



Extração de Características utilizando Redes Neurais Convolucionais

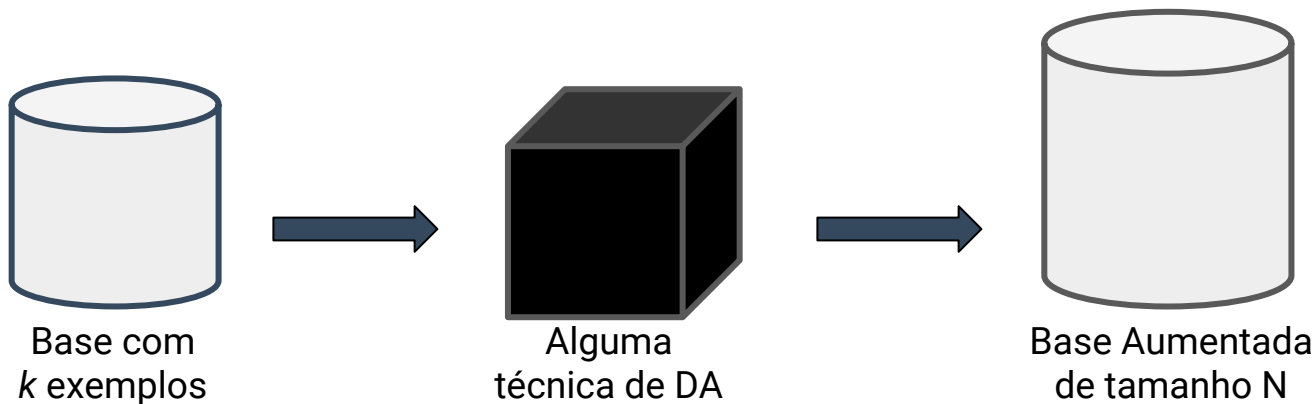
Willian Dihanster Gomes de Oliveira RA: 112269

Introdução

- Em Aprendizado de Máquina (AM), visando ter bons resultados, é importante ter bastante dados.
- Só que, às vezes conseguir dados é uma tarefa difícil e custosa.
 - Como obter mais dados?

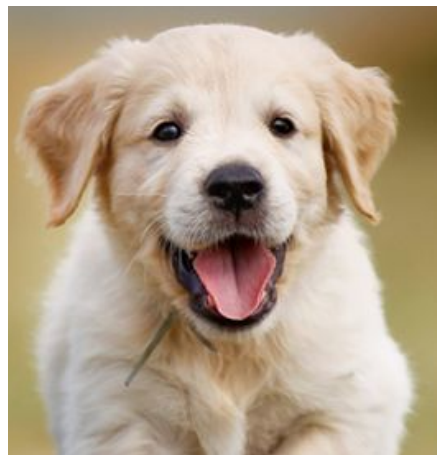
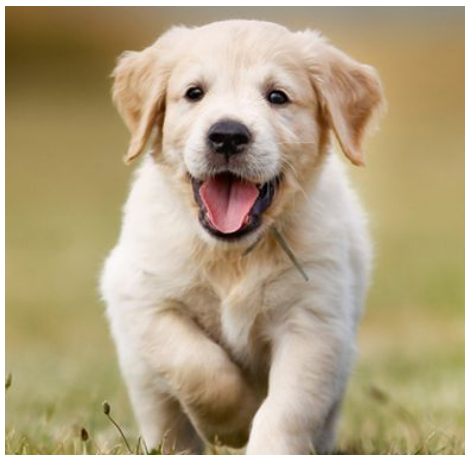
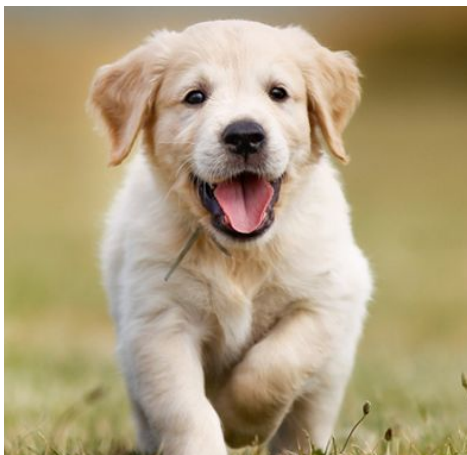
Data Augmentation

- São técnicas computacionais para aumentar bases de dados.



Data Augmentation

- Exemplo utilizando *flip* e corte.
- No projeto: corte, brilho, rotação, *zoom*, *flip*, *shift*...

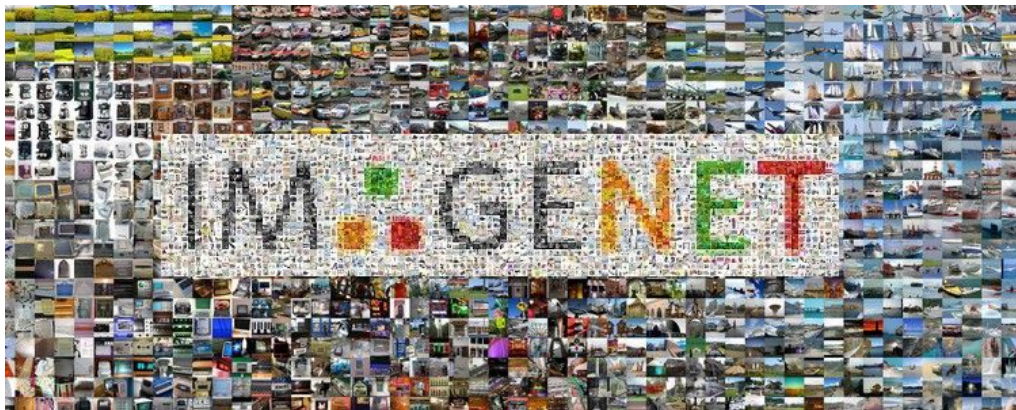


Extração de Características

- As vezes, é necessários representar imagens por um vetor de características.
- Há diversos algoritmos descritores de imagens.
- Pode-se também utilizar os vetores de *features* de uma rede convolucional.

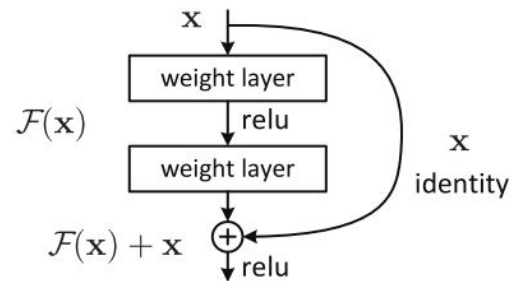
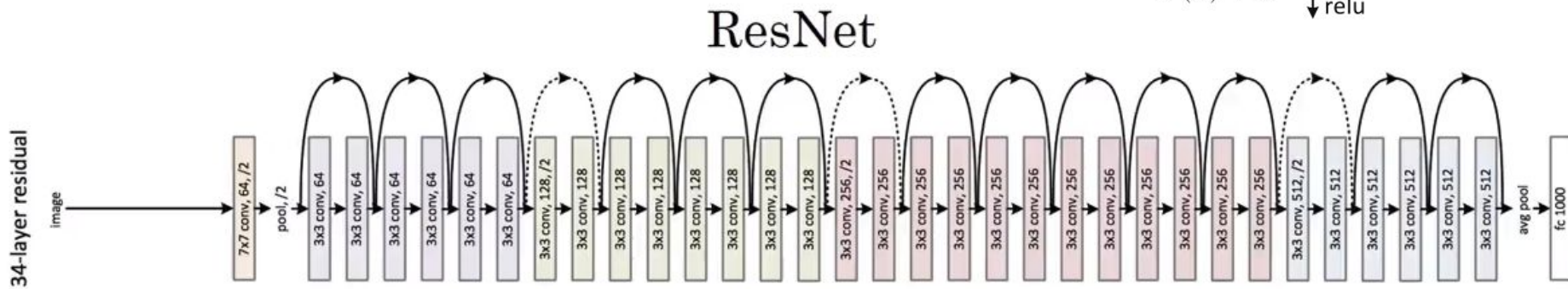
Redes Convolucionais

- Foram utilizadas 4 redes convolucionais diferentes
 - ResNet50, VGG16, VGG19, Xception.
- Pré-treinadas com a *ImageNet* [1]
 - 1000 classes; 1.2 milhões treino; 100mil teste.



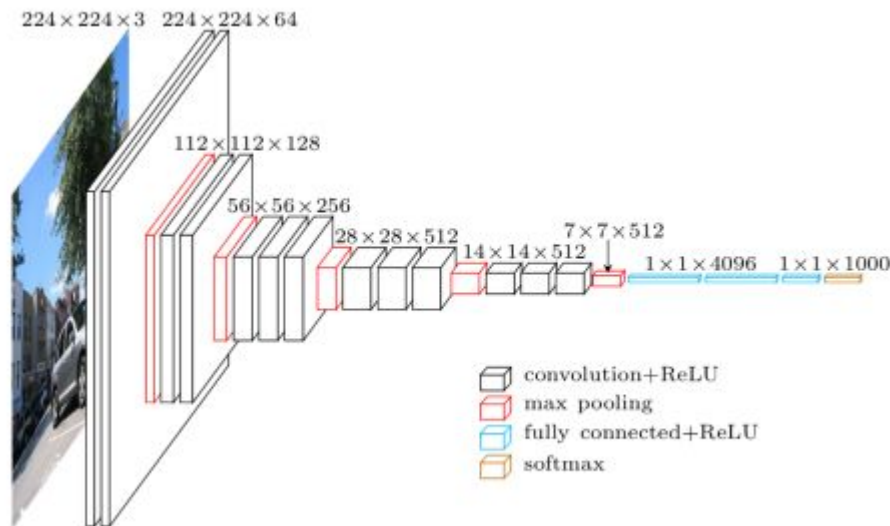
ResNet50

- É uma rede residual de 50 camadas. [2]
- O vetor de características da última camada, antes da classificação, é (1, 2048).



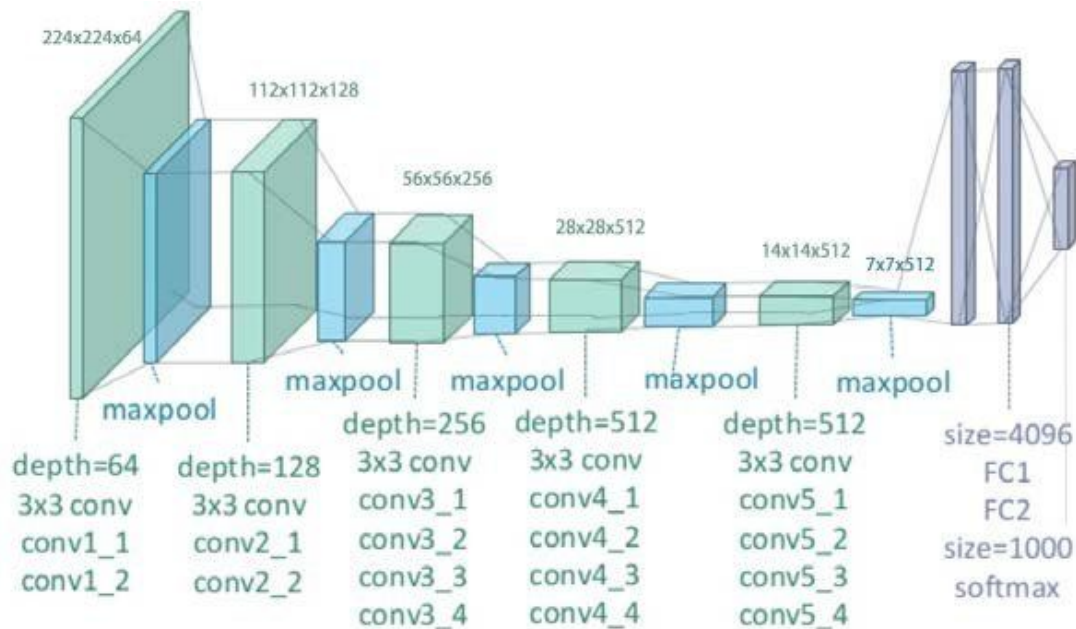
VGG16

- É uma CNN de 16 camadas de convolução. [3]
- O vetor de características é (1, 4096).



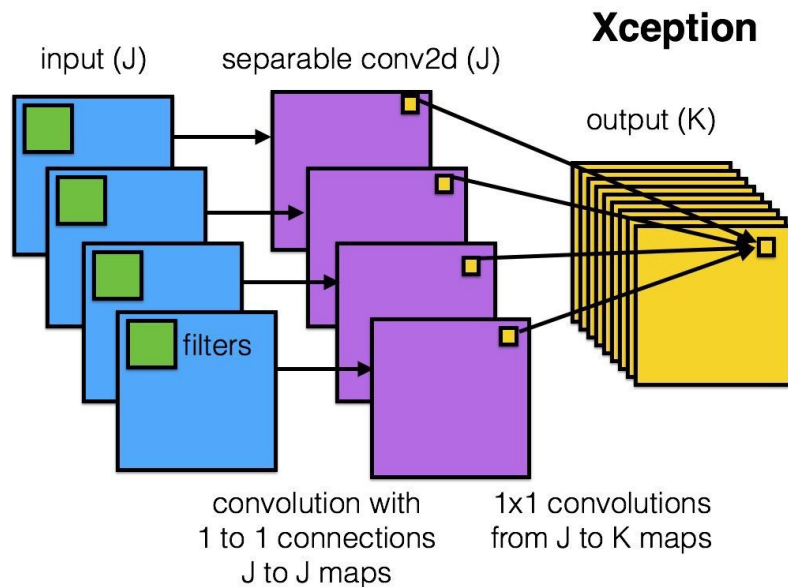
VGG19

- É uma CNN de 19 camadas de convolução. [3]
- O vetor de características é (1, 4096).



Xception

- Xception - Xtreme Inception, com convoluções separadas. [4]
- Vetor de características (1, 4096).



Bases de Dados: Brazilian Coffee Scenes Dataset

- Imagens de satélite de plantações de café, em cidades de Minas Gerais pelo sensor SPOT. [5]
- Cada imagem possui 64x64 pixels. 2876 imagens, 50% de café e 50% não café. 2010 treino e 866 de teste.



Exemplos rotulados como “coffee”.



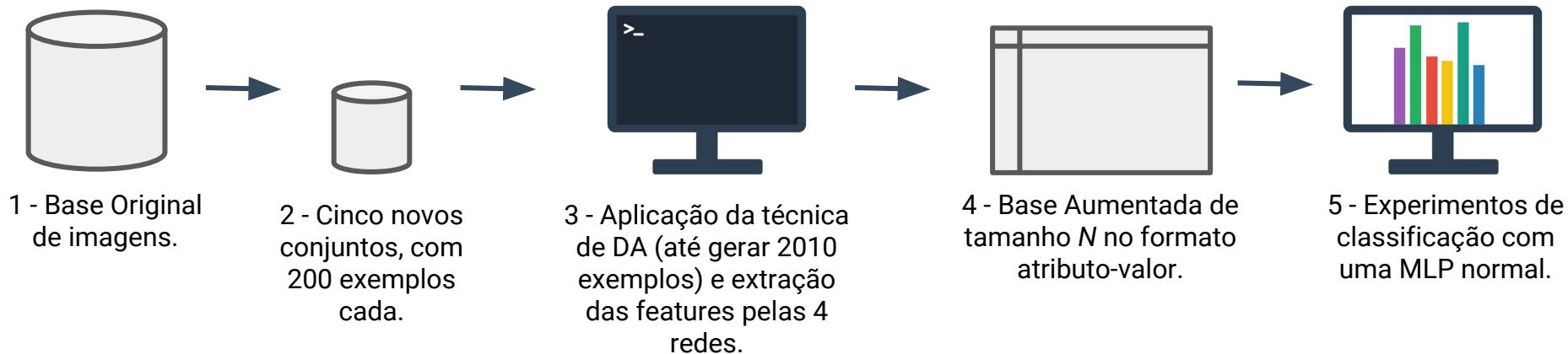
Exemplos rotulados como “noncoffee”.

Ferramentas

- Python 3.6
- TensorFlow
- Keras
- Numpy

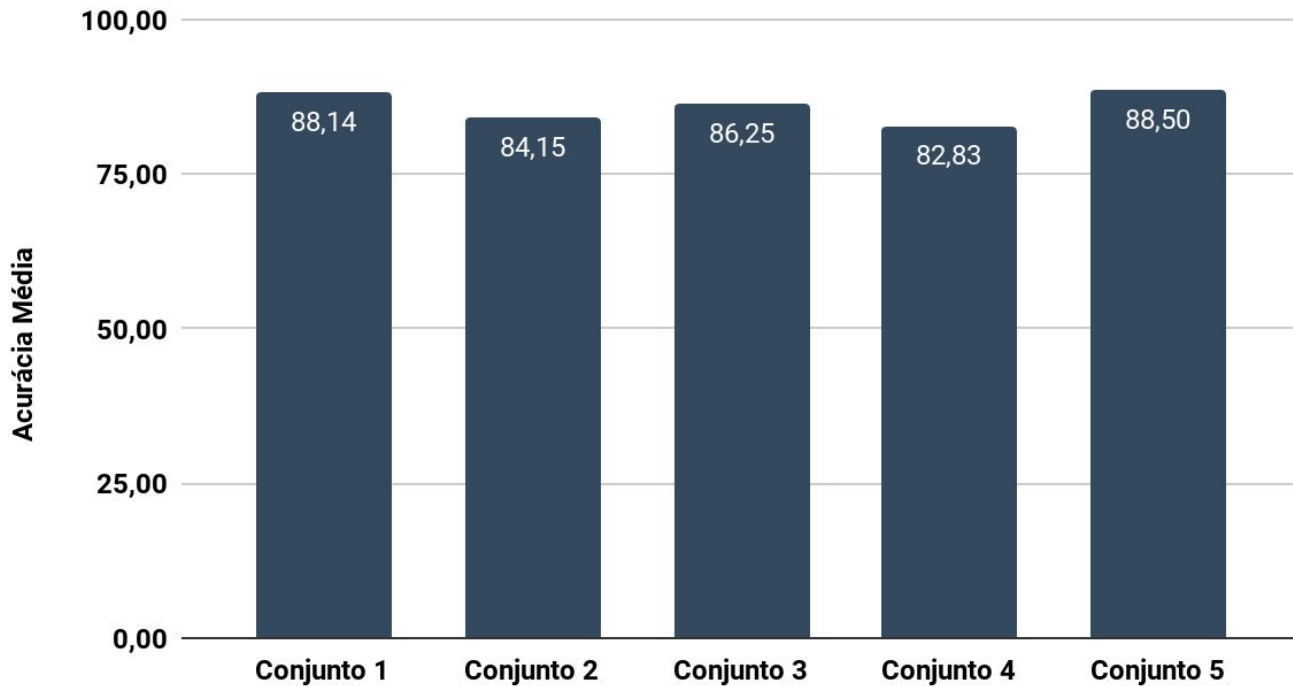


Metodologia



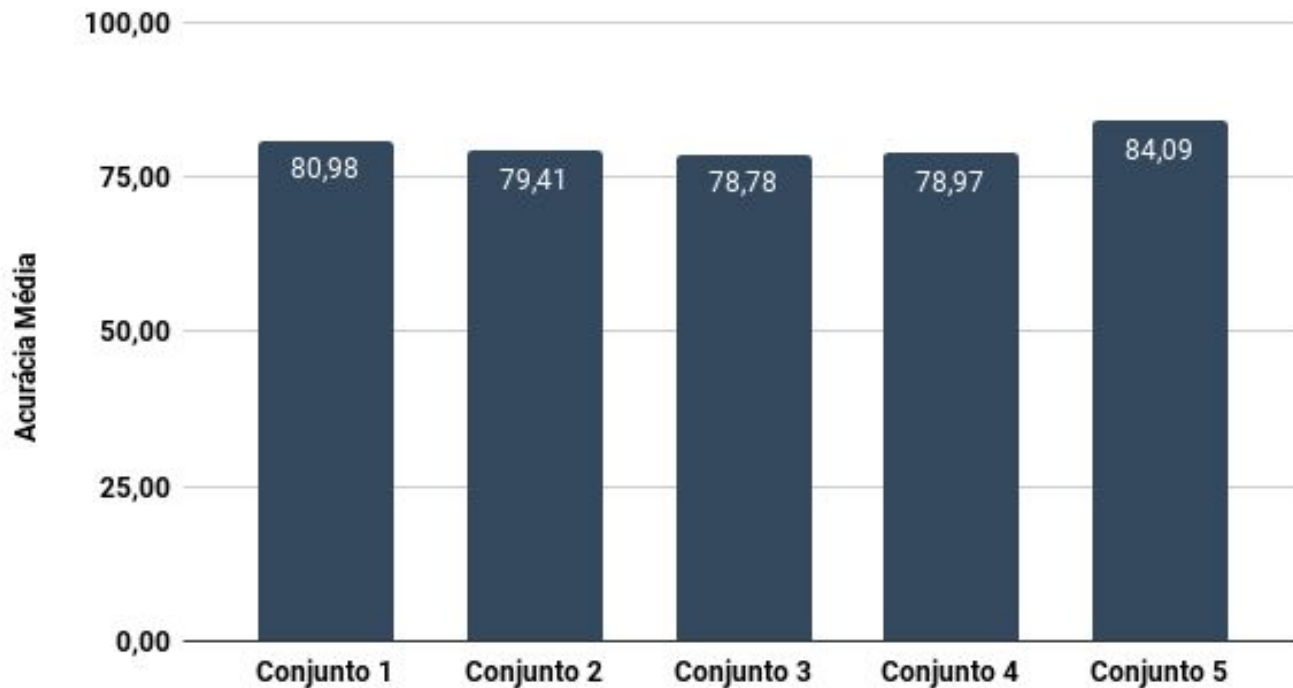
Resultados (ResNet50)

ResNet50



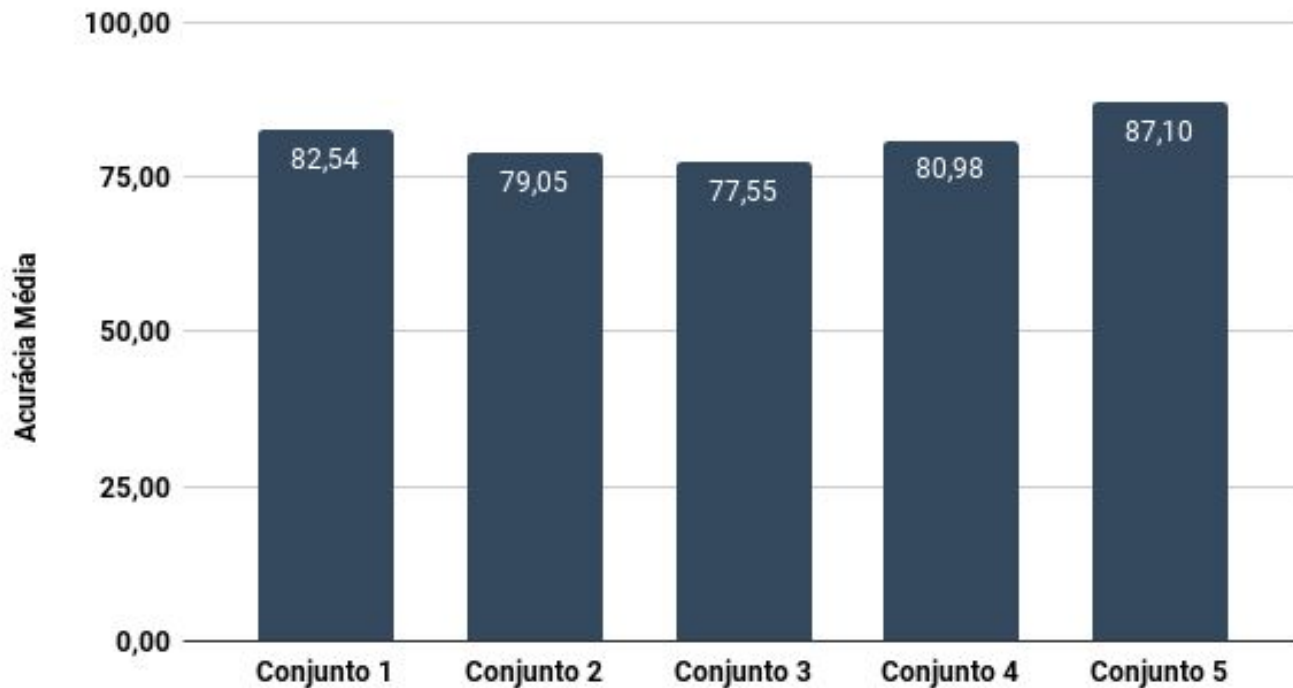
Resultados (VGG16)

VGG16

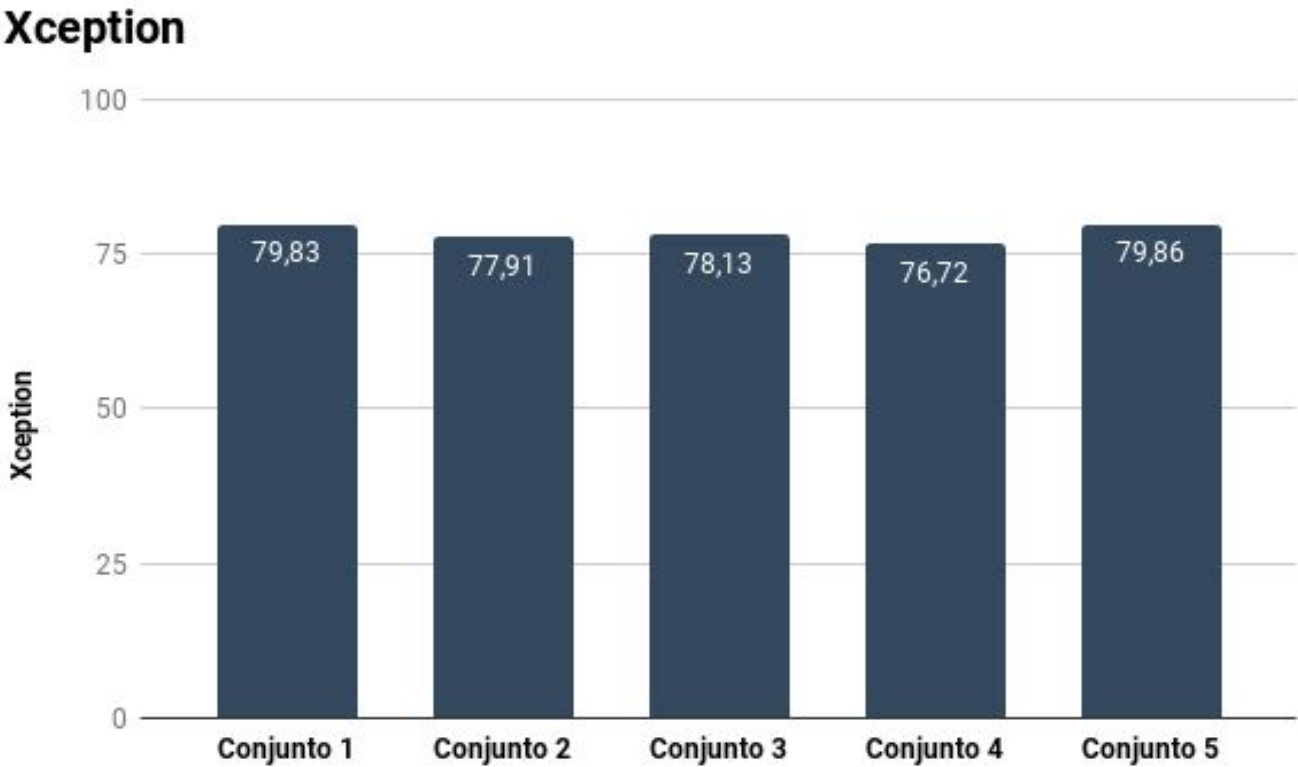


Resultados (VGG19)

VGG19

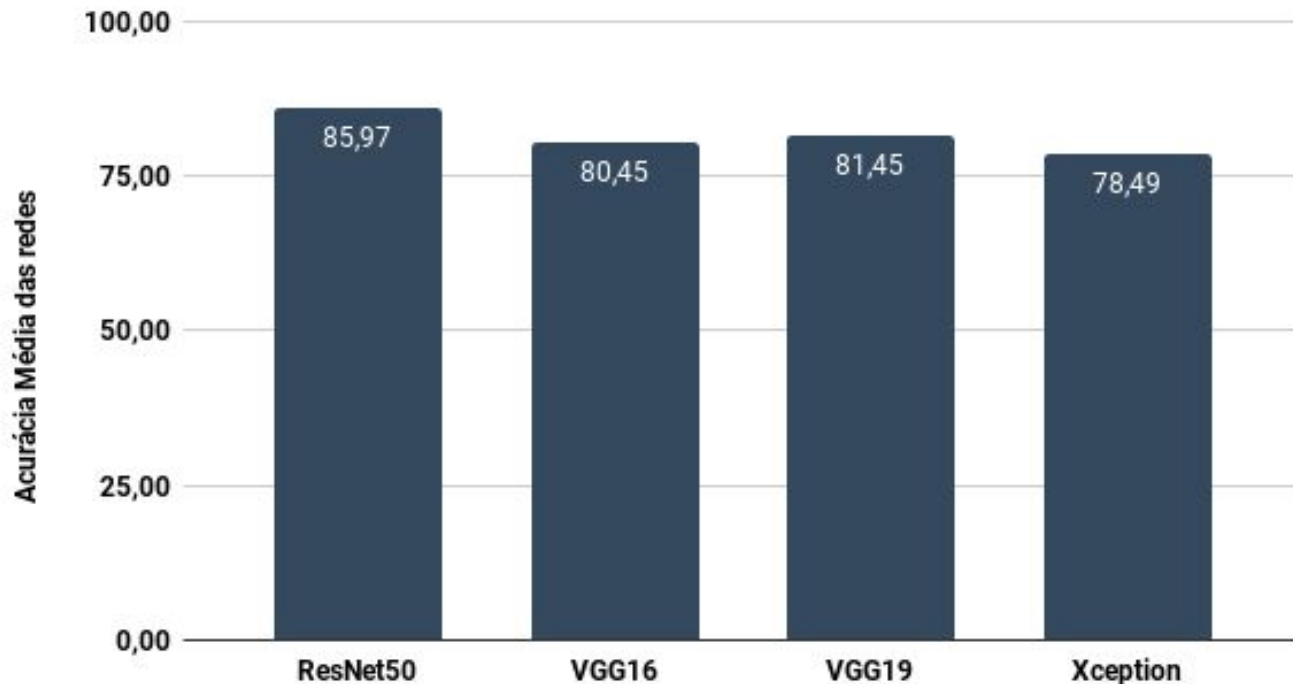


Resultados (Xception)



Resultados das redes para os conjuntos

Acurácia Média da Redes para os conjuntos



Conclusões

- Técnicas de DA podem ajudar no aumento da acurácia.
 - Vantagem: para qualquer conjunto de imagens.
 - Desvantagem: perda de sentido ou redundância.
- Redes convolucionais podem também ser utilizadas como um bom extrator de características.
 - Em especial, a ResNet que obteve melhores resultados neste domínio.

Referências

- [1] Deng, Jia, et al. "Imagenet: A large-scale hierarchical image database." *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. Ieee, 2009.
- [2] Targ, Sasha, Diogo Almeida, and Kevin Lyman. "Resnet in Resnet: generalizing residual architectures." *arXiv preprint arXiv:1603.08029* (2016).
- [3] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." *arXiv preprint arXiv:1409.1556* (2014).
- [4] Chollet, François. "Xception: Deep learning with depthwise separable convolutions." *arXiv preprint* (2017): 1610-02357.
- [5] Penatti, Otávio AB, Keiller Nogueira, and Jefersson A. dos Santos. "Do deep features generalize from everyday objects to remote sensing and aerial scenes domains?." *Computer Vision and Pattern Recognition Workshops (CVPRW), 2015 IEEE Conference on*. IEEE, 2015.



Obrigado!