



Processamento de imagens faciais usando aprendizado profundo de máquina com a biblioteca dlib

Natan R. de Oliveira*, Roberto A. Lotufo

Resumo

Esta pesquisa tem como objetivo reconhecer imagens faciais em fotos e vídeos utilizando aprendizado profundo de máquina com o auxílio da biblioteca de código aberto dlib¹. Devido à complexidade do tratamento de fluxo de dados com algoritmos de clusterização hierárquica, esta ferramenta tornou-se inviável para o escopo deste projeto. O desenvolvimento de redes neurais profundas convolucionais para reconhecimento de faces é uma alternativa para atingir o objetivo deste projeto.

Palavras-chave:

Aprendizado de Máquina, Deep Learning, Reconhecimento Facial.

Introdução

O reconhecimento facial enquanto técnica tem muitas aplicações nos dias atuais, principalmente no que tange segurança de estabelecimentos e usuários. Com a popularização de bibliotecas de código aberto que implementam diversas aplicações em aprendizado de máquina é possível construir sistemas de segurança eficientes e baratos.

A biblioteca de código aberto dlib¹ é uma candidata louvável para a construção destes sistemas devido aos excelentes resultados em respeitáveis *benchmarks* como *Labeled Faces in the Wild*. No entanto, após o treino do sistema com datasets reconhecidos², era necessário agrupar faces similares em clusters.

Contudo tal técnica tornou-se inviável para o escopo do projeto, pois a clusterização hierárquica não suporta nativamente o tratamento de novos dados inseridos no sistema e, por isso, seria necessário clusterizar os dados novamente a cada nova face capturada, demandando muito tempo para para classificar imagens de um vídeo.

Com objetivo de contornar o problema, estudou-se redes neurais profundas convolucionais[1] para entender as principais técnicas usadas para reconhecer e clusterizar imagens similares.

Resultados e Discussão

A identificação de faces em imagens é feita detectando posições de rostos e, em seguida, extraindo uma codificação adequada para tal face. Assim, cada face possui um ID que é sua referência. Para melhorar a acurácia deste processo de indexação, algoritmos de clusterização hierárquica foram utilizados para distanciar faces não similares, simplificando a classificação de uma imagem recém chegada ao sistema.

No entanto, foi verificado a ineficácia da classificação na entrada de novos dados nos algoritmos construídos com tais ferramentas. Na ocasião de

classificação de novas faces, existia a necessidade de executar o algoritmo novamente, ou seja, executar a clusterização para todo novo fluxo de dados, impossibilitando a criação de um classificador de novas faces, objetivo proposto no início deste projeto. Então, uma mudança de abordagem se viu obrigatória.

Com a construção de uma rede neural profunda específica, é possível utilizar as técnicas de detecção e indexação de faces e viabilizar a classificação de novas faces, seguindo os exemplos pesquisados, como DeepFace[2], DeepID[3] e FaceNet[4], comentados em artigo recente[1] sobre um sistema de estado-da-arte para detecção, identificação e verificação de faces.

Conclusões

A construção de sistemas de segurança eficientes e baratos que implementam técnicas do estado-da-arte em reconhecimento facial trás muitos benefícios para a indústria e seu desenvolvimento.

A biblioteca dlib pode ser uma boa ferramenta para tais sistemas, mas ainda é necessário um bom algoritmo de clusterização de dados para um bom sistema de reconhecimento de faces.

A construção de redes neurais consegue uma alta precisão em identificação e reconhecimento de imagens faciais extraídas de vídeo.

Agradecimentos

Agradeço à CNPq pela viabilização de recursos para o comprimento deste projeto. Também agradeço ao Erick Seolin, colaborador da NeuralMind, pelo suporte na compreensão e uso da dlib.

[1] R. Ranjan, A. Bansal, J. Zheng, H. Xu, J. Gleason, B. Lu, A. Nanduri, J.-C. Chen, C. Castillo, and R. Chellappa, "A fast and accurate system for face detection, identification, and verification," *CoRR*, vol. abs/1809.07586, 2018.

[2] Y. Taigman, M. Yang, M. A. Ranzato, and L. Wolf, "Deepface: Closing the gap to human-level performance in face verification," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 1701–1708.

[3] Y. Sun, X. Wang, and X. Tang, "Deep learning face representation from predicting 10000 classes," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014, pp. 1891–1898.

[4] F. Schroff, D. Kalenichenko, and J. Philbin, "Facenet: A unified embedding for face recognition and clustering," *arXiv preprint arXiv:1503.03832*, 2015.

¹ <http://dlib.net/>

²PIPA<<https://people.eecs.berkeley.edu/~nzhang/piper.html>> e LFW <<http://vis-www.cs.umass.edu/lfw/>>