realizar um guia para futuras implementações.

RESULTADOS E DISCUSSÃO

1

Para o desenvolvimento da solução, modelos prontos e configurações padrões serão adotadas para agilizar o processo de implementação da rede. Com isso, escolhemos a implementação da rede MS-COCO, construída com estrutura de ResNet101 e treinada no *dataset* COCO [9, 10]. Utilizando de modelos pré-treinados, os *weights* são especialmente úteis pela sua capacidade de aproveitar todo treinamento já realizado e abre possibilidade de interrupção para serem continuados em um momento posterior [8], visto que esse processo é demorado.

Com a rede treinada e configurada, a imagem que se deseja alimentar no sistema é passada como uma matriz de pontos RGB [8] e processada com as classes importadas do modelo treinado, presente no arquivo *weight* importado. Aproveitando-se das classes já presentes no *dataset* COCO, utilizamos imagens que contenham objetos reconhecidos pela rede.

A rede retorna as instâncias encontradas, os pontos pertencentes à essa mesma instância e quatro pontos para a caixa de delimitadora [8]. Uma exibição padrão desse resultado é a aplicação de uma “máscara” de uma cor para cada instância na imagem original, sobrepondo-a com o resultado da rede neural.



FIGURA 1. Foto de alunos disponível no site <conict.ifsp.edu.br> (esquerda) e seu resultado depois de passar pela rede neural (direita).



FIGURA 2. Foto do câmpus disponível no site <conict.ifsp.edu.br> (esquerda) e seu resultado depois de passar pela rede neural (direita).

CONCLUSÕES

Seguindo as ferramentas disponíveis no mercado para sua implementação, é possível verificar que a tecnologia de classificação, detecção de objetos, localização e segmentação de instância são hoje em dia facilmente implementadas. O pesquisador consegue, com o auxílio de documentos e *frameworks*, desenvolver sua rede neural convolucional em pouco tempo, facilitando que novas pesquisas e aplicações sobre o conteúdo sejam desenvolvidas, impulsionando o desenvolvimento da tecnologia com maior acessibilidade ao público.

REFERÊNCIAS

- ABDULLA, W. *Mask R-CNN for Object Detection and Instance Segmentation on Keras and TensorFlow*. 2017.
- BENGIO, Y. LECUN, Y. *Convolutional Networks for Images, Speech and Time-Series*. 1997.
- FLORINDO, João B. *Redes Neurais Convolucionais – Deep Learning*. Instituto de Matemática, Estatística e Computação Científica, Universidade Estadual de Campinas (UNICAMP), Brasil
- HE, Kaiming. GKIOXARI, Georgia. DOLLÁR, Piotr. GIRSHICK, Ross. *Facebook AI Research (FAIR)*. 2017
- KARPATHY, A. CS231n *Convolutional Neural Networks for Visual Recognition*, 2016.
- Keras. Disponível em: <<https://www.keras.io>>. Acesso em: 20 ago 2019.
- LIN, Tsung-Yi. MAIRE, Michael. BELONGIE, Serge. BOURDEV, Lubomir. GIRSHICK, Ross. HAYS, James. PERONA, Pietro. RAMANAN, C. Lawrence Zitnick. DOLLÁR, Piotr. Microsoft COCO: *Common Objects in Context*. 2014.
- RIPLEY, B. D. *Pattern recognition and Neural Networks*. Cambridge University Press. ISBN 0 521 46086.7 1996.
- Tensorflow. Disponível em: <<https://www.tensorflow.org/>>. Acesso em: 20 ago 2019.