

VGG16 com Adaptação para Detecção de Pneumonia Através de Imagens de Raios-X

Lavinia Mitiko Takarabe, André Kazuo Takahata

Universidade Federal do ABC, Av. dos Estados, 5001
lavinia.mitiko@aluno.ufabc.edu.br, andre.t@ufabc.edu.br

Abstract: Neste artigo é detalhado o processo de desenvolvimento de um algoritmo capaz de identificar casos de pneumonia a partir de raios-x da região torácica, se utilizando de uma rede neural convolucional VGG16, em que propomos a realização de adaptações. Foi utilizada uma base de dados do Kaggle contando com imagens para 14 patologias torácicas diferentes. Ao fim, obtivemos um F1-score de 0,462, valor compatível com o estado da arte da literatura.

1. Introdução

Nos últimos anos, houve um grande desenvolvimento das técnicas de deep learning para a classificação de imagens, incluindo imagens médicas [1]. Nesse contexto, realizamos uma análise inicial para o desempenho de uma rede VGG-16 [3] com adaptações para a detecção de pneumonia por meio de imagens de raio-X e comparamos com resultados do estado da arte na literatura [2, 4, 5].

2. Metodologia

O banco de dados utilizado foi retirado do Kaggle¹ contando com 112.120 imagens de raio-X frontais do tórax a partir de duas vistas diferentes AP (anteroposterior) e PA (posteroanterior) de 30.805 sujeitos únicos, abrangendo 14 diferentes patologias torácicas comuns (atelectasia 10,3%, consolidação 4,16%, infiltração 17,74%, pneumotórax 4,72%, edema 2,05%, efisema 2,44%, fibrose 1,50%, efusão 11,88%, pneumonia 1,28%, espessamento pleural 3,02%, cardiomegalia 2,47%, nódulo 5,65%, massa 5,15% e hérnia 0,20%) e um grupo de pacientes examinados não relacionados a nenhuma das patologias 53,84%. Para garantir o balanceamento das imagens para a Pneumonia na etapa de treinamento, foram utilizados 2.290 imagens (de forma que 50% destas estivessem associadas à condição e 50% não) e para a validação 1430 aleatoriamente escolhidas.

Em um modelo VGG16 (Fig. 1) as primeiras camadas convolutivas capturam características baseadas em linhas e bordas, uma vez que estas tem campos receptivos muito pequenos, podendo ser utilizadas para qualquer classificação de imagens [3]. Desta forma, é interessante reutilizar ou congelar os pesos pré-treinados das primeiras camadas. Assim só resta treinar os pesos das últimas camadas que são utilizados para distinguir casos específicos. Este ajuste de pesos nas camadas finais é chamado de ajuste fino. Para o ajuste fino foram congeladas as doze primeiras camadas convolucionais pré-treinadas, e as últimas block5conv3 e block5pool tiveram seus parâmetros ajustados, durante o treinamento. Além disso, foram adicionadas sequencialmente camadas do tipo: *flatten*, *dropout*, *dense*, *dropout* e *dense*.

3. Resultados

Tomou-se como limiar para a decisão o ponto que maximiza o F1-score. Esta escolha resultou na sensibilidade de 0,753 e especificidade de 0,582, isso indica que terão mais falsos positivos e menos falso negativos. O F1-score é dado por $2 \cdot \frac{(\text{precisão} \cdot \text{revocação})}{(\text{precisão} + \text{revocação})}$. Pode-se dizer que o algoritmo teve um desempenho comparável numericamente aos valores apresentados em [2], apresentando F1-score de 0,462, ultrapassando o F1-score de radiologistas que apresentaram coeficiente médio de 0,387 e também ao resultado da rede CheXNet de 0,435 [2]. O AUC do classificador de detecção de pneumonia proposto foi de 0,782, ultrapassando valores da literatura de 0,633 em Wang et al. [5], 0,713 em Yao et al. [4] e Rajpurkar et al. [2] de 0,768. E a rede com adaptação proposta desempenha melhor que a VGG16 sem adaptação que, treinada com o mesmo conjunto de dados e sob a mesma estratégia de divisão, apresentou F1-score de 0,399 e AUC de 0,601.

¹<https://www.kaggle.com/nih-chest-xray>

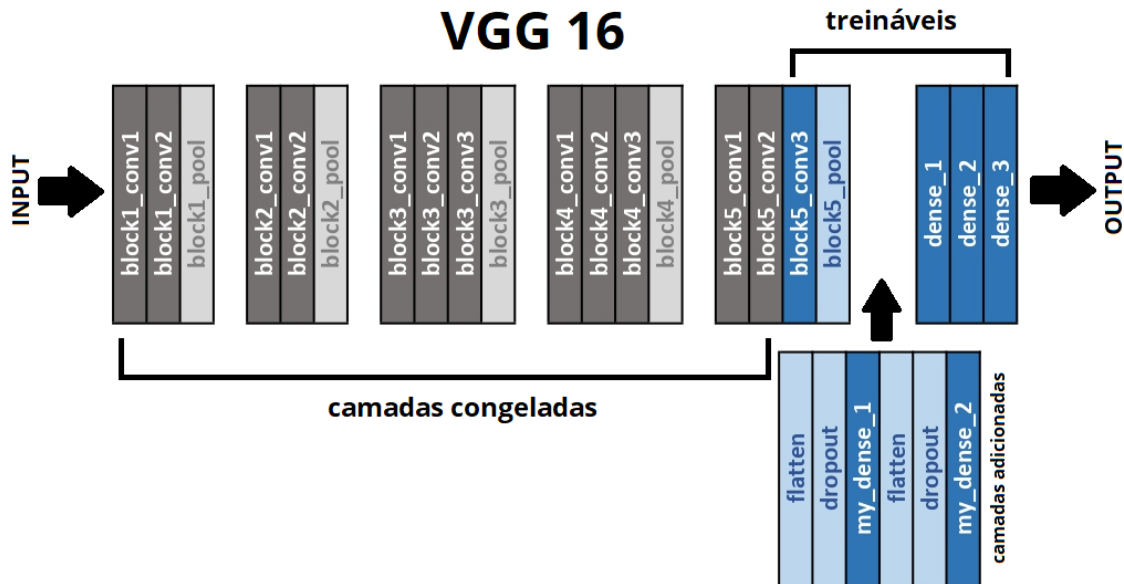


Fig. 1. Arquitetura da rede VGG 16 baseada no artigo [3]. Esta figura ilustra a arquitetura da rede VGG16, com 13 camadas do tipo convolucional e as três camadas do tipo *fully connected* (FC), indicando as adaptações realizadas, o congelamento das 12 primeiras camadas, as camadas treináveis e as camadas adicionadas.

4. Discussão

É necessário explicitar que estes são resultados preliminares e ainda são necessários mais testes e desenvolvimento para melhoria de sensibilidade e especificidade, assim como uma melhor comparação com a literatura, visto que a base de dados e a estratégia de validação neste trabalho e no desenvolvimento da CheXNet são diferentes. Ainda assim, o resultado obtido demonstra que uma rede VGG16 com adaptações possui potencial para apresentar um desempenho comparável com ao estado da arte a partir de 18 camadas treináveis em contraste às 121 camadas do modelo CheXNet.

5. References

- [1] Currie, Geoff, et al. "Machine learning and deep learning in medical imaging: intelligent imaging." *Journal of Medical Imaging and Radiation Sciences* 50.4 (2019): 477-487.
- [2] Rajpurkar, Pranav, et al. "Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning." *arXiv preprint arXiv:1711.05225* (2017).
- [3] Zhang, Xiangyu, et al. "Accelerating very deep convolutional networks for classification and detection." *IEEE transactions on pattern analysis and machine intelligence* 38.10 (2015): 1943-1955.
- [4] Yao, Li, et al. "Learning to diagnose from scratch by exploiting dependencies among labels." *arXiv preprint arXiv:1710.10501* (2017).
- [5] Wang, X., et al. "Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases." *IEEE CVPR*. 2017..