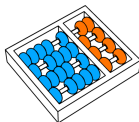


# Aumentação de Dados para Classificação de Imagens Utilizando Redes Neurais Profundas

Marianna de Pinho Severo



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

10 de Dezembro de 2020

# Agenda

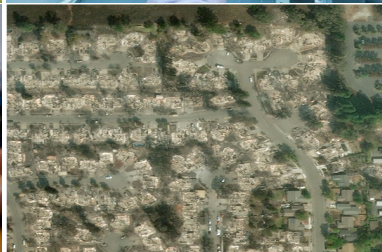
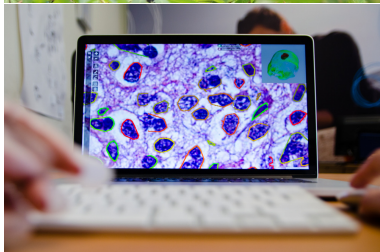
- 1 Introdução
- 2 Revisão Bibliográfica
- 3 Materiais
- 4 Métodos
- 5 Resultados Preliminares
- 6 Plano de Trabalho

# Agenda

- 1 Introdução
- 2 Revisão Bibliográfica
- 3 Materiais
- 4 Métodos
- 5 Resultados Preliminares
- 6 Plano de Trabalho

# Introdução

## Visão Computacional

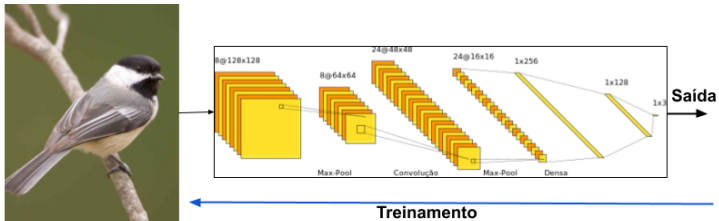
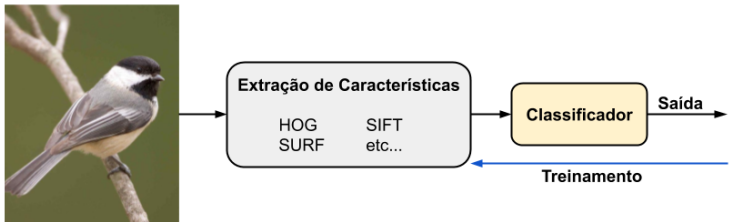


Fontes: <https://plantvillage.psu.edu/projects>, <https://cancercenter.ai/>, <https://xview2.org/dataset> e <https://www.factoryworx.com.au/predictive-analytics-through-artificial-intelligence/>.



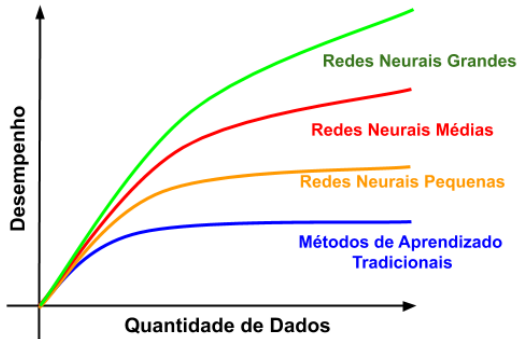
# Introdução

## Aprendizado de Máquina Profundo



Classificação sem e com Aprendizado Profundo.

- Aumento do poder computacional.
- Desenvolvimento de algoritmos mais sofisticados.
- Acesso a uma grande quantidade de dados.



Quantidade de dados e desempenho das técnicas.

# Introdução

## Caracterização do Problema

- Volume.
- Diversidade.
- Invariâncias.



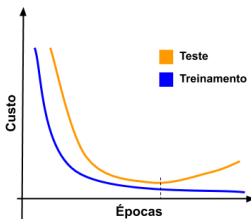
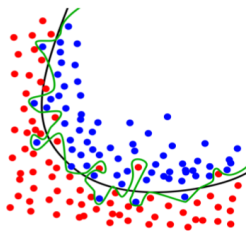
Exemplos da diversidade do mundo real.

Fontes: <https://bit.ly/3qhx3en>, <https://bit.ly/3osl8OV>, <https://bit.ly/3mP8ZTZ>, <https://bit.ly/2VKwBx3>,  
<https://bit.ly/39RplaX> e <https://bit.ly/36Mrida>.

# Introdução

## Caracterização do Problema

- Pequena quantidade de dados:
  - Raridade de certas classes.
  - Elevado custo de geração das imagens.
  - Necessidade de especialistas.
- *Overfitting*.
- Aumentação de dados.



Exemplo de curvas que indicam *overfitting*.

### Objetivo Geral

Investigar técnicas de aumento de dados do estado da arte, assim como explorar suas combinações, vantagens, desvantagens e capacidade de generalização para diferentes domínios de aplicação em tarefas de classificação de imagens.

### Objetivos Específicos

- Estudar as abordagens do estado da arte para classificação de imagens e para aumento de dados.
- Escolher e avaliar os métodos de aumento em bases de dados voltadas para diferentes domínios de aplicação.
- Investigar os benefícios, desvantagens e capacidade de generalização de combinações das técnicas escolhidas.

### Objetivos Específicos

- Propor uma nova solução de aumento de dados, baseada nos resultados das investigações.
- Comparar o método proposto com outras abordagens disponíveis.
- Documentar e publicar os resultados.

### Contribuições

Ajudar na tomada de decisões em trabalhos de classificação de imagens, determinar as abordagens de aumento que melhor se adequam aos casos de uso e outros possíveis cenários, e propor uma nova metodologia de aumento que contribua para o desenvolvimento da área.

### Questões de Pesquisa

- Conseguimos obter **resultados comparáveis aos de trabalhos recentes** de classificação de imagens, aplicando aumento de dados sobre conjuntos que possuem **poucas amostras de treinamento**?

### Questões de Pesquisa

- Conseguimos obter **resultados comparáveis aos de trabalhos recentes** de classificação de imagens, aplicando aumento de dados sobre conjuntos que possuem **poucas amostras de treinamento**?
- Podemos combinar **técnicas complementares** de aumento de dados, de maneira que as limitações individuais sejam amenizadas, conduzindo à construção de **melhores classificadores**?



### Questões de Pesquisa

- Conseguimos obter **resultados comparáveis aos de trabalhos recentes** de classificação de imagens, aplicando aumento de dados sobre conjuntos que possuem **poucas amostras de treinamento**?
- Podemos combinar **técnicas complementares** de aumento de dados, de maneira que as limitações individuais sejam amenizadas, conduzindo à construção de **melhores classificadores**?
- É possível a criação de uma **metodologia de aumento** de dados com grande capacidade de **generalização** entre bases de dados pertencentes a **diferentes domínios** de aplicação?

### Questões de Pesquisa

- Conseguimos obter **resultados comparáveis aos de trabalhos recentes** de classificação de imagens, aplicando aumento de dados sobre conjuntos que possuem **poucas amostras de treinamento**?
- Podemos combinar **técnicas complementares** de aumento de dados, de maneira que as limitações individuais sejam amenizadas, conduzindo à construção de **melhores classificadores**?
- É possível a criação de uma **metodologia de aumento** de dados com grande capacidade de **generalização** entre bases de dados pertencentes a **diferentes domínios** de aplicação?
- Como as técnicas de aumento de dados se **comparam** a **outras abordagens** que tratam o problema de **escassez de amostras**, tais como a regularização *dropout*, o pré-treinamento e novos paradigmas de aprendizado?

# Agenda

- 1 Introdução
- 2 Revisão Bibliográfica**
- 3 Materiais
- 4 Métodos
- 5 Resultados Preliminares
- 6 Plano de Trabalho

## Métodos de aumento de dados

- Métodos básicos de aumento de dados.
- Aumento baseado em aprendizado profundo:
  - Espaço de Características.
  - Redes Adversariais Generativas.
  - Transferência de Estilo Neural.
  - Meta-Aprendizado.

Trabalho	Principais contribuições
Perez et al. [11]	Combinação de métodos tradicionais, transformações elásticas e combinação de imagens. Aumentação do conjunto de treinamento e teste. Melhor abordagem: transformações geométricas e fotométricas.
DeVries et al. [5]	Incentivo ao melhor aprendizado do contexto da imagem.
Lopes et al. [9]	Maior robustez a ruídos e a outros tipos de corrupções. Maior capacidade de generalização.
Summers et al. [12]	União de métodos não lineares e lineares de combinação de imagens para aumento de dados.

Resumo de trabalhos de métodos básicos de aumento de dados.

Trabalho	Principais contribuições
DeVries et al. [4]	Diversidade resultante de operações simples sobre o espaço de características (ruído, interpolação e extrapolação). Independente de domínio. Complementar a técnicas específicas de domínio.
Wang et al. [14]	Criação de distribuições a partir das quais transformações significativas podem ser amostradas. Função de perda robusta e que otimiza o treinamento.

Resumo de trabalhos de aumento no espaço de características.

# Revisão Bibliográfica

Aumentação Baseada em Aprendizado Profundo - Redes Adversariais Generativas

Trabalho	Principais contribuições
Frid-Adar et al. [6]	Mostrou que os métodos tradicionais de aumento possuem limitações e que a adição de amostras sintéticas produzidas por DCGANs conduzem a melhorias ainda maiores dos modelos.
Zhang et al. [15]	Criação de uma função de perda ( $2k$ loss) para o treinamento de ACGANs, que geram amostras mais diversas.
Zhu et al. [18]	Amostras sintéticas produzidas por CycleGANs complementam o <i>manifold</i> , estabelecendo melhores fronteiras de decisão. A aumento de uma classe melhora a acurácia das restantes.

Resumo de trabalhos de aumento baseada em Redes Adversariais Generativas.

Trabalho	Principais contribuições
Zheng et al. [16]	Transferência de estilo é uma técnica significativa de aumento. Conduz a melhores resultados em combinação com técnicas tradicionais de aumento de dados.
Jackson et al. [8]	Transferência de estilo a partir de vetores de estilo aleatórios. Combate ao <i>overfitting</i> sobre textura e consequente deslocamento de domínio.

Resumo de trabalhos de aumento baseada em transferência de estilo neural.



# Revisão Bibliográfica

Aumentação Baseada em Aprendizado Profundo - Meta-Aprendizado

Trabalho	Principais contribuições
Cubuk et al. [3]	Automatização da escolha das técnicas de aumento de dados. Transferência de técnicas entre bases de dados diferentes é efetiva. Cada base de dados possui um conjunto de aumentações que são mais relevantes.
Geng et al. [7]	Atualização do trabalho de Cubuk et al. [3], com a adoção de um espaço de buscas contínuo. Melhoria da busca de técnicas e escolha de melhores políticas de aumento.
Mihn et al. [10]	Processo iterativo e mais facilmente explicável. Automatização do pré-processamento dos dados. Sequência de aumentações para cada amostra.

Resumo de trabalhos de aumento baseada em meta-aprendizado.

# Agenda

- 1 Introdução
- 2 Revisão Bibliográfica
- 3 Materiais**
- 4 Métodos
- 5 Resultados Preliminares
- 6 Plano de Trabalho

### Bases

- **Digipathos Embrapa** [1]: identificação de espécies doentes em plantações.
- **Places365-Standard** [17]: reconhecimento de cenas.
- **VGGFace2** [2]: reconhecimento de faces (identificação e verificação).
- **COVIDx** [13]: identificação de pacientes contaminados com o novo coronavírus SARS-CoV-2.

Quantidade de amostras por classe

Base de dados	Subconjunto	Número de classes	Total de amostras	Mínimo	Média	Máximo
Digipathos Embrapa	PDDDB	171	2.337	1	13	88
	XDB	93	46.101	1	495	3.791
Places365 - Standard	Treinamento	365	1.803.460	3.068	4.940	5.000
	Validação		36.500	100	100	100
	Teste		328.500	900	900	900
VGGFace2	Treinamento	8.631	3.141.890	87	364	843
	Teste	500	169.396	98	338	761

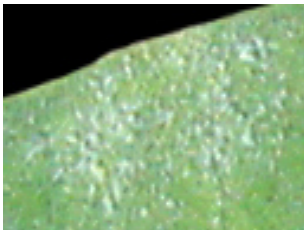
Estatísticas sobre as amostras para as bases de dados.

Base de dados	Subconjunto	Classe	Total de amostras
COVIDx	Treinamento	COVID-19	505
		Pneumonia	5.464
		Normal	7.966
	Teste	COVID-19	100
		Pneumonia	594
		Normal	885

Estatísticas sobre a distribuição de amostras da base de dados COVIDx.

# Materiais

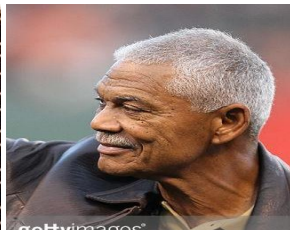
## Bases de Dados



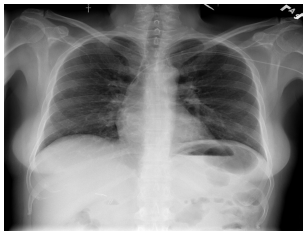
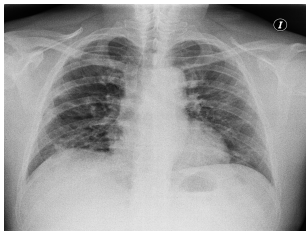
Exemplos de amostras da classe Mandioca, da base de dados Digipathos Embrapa.



Exemplos de amostras da classe Cafeteria, da base de dados Places365 - Standard.



Exemplos de amostras da classe Felipe Alou, da base de dados VGGFace2.



Exemplos de COVID-19, Normal e Pneumonia, respectivamente, da base COVIDx.

## Laboratório de Informática Visual (LIV)

- Intel i7-3770 3.50 GHz.
- 32 GB de memória RAM.
- GPU NVIDIA TITAN V.

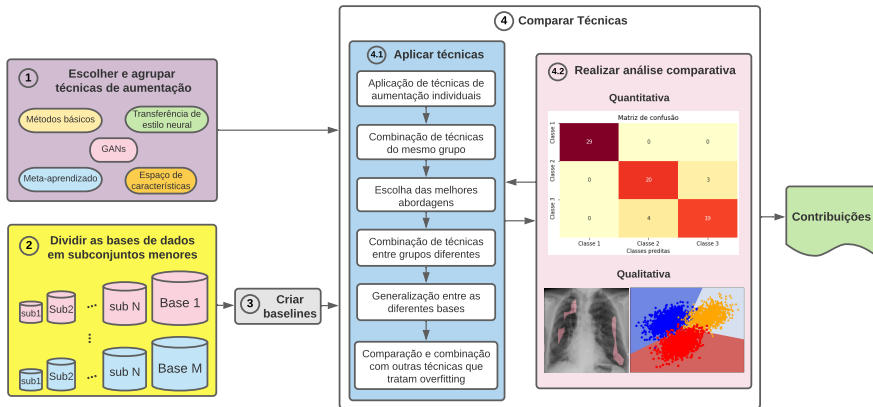
## Google Colab

- CPU, GPU, TPU, memória e armazenamento.

# Agenda

- 1 Introdução
- 2 Revisão Bibliográfica
- 3 Materiais
- 4 Métodos**
- 5 Resultados Preliminares
- 6 Plano de Trabalho





Principais etapas que compõem a metodologia deste projeto.

### Mais detalhes

- Validação cruzada *k-fold* e *holdout*.
- Modelos criados a partir do conjunto de treinamento e da junção deste com o conjunto de validação.
- Separação dos dados de acordo com o reportado em seus artigos.
- Modelos iniciais: SE-ResNet-50, VGG16, ResNet-152, GoogLeNet, arquiteturas Inception mais recentes, ResNet-50 e COVID-Net.
- Otimização de hiperparâmetros dos classificadores e dos métodos de aumentação baseados em aprendizado profundo.
- Técnicas de parada antecipada.

### Métricas

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

$$\text{Precisão} = \frac{TP}{TP+FP}$$

$$\text{Acurácia} = \frac{TP+TN}{TP+TN+FP+FN}$$

$$\text{Acurácia Balanceada} = \frac{1}{N} \sum_{i=1}^N \text{Sensibilidade}(\text{classe}[i])$$

### Métricas

**Identificação:** TPIR, FNIR, FPIR e *Rank* de Identificação.

**Verificação:** FAR e TAR.

**Curvas:** ROC (FAR×TAR), DET (FPIR×FNIR) e CMC (*Rank*×TPIR).

# Agenda

- 1 Introdução
- 2 Revisão Bibliográfica
- 3 Materiais
- 4 Métodos
- 5 Resultados Preliminares**
- 6 Plano de Trabalho

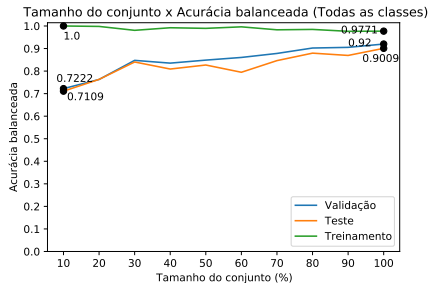
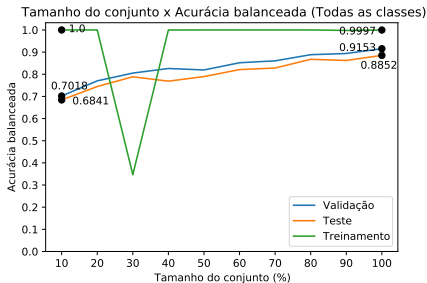
## Configurações de Treinamento

- Base COVIDx [13].
- Classificador ResNet-50.
- Otimizador SGD.
- Taxa de aprendizagem 0,001.
- *Momentum* Nesterov 0,9.
- Entropia Cruzada Categórica.
- 100 épocas e *early stopping* com paciência 10.
- Métricas Precisão, Sensibilidade e Acurácia balanceada.

## Configurações de Treinamento

- Experimentos com e sem aumentoção.
- Rotação, Translação, Brilho, Escala e Espelhamento horizontal.
- Cenários COVID, Balanceado e Todas as Classes.
- 10 subconjuntos.

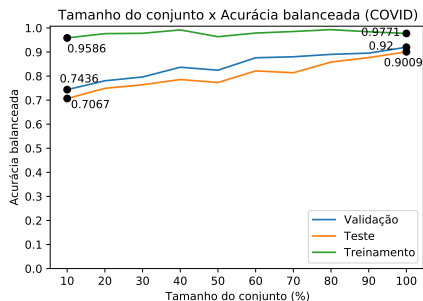
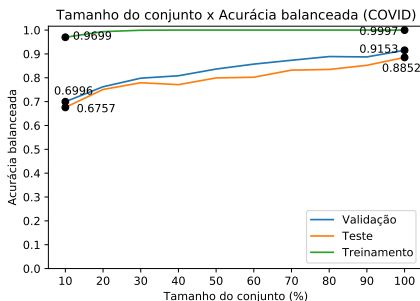
# Resultados Preliminares



Acurácia balanceada para experimentos sem (esquerda) e com aumento de dados (direita), cenário Todas as Classes.

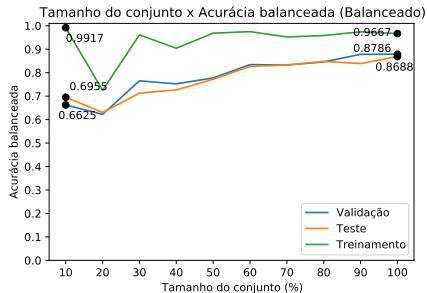
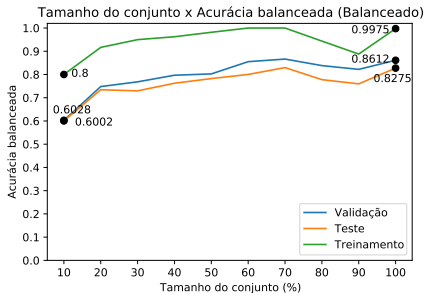


# Resultados Preliminares



Acurácia balanceada para experimentos sem (esquerda) e com aumento de dados (direita), cenário COVID.

# Resultados Preliminares



Acurácia balanceada para experimentos sem (esquerda) e com aumento de dados (direita), cenário Balanceado.

# Resultados Preliminares

Precisão									
Tamanho (%)	Todas as Classes			COVID			Balanceado		
	Normal	Pneumonia	COVID-19	Normal	Pneumonia	COVID-19	Normal	Pneumonia	COVID-19
10	0,9144	0,8171	0,9565	0,9108	0,8476	0,9444	0,8856	0,6684	0,8235
20	0,9311	0,8195	0,9302	0,9313	0,8616	<b>0,9500</b>	0,8730	0,7798	0,5091
30	0,9361	0,8657	0,9259	0,9301	0,8860	0,9200	0,8851	0,8160	0,7333
40	0,9233	0,8569	0,9200	0,9344	0,8673	0,9362	0,9051	0,8285	0,6842
50	0,9320	0,8688	0,8947	0,9405	0,8736	0,9455	0,8924	0,8152	0,7722
60	0,9428	0,8838	<b>0,9667</b>	0,9323	0,8790	0,9310	0,9295	0,8384	0,6854
70	0,9376	0,8883	0,9531	0,9384	0,8864	0,9000	0,9174	<b>0,8552</b>	0,6460
80	0,9420	0,9010	0,9481	0,9478	0,8778	0,9265	0,9106	0,8347	0,8182
90	<b>0,9506</b>	0,8792	0,9114	<b>0,9502</b>	0,8926	0,9437	0,9050	0,8331	0,8033
100	0,9422	<b>0,9191</b>	0,9070	0,9422	<b>0,9191</b>	0,9070	<b>0,9357</b>	0,8383	<b>0,8375</b>

Precisão para os experimentos sem aumento de dados.

Sensibilidade									
Tamanho (%)	Todas as Classes			COVID			Balanceado		
	Normal	Pneumonia	COVID-19	Normal	Pneumonia	COVID-19	Normal	Pneumonia	COVID-19
10	0,9299	0,9024	0,2200	0,9582	0,8990	0,1700	0,7785	0,8822	0,1400
20	0,9164	0,9175	0,4000	0,9503	0,9226	0,3800	0,8621	0,7811	0,5600
30	0,9435	0,9226	0,5000	<b>0,9616</b>	0,9158	0,4600	0,9051	0,8434	0,4400
40	0,9390	0,9074	0,4600	0,9492	0,9242	0,4400	0,9051	0,8620	0,5200
50	0,9446	0,9141	0,5100	0,9469	0,9310	0,5200	0,8904	0,8468	0,6100
60	0,9492	<b>0,9343</b>	0,5800	0,9492	0,9175	0,5400	0,9085	0,8822	0,6100
70	0,9503	0,9242	0,6100	0,9469	0,9192	0,6300	0,9040	0,8552	<b>0,7300</b>
80	0,9537	0,9192	0,7300	0,9435	0,9310	0,6300	0,9096	0,8838	0,5400
90	0,9356	0,9310	0,7200	0,9492	<b>0,9377</b>	0,6700	<b>0,9153</b>	0,8737	0,4900
100	<b>0,9582</b>	0,9175	<b>0,7800</b>	0,9582	0,9175	<b>0,7800</b>	0,9051	<b>0,9074</b>	0,6700

Sensibilidade para os experimentos sem aumento de dados.

# Resultados Preliminares

Matriz de confusão para 10% dos dados (Todas as Classes)

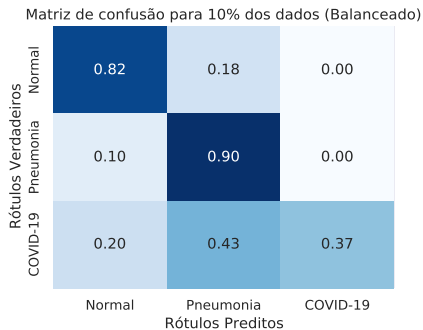
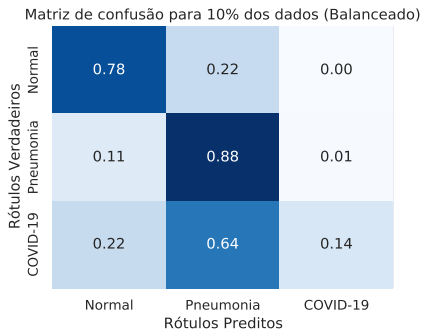
Rótulos Verdadeiros	Normal	0.93	0.07	0.00
	Pneumonia	0.10	0.90	0.00
	COVID-19	0.20	0.58	0.22
	Rótulos Preditos	Normal	Pneumonia	COVID-19

Matriz de confusão para 10% dos dados (Todas as Classes)

Rótulos Verdadeiros	Normal	0.95	0.05	0.00
	Pneumonia	0.09	0.91	0.00
	COVID-19	0.19	0.54	0.27
	Rótulos Preditos	Normal	Pneumonia	COVID-19

Matrizes de confusão para experimentos sem (esquerda) e com aumento de dados (direita), cenário Todas as Classes.

# Resultados Preliminares



Matrizes de confusão para experimentos sem (esquerda) e com aumento de dados (direita), cenário Balanceado.

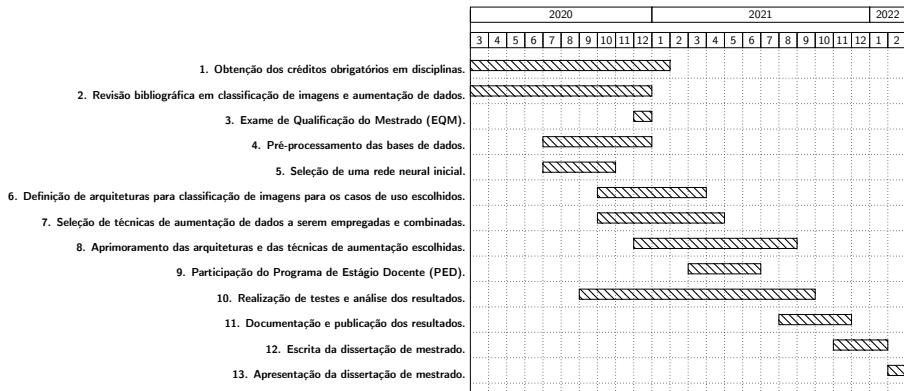
## Conclusões

- O desempenho dos classificadores melhora ao incrementarmos o tamanho do conjunto de dados de treinamento.
- Com a adoção de técnicas de aumento de dados tradicionais, conseguimos reduzir o impacto do *overfitting*, melhorando a acurácia de teste, inclusive sobre conjuntos com poucas amostras.
- Além da acurácia, a aumento de dados também melhorou a precisão e a sensibilidade dos classificadores.
- Mesmo ao empregarmos amostras aumentadas, a sensibilidade da categoria COVID-19 permaneceu extremamente baixa.

# Agenda

- 1 Introdução
- 2 Revisão Bibliográfica
- 3 Materiais
- 4 Métodos
- 5 Resultados Preliminares
- 6 Plano de Trabalho**

# Plano de Trabalho



Cronograma de atividades dividido em meses.



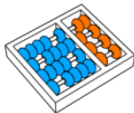
- [1] J. G. A. Barbedo.  
Plant disease identification from individual lesions and spots using deep learning.  
*Biosystems Engineering*, 180:96–107, 2019.
- [2] Q. Cao, L. Shen, W. Xie, O. M. Parkhi, and A. Zisserman.  
VGGFace2: A dataset for recognising faces across pose and age.  
In *IEEE International Conference on Automatic Face and Gesture Recognition (FG)*, pages 67–74, 2018.
- [3] E. D. Cubuk, B. Zoph, D. Mane, V. Vasudevan, and Q. V. Le.  
AutoAugment: Learning Augmentation Strategies from Data.  
In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 113–123, 2019.
- [4] T. DeVries and G. W. Taylor.  
Dataset Augmentation in Feature Space.  
*arXiv:1702.05538*, 2017.

- [5] T. DeVries and G. W. Taylor.  
Improved Regularization of Convolutional Neural Networks with Cutout.  
*arXiv:1708.04552*, 2017.
- [6] M. Frid-Adar, I. Diamant, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan.  
GAN-based Synthetic Medical Image Augmentation for increased CNN Performance in Liver Lesion Classification.  
*Neurocomputing*, 321:321–331, 2018.
- [7] M. Geng, K. Xu, B. Ding, H. Wang, and L. Zhang.  
Learning data augmentation policies using augmented random search.  
*arXiv:1811.04768*, 2018.
- [8] P. T. Jackson, A. A. Abarghouei, S. Bonner, T. P. Breckon, and B. Obara.  
Style Augmentation: Data Augmentation via Style Randomization.  
In *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPR Workshops)*, pages 83–92, 2019.

- [9] R. G. Lopes, D. Yin, B. Poole, J. Gilmer, and E. D. Cubuk.  
Improving Robustness Without Sacrificing Accuracy with Patch Gaussian Augmentation.  
*arXiv:1906.02611*, 2019.
- [10] T. N. Minh, M. Sinn, H. T. Lam, and M. Wistuba.  
Automated Image Data Preprocessing with Deep Reinforcement Learning.  
*arXiv:1806.05886*, 2018.
- [11] F. Perez, C. Vasconcelos, S. Avila, and E. Valle.  
Data Augmentation for Skin Lesion Analysis.  
In *OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis*, pages 303–311. Springer, 2018.
- [12] C. Summers and M. J. Dinneen.  
Improved Mixed-Example Data Augmentation.  
In *IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 1262–1270, 2019.

- [13] L. Wang and A. Wong.  
COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest X-Ray Images.  
*arXiv:2003.09871*, 2020.
- [14] Y. Wang, X. Pan, S. Song, H. Zhang, G. Huang, and C. Wu.  
Implicit Semantic Data Augmentation for Deep Networks.  
*In Advances in Neural Information Processing Systems (NIPS)*, pages 12635–12644, 2019.
- [15] X. Zhang, Z. Wang, D. Liu, and Q. Ling.  
DADA: Deep Adversarial Data Augmentation for Extremely Low Data Regime Classification.  
*In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2807–2811, 2019.

- [16] X. Zheng, T. Chalasani, K. Ghosal, S. Lutz, and A. Smolic.  
STaDA: Style Transfer as Data Augmentation.  
*International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP)*, 2019.
- [17] B. Zhou, A. Lapedriza, A. Khosla, A. Oliva, and A. Torralba.  
Places: A 10 million Image Database for Scene Recognition.  
*IEEE Transactions on Pattern Analysis and Machine Intelligence*,  
40(6):1452–1464, 2017.
- [18] X. Zhu, Y. Liu, J. Li, T. Wan, and Z. Qin.  
Emotion Classification with Data Augmentation Using Generative Adversarial Networks.  
*In Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD)*, pages 349–360, 2018.



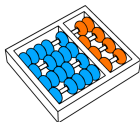
**LIV**

Laboratório de Informática Visual

**FITec**  
*Inovações Tecnológicas*

# Aumentação de Dados para Classificação de Imagens Utilizando Redes Neurais Profundas

Marianna de Pinho Severo



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

10 de Dezembro de 2020