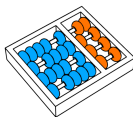


Classificação de Imagens de Radiografias de Tórax Utilizando Redes Neurais Profundas

Vinicius Teixeira de Melo



Orientador: Prof. Dr. Zanoni Dias
Coorientador: Prof. Dr. Hélio Pedrini

17 de Outubro de 2019

Agenda

- 1 Introdução
- 2 Conceitos e Trabalhos Relacionados
- 3 Material e Métodos
- 4 Resultados Preliminares
- 5 Plano de Trabalho

- 1 Introdução
- 2 Conceitos e Trabalhos Relacionados
- 3 Material e Métodos
- 4 Resultados Preliminares
- 5 Plano de Trabalho

Caracterização do Problema

- Radiografia de tórax é um dos exames mais comuns globalmente.
- Identificar condições cardiopulmonares agudas e crônicas.
- Interpretação automatizada da radiografia de tórax.
- *CheXpert* [1] e *MIMIC-CXR* [2].

Motivação

- Distinguir as diferentes patologias torácicas é uma tarefa difícil.
- Infiltrações no pulmão, dispositivos de suporte e anormalidades de tamanho ou contorno do coração [3].
- Estima-se que os erros de diagnóstico estejam entre 14 e 38% [4].

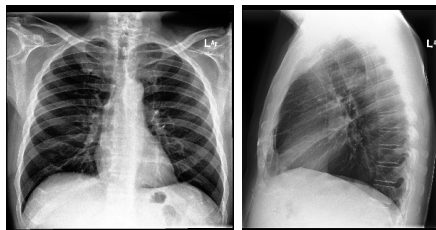


Figura: Radiografia frontal e lateral de um mesmo paciente.

Objetivo Geral

Investigar técnicas recentes de classificação de imagens e propor uma abordagem baseada em Redes Neurais Profundas para classificação de radiografias de tórax.

Objetivos Específicos

- Estudar as abordagens de classificação de imagens.
- Propor uma solução baseada em rede neural profunda.
- Validar o método proposto em diferentes bases de dados.
- Comparar e avaliar o método proposto com outras abordagens.
- Publicar os resultados.

Contribuições

Desenvolvimento de uma metodologia para classificação de patologias em imagens de radiografias de tórax a partir da extração de áreas de interesse das imagens de entrada.

Questões de Pesquisa

- É possível obter resultados competitivos em relação aos trabalhos recentes em classificação de imagens de radiografias de tórax utilizando redes neurais profundas?
- Considerando a classificação multirrótulo e a divisão em classificadores binários, qual método obtém melhores resultados no problema de classificação de imagens de radiografia de tórax?
- Quais são as implicações de usar imagens de alta resolução como entrada para a rede neural?

Questões de Pesquisa

- Como os pesos de redes pré-treinadas na base *ImageNet* e, em seguida, aperfeiçoados para o domínio médico, se comparam aos pesos aprendidos somente a partir de imagens médicas?
- Como tratar a incerteza sobre os rótulos em imagens da base de dados *CheXpert* [1]?
- É possível generalizar nosso método para outras bases de dados de radiografias?

Agenda

- 1 Introdução
- 2 Conceitos e Trabalhos Relacionados**
- 3 Material e Métodos
- 4 Resultados Preliminares
- 5 Plano de Trabalho

Rede Neural Convolutacional Profunda

- Modelos inspirados no córtex visual do cérebro humano.
- Extração de características.
- Hierarquia de camadas.
- Conectividade esparsa e compartilhamento de parâmetros.

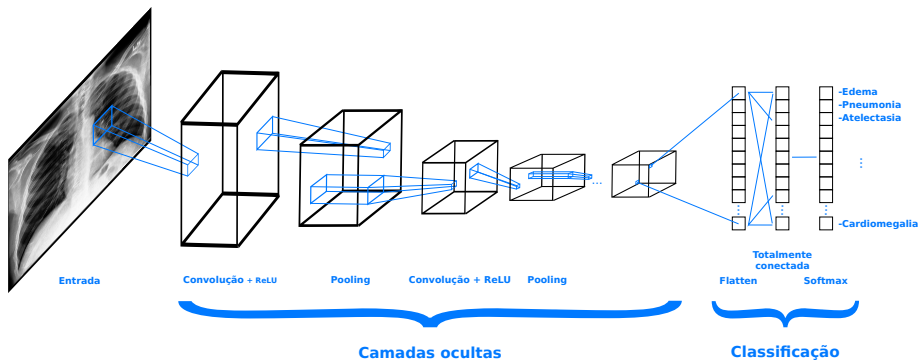


Figura: Classificação usando uma Rede Neural Convolutacional Profunda.

Rede Neural Recorrente

- Processamento de informações sequenciais ou dependentes.
- Utiliza os resultados já computados no estado atual.
- Entrada esperada é um vetor codificado.
- Memória de Curto Longo Prazo (LSTM).

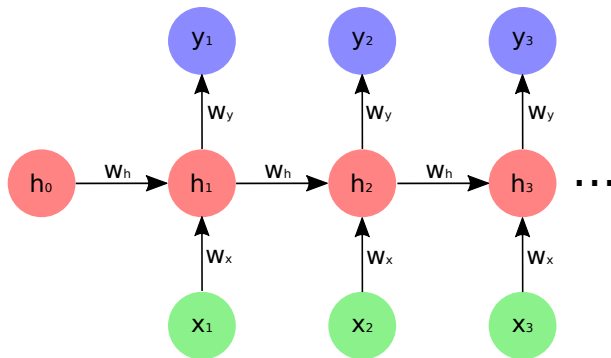


Figura: Exemplo de RNN.

Transferência de Aprendizado

- Reuso de modelos.
- Diminuição do tempo de treinamento.

Aumentação de Dados

- Busca solucionar o problema de falta de dados.
- Aplica transformações nas imagens.
- Translação, rotação e escala.

Wang et al. (2017) [5]

- Classificação e localização de patologias.
- Modelos pré-treinados na *ImageNet*: *AlexNet*, *GoogLeNet*, *VGGNet-16* e *ResNet-50*.
- Utilizaram a base de dados *ChestX-ray8*.
- *ResNet-50* obteve AUROC 0,696 para classificação e acurácia 0,628 para localização.

Trabalhos Correlatos

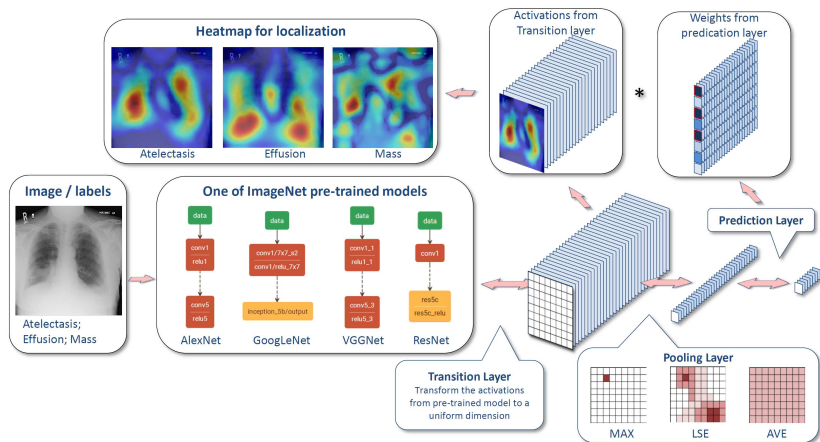


Figura: Fluxograma geral da estrutura DCNN unificada e o processo de localização das áreas de interesse, proposto por Wang et al. (2017) [5].

Yao et al. (2017) [6]

- Maximização da necessidade de extrair características relevantes.
- Baseado na arquitetura das redes *DenseNets* para codificação.
- A decodificação é feita utilizando RNN.
- Utilizou a base *ChestX-ray14*.
- Obteve AUROC de 0,802.

Trabalhos Correlatos

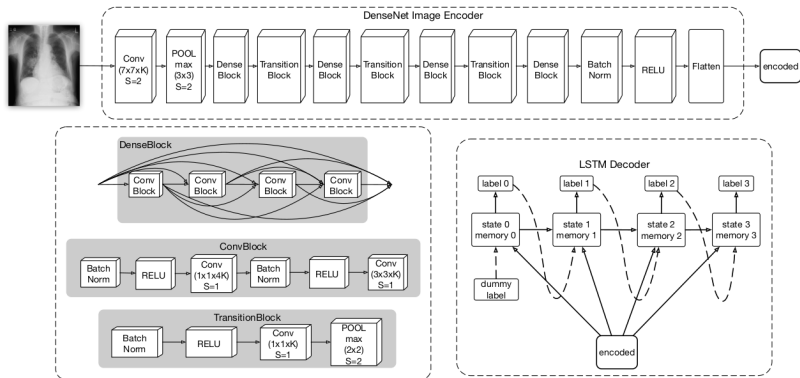


Figura: Representação do processo de codificação e decodificação baseado na arquitetura da rede *DenseNet*, proposta por Yao et al. (2017) [6].

Rajpurkar et al. (2017) [7]

- Algoritmo para detectar pneumonia em radiografias.
- Baseado na *DenseNet121*.
- Classificação Binária × Classificação Multirrótulo.
- Utilizou a base *ChestX-ray14*.
- Obteve AUROC de 0,841.

Agenda

- 1 Introdução
- 2 Conceitos e Trabalhos Relacionados
- 3 Material e Métodos**
- 4 Resultados Preliminares
- 5 Plano de Trabalho

Material e Métodos

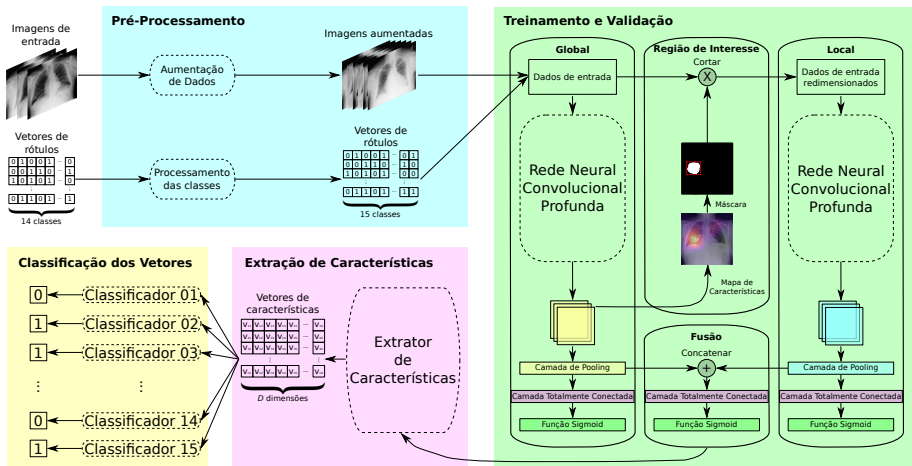


Figura: Etapas que compõem a metodologia deste projeto.

Métricas de Avaliação

$$\text{Acurácia} = \frac{TP + TN}{P + N}$$

$$\text{Acurácia balanceada} = \frac{1}{2} \left(\frac{TP}{P} + \frac{TN}{N} \right)$$

$$\text{Sensibilidade} = \frac{TP}{TP + FN}$$

$$\text{Especificidade} = \frac{TN}{TN + FP}$$

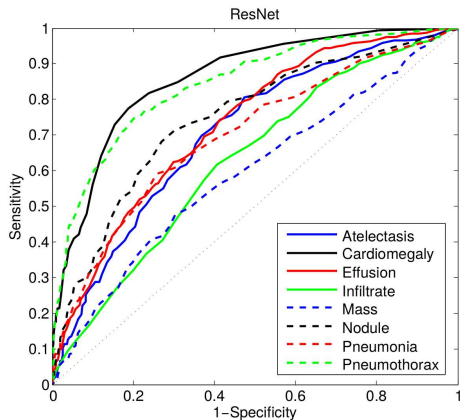


Figura: Exemplo de AUROC obtido com a DCNN *ResNet-50* [5].

CheXpert

- Foi disponibilizada para um desafio¹.
- Possui 224316 imagens de 65240 pacientes.
- Dividida em 14 tipos de observações.
- As classes podem ser inferidas como incerta.

¹<https://stanfordmlgroup.github.io/competitions/chexpert/>

Classe	Treinamento			Validação			Teste		
	Positivo	Incerto	Negativo	Positivo	Incerto	Negativo	Positivo	Incerto	Negativo
Normal	15576	0	141051	2163	0	20408	4680	0	39770
Cardiomiastinal aumentado	7632	8596	140399	1079	1358	20134	2196	2449	39805
Cardiomegalia	18840	5657	132130	2783	828	18960	5445	1602	37403
Opacidade pulmonar	74154	3837	78636	10766	559	11246	20787	1202	22461
Lesão pulmonar	6434	1058	149135	963	141	21467	1790	289	42371
Edema pulmonar	36775	9211	110641	5324	1241	16006	10192	2532	31726
Consolidação pulmonar	10412	19472	126743	1479	2822	18270	2925	5448	36077
Pneumonia	4277	13184	139166	625	1975	19971	1145	3611	39694
Atelectasia	23464	23588	109575	3317	3377	15877	6675	6774	31001
Pneumotórax	13564	2187	140876	1974	333	20264	3918	625	39907
Derrame pleural	60147	8217	88263	8746	1098	12727	17361	2313	24776
Outros problemas pleurais	2409	1834	152384	346	248	21977	769	571	43110
Fratura	6329	471	149827	935	61	21575	1776	110	42564
Dispositivos de suporte	81363	758	74506	11893	107	10571	22852	214	21384

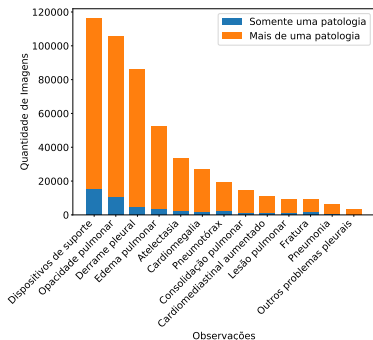
Tabela: Divisão da base de dados *CheXpert*.

ChestX-ray14

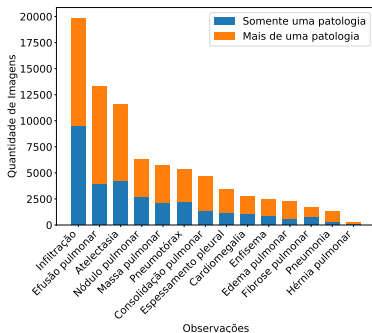
- Possui 112120 imagens de 30805 pacientes.
- Dividida em 14 tipos de patologias.
- As classes só podem ser inferidas como positiva ou negativa.

Classe	Treinamento		Validação		Teste	
	Positivo	Negativo	Positivo	Negativo	Positivo	Negativo
Atelectasia	8103	70381	1128	10083	2304	20121
Cardiomegalia	1940	76544	268	10943	564	21861
Efusão pulmonar	9347	69137	1285	9926	2675	19750
Infiltração	13965	64519	1952	9259	3953	18472
Massa pulmonar	4017	74467	581	10630	1148	21277
Nódulo pulmonar	4452	74032	617	10594	1254	21171
Pneumonia	948	77536	139	11072	266	22159
Pneumotórax	3716	74768	526	10685	1056	21369
Consolidação pulmonar	3258	75226	493	10718	916	21509
Edema pulmonar	1634	76850	227	10984	442	21983
Enfisema	1769	76715	250	10961	497	21928
Fibrose pulmonar	1188	77296	146	11065	352	22073
Espessamento pleural	2397	76087	323	10888	665	21760
Hérnia pulmonar	167	78317	18	11193	42	22383

Tabela: Divisão da base de dados *ChestX-ray14* [8].



(a) *CheXpert*.



(b) *ChestX-ray14*.

Figura: Distribuição do número de patologias por classe.

Recursos Computacionais

- Laboratório de Informática Visual (LIV).
- Intel i7-3770 3.50 GHz.
- 32 GB de memória RAM.
- GPU NVIDIA TITAN V.

Agenda

- 1 Introdução
- 2 Conceitos e Trabalhos Relacionados
- 3 Material e Métodos
- 4 Resultados Preliminares**
- 5 Plano de Trabalho

Configurações do Treinamento

- Otimizador *Adam*.
- Taxa de aprendizagem 0,001.
- Entropia Cruzada Binária.
- 100 épocas.
- Rotação Horizontal Aleatória.
- Métrica AUROC.

Resultados Preliminares

Classe	Wang et al. [8]	Yao et al. [6]	Rajpurkar et al. [7]
Atelectasia	0,7003	0,772	0,8094
Cardiomegalia	0,8100	0,904	0,9248
Efusão pulmonar	0,7585	0,859	0,8638
Infiltração	0,6614	0,695	0,7345
Massa pulmonar	0,6933	0,792	0,8676
Nódulo pulmonar	0,6687	0,717	0,7802
Pneumonia	0,6580	0,713	0,7680
Pneumotórax	0,7993	0,841	0,8887
Consolidação pulmonar	0,7032	0,788	0,7901
Edema pulmonar	0,8052	0,882	0,8878
Enfisema	0,8330	0,829	0,9371
Fibrose pulmonar	0,7859	0,767	0,8047
Espessamento pleural	0,6835	0,765	0,8062
Hérnia pulmonar	0,8717	0,914	0,9164
Média da validação	0,7451	0,8027	0,8413

Tabela: Resultados da literatura obtidos na base de dados *ChestX-ray14* [8].

Resultados Preliminares

Classe	DenseNet169	DenseNet201	DenseNet121	GoogLeNet	VGG16	ResNet50	AlexNet
Atelectasia	0,8319	0,8324	0,8285	0,8214	0,8187	0,8141	0,7714
Cardiomegalia	0,9085	0,8988	0,9009	0,8893	0,8729	0,8819	0,8851
Efusão pulmonar	0,8922	0,8940	0,8904	0,8859	0,8884	0,8845	0,8607
Infiltração	0,7268	0,7205	0,7149	0,7065	0,7145	0,7077	0,6950
Massa pulmonar	0,8712	0,8779	0,8591	0,8506	0,8510	0,8496	0,7723
Nódulo pulmonar	0,7844	0,7799	0,7628	0,7604	0,7683	0,7514	0,6784
Pneumonia	0,7872	0,7771	0,7655	0,7713	0,7402	0,7615	0,7283
Pneumotórax	0,8988	0,8891	0,8812	0,8732	0,8780	0,8678	0,8342
Consolidação pulmonar	0,8137	0,8147	0,8099	0,8085	0,7944	0,7972	0,7802
Edema pulmonar	0,9072	0,9043	0,8971	0,8924	0,8854	0,8921	0,8767
Enfisema	0,9229	0,9215	0,9063	0,9061	0,9092	0,9080	0,8387
Fibrose pulmonar	0,8197	0,8130	0,7784	0,7978	0,7690	0,7747	0,7414
Espessamento pleural	0,8040	0,8058	0,7950	0,7848	0,7890	0,7839	0,7337
Hérnia pulmonar	0,8968	0,9054	0,8925	0,8680	0,8767	0,8512	0,8215
Média da validação	0,8475	0,8453	0,8345	0,8298	0,8255	0,8233	0,7870
Taxa de perda	0,1412	0,1423	0,1449	0,1459	0,1456	0,1502	0,1548

Tabela: Resultados obtidos na base de dados *ChestX-ray14* [8].

Resultados Preliminares

Classe	DenseNet201	DenseNet121	DenseNet169	GoogLeNet	VGG16	ResNet50	AlexNet
Normal	0,8703	0,8707	0,8720	0,8703	0,8671	0,8645	0,8479
Cardiomiastinal aumentado	0,6760	0,6680	0,6683	0,6562	0,6458	0,6604	0,6323
Cardiomegalia	0,8576	0,8506	0,8457	0,8503	0,8463	0,8526	0,8259
Opacidade pulmonar	0,7334	0,7299	0,7285	0,7303	0,7326	0,7231	0,7157
Lesão pulmonar	0,7887	0,7941	0,7880	0,7754	0,7588	0,7743	0,7252
Edema pulmonar	0,8414	0,8430	0,8411	0,8397	0,8390	0,8371	0,8226
Consolidação pulmonar	0,7510	0,7511	0,7448	0,7414	0,7416	0,7397	0,7078
Pneumonia	0,7379	0,7627	0,7525	0,7511	0,7537	0,7525	0,7014
Atelectasia	0,7035	0,6949	0,6919	0,6925	0,6912	0,6919	0,6702
Pneumotórax	0,8657	0,8618	0,8617	0,8521	0,8653	0,8503	0,8029
Derrame pleural	0,8807	0,8778	0,8812	0,8781	0,8782	0,8732	0,8550
Outros problemas pleurais	0,8120	0,8042	0,8205	0,7990	0,7872	0,7884	0,7549
Fratura	0,7516	0,7550	0,7559	0,7441	0,7288	0,7373	0,6738
Dispositivos de suporte	0,8897	0,8873	0,8876	0,8762	0,8901	0,8811	0,8410
Média da validação	0,7971	0,7965	0,7957	0,7898	0,7876	0,7876	0,7555
Taxa de perda	0,2746	0,2761	0,2765	0,2777	0,2756	0,2803	0,2925

Tabela: Resultados obtidos na base de dados *CheXpert* [1].

Agenda

- 1 Introdução
- 2 Conceitos e Trabalhos Relacionados
- 3 Material e Métodos
- 4 Resultados Preliminares
- 5 Plano de Trabalho**

Plano de Trabalho

1. Obtenção dos créditos obrigatórios em disciplinas do programa.

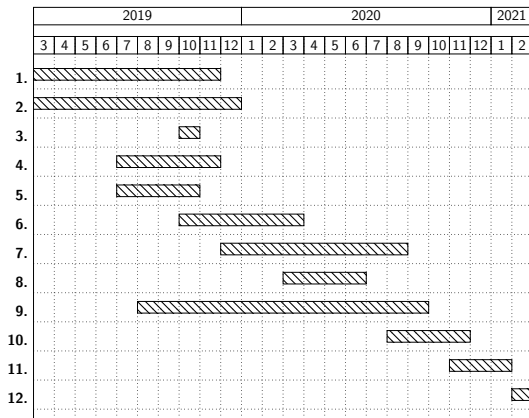


Tabela: Cronograma de atividades dividido em meses.

Plano de Trabalho

2. Revisão bibliográfica.

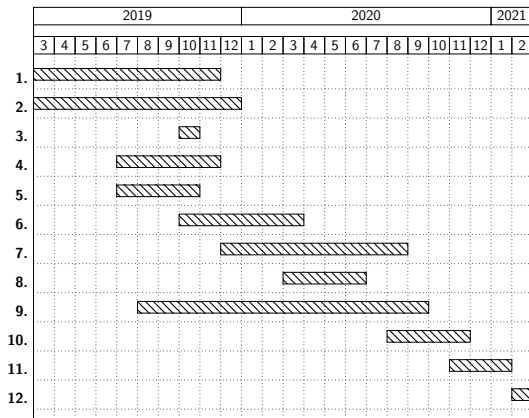


Tabela: Cronograma de atividades dividido em meses.

3. Exame de Qualificação do Mestrado (EQM).

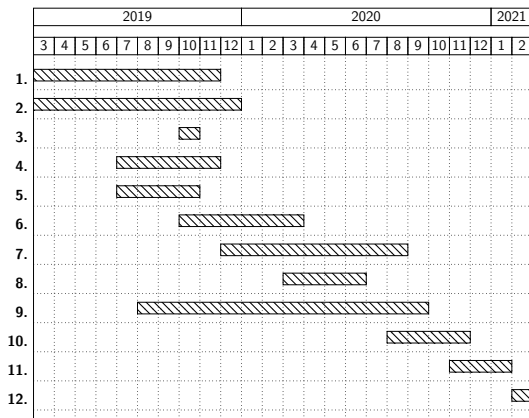


Tabela: Cronograma de atividades dividido em meses.

4. Pré-processamento das bases de dados.

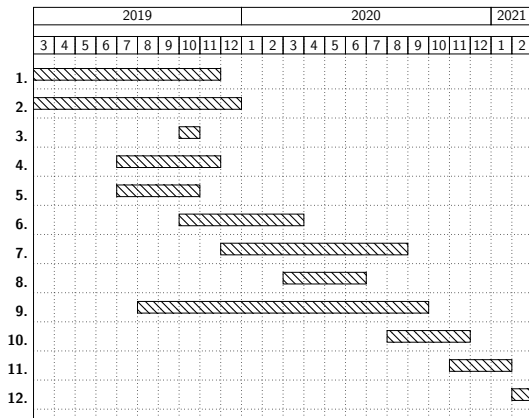


Tabela: Cronograma de atividades dividido em meses.

6. Definição de arquitetura para classificação de radiografias de tórax.

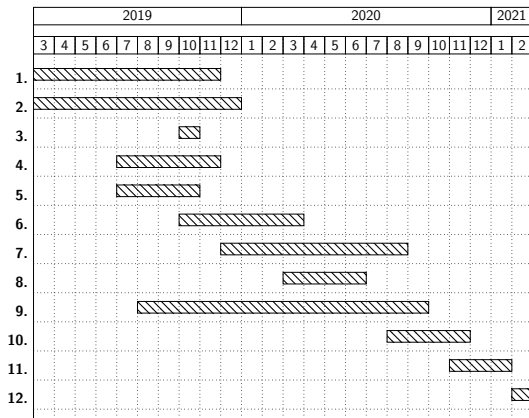


Tabela: Cronograma de atividades dividido em meses.

7. Aprimoramento da arquitetura.

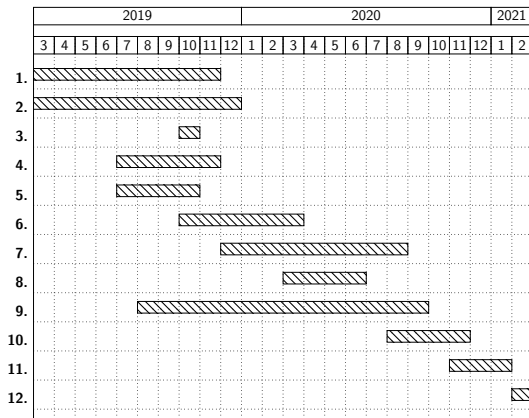


Tabela: Cronograma de atividades dividido em meses.

8. Participação do Programa de Estágio Docente (PED).

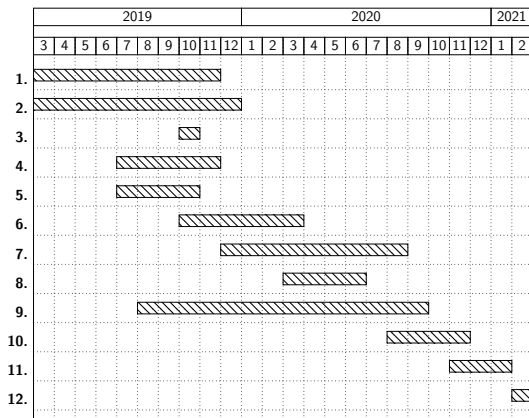


Tabela: Cronograma de atividades dividido em meses.

9. Realização de testes e análise dos resultados.

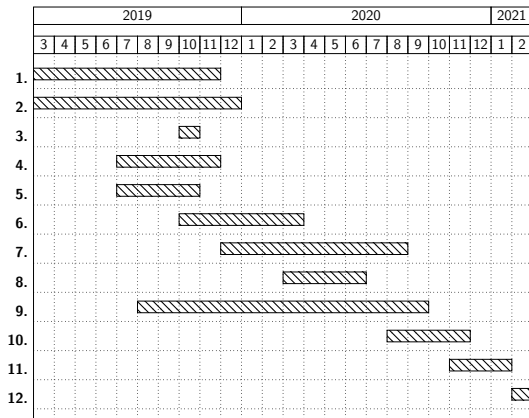


Tabela: Cronograma de atividades dividido em meses.

10. Documentação e publicação dos resultados.

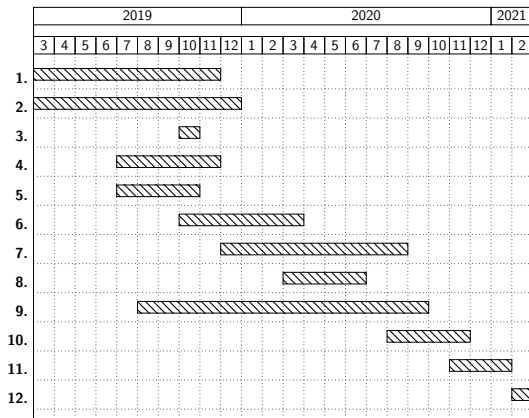


Tabela: Cronograma de atividades dividido em meses.

11. Escrita da dissertação de mestrado.

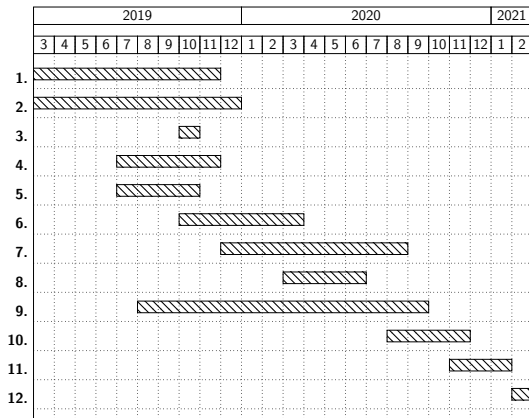


Tabela: Cronograma de atividades dividido em meses.

12. Apresentação da dissertação de mestrado.

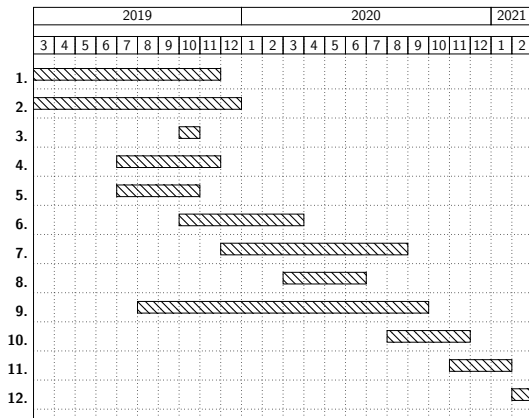


Tabela: Cronograma de atividades dividido em meses.

- [1] J. Irvin, P. Rajpurkar, M. Ko, Y. Yu, S. Ciurea-Illcus, C. Chute, H. Marklund, B. Haghgoo, R. Ball, K. Shpanskaya, J. Seekins, D. A. Mong, S. S. Halabi, J. K. Sandberg, R. Jones, D. B. Larson, C. P. Langlotz, B. N. Patel, M. P. Lungren, and A. Y. Ng.
CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison.
In *arXiv:1901.07031*, 2019.
- [2] A. E. W. Johnson, T. J. Pollard, S. Berkowitz, N. R. Greenbaum, M. P. Lungren, C. Deng, R. G. Mark, and S. Horng.
MIMIC-CXR: A Large Publicly Available Database of Labeled Chest Radiographs.
arXiv:1901.07042, 2019.

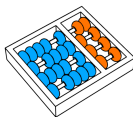
- [3] B. van Ginneken, L. Hogeweg, and M. Prokop.
Computer-aided Diagnosis in Chest Radiography: Beyond Nodules.
European Journal of Radiology, 72(2):226–230, 2009.
- [4] G. R. Tudor, D. Finlay, and N. Taub.
An assessment of inter-observer agreement and accuracy when reporting plain radiographs.
Clinical Radiology, 52(3):235–238, 1997.
- [5] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers.
ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases.
In *30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2097–2106, 2017.

- [6] L. Yao, E. Poblenz, D. Dagunts, B. Covington, D. Bernard, and K. Lyman.
Learning to diagnose from scratch by exploiting dependencies among labels.
arXiv:1710.10501, 2017.
- [7] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, R. L. Ball, C. Langlotz, K. Shpanskaya, M. P. Lungren, and A. Y. Ng.
CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning.
arXiv:1711.05225, 2017.

- [8] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers. ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases.
arXiv:1705.02315, 2017.

Classificação de Imagens de Radiografias de Tórax Utilizando Redes Neurais Profundas

Vinicius Teixeira de Melo



Orientador: Prof. Dr. Zanoni Dias
Coorientador: Prof. Dr. Hélio Pedrini

17 de Outubro de 2019