

# INF-0619: Projeto Final

## Tema: Análise da Gravidade de Artrose no Joelho

### Equipe:

Natural Intelligence

### Membros:

Elcio Keniti Suzuki

Hamilton Gonçalves de Araújo

José Renato de Oliveira

Maurício Luiz Sobrinho

Tomais Gonçalves Peluso

### Descrição do Problema

A osteoartrose do joelho é uma doença de caráter inflamatório e degenerativo que provoca a destruição da cartilagem articular e leva a uma deformidade da articulação. A etiologia do processo degenerativo é complexa e inicia-se com o envelhecimento, porém, durante a vida do paciente podem ocorrer fatos que provoquem o início precoce desse processo degenerativo natural.

O diagnóstico é feito associando se os sintomas com as alterações presentes nos exames de imagem, especialmente na radiografia (raio-X). Para isso, usa-se a escala KELLGREN-LAWRENCE:

Grau	Alterações Radiográficas
<b>0</b>	Saudável
<b>I</b>	Questionável
<b>II</b>	Mínima
<b>III</b>	Moderada
<b>IV</b>	Severa

*Tabela 1: Escala de Kellgreen-Lawrence*

## Fluxo de Execução

O sistema segue o seguinte fluxo de execução, conforme ilustrado na Figura 5:

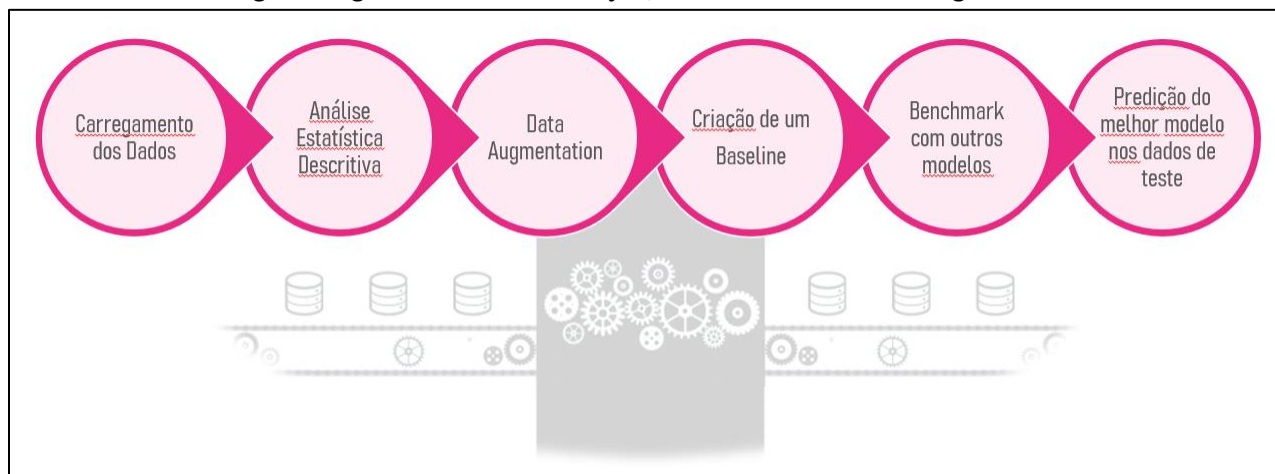


Figura 1: Fluxo de Execução

## Carregamento dos Dados e Análise Estatística Descritiva

A base de dados deste projeto consiste em cerca de 6 mil imagens de raio-X de joelhos de pacientes separadas em treinamento, validação e teste e também pelo grau de severidade na Osteoartrose de acordo com a escala KELLGREN-LAWRENCE.

### Distribuição das Classes

Os dados de treinamento, validação e teste estão altamente desbalanceados, conforme ilustrados na Figura 1.

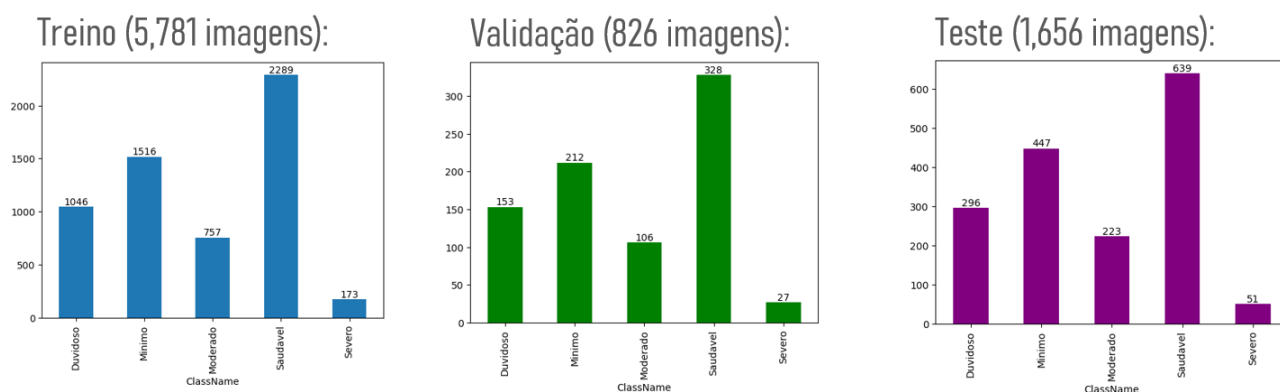


Figura 2: Distribuição das Classes nos Dados de Treinamento, Validação e Teste

Dados desbalanceados podem ser definidos pela pequena incidência de uma categoria dentro de um data set (classe minoritária) em comparação com as demais categorias (classes majoritárias) e podem acarretar problemas na construção de modelos e na geração de previsões.

## Reestruturação dos Dados

Uma forma de tirar o viés causado pela diferença de proporção das categorias consiste em manipular a quantidade de dados que são efetivamente utilizados pelo modelo de Machine Learning, tentando igualar o número de observações entre as classes.

A Figura 2 ilustra a técnica utilizada neste trabalho para balancear as classes nos dados de treinamento. Foi feito undersampling das classes majoritárias (saudável, duvidoso, mínimo, moderado) e oversampling da classe minoritária (severo) de modo que todas as classes fiquem com 600 observações.

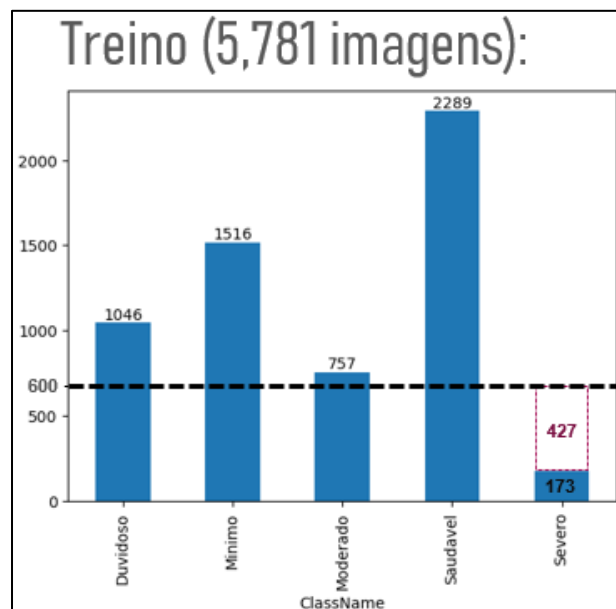


Figura 3: Undersampling e Oversampling dos Dados de Treinamento

## Data Augmentation

O oversampling foi feito gerando dados sintéticos com as seguintes técnicas de Data Augmentation (Figura 3):

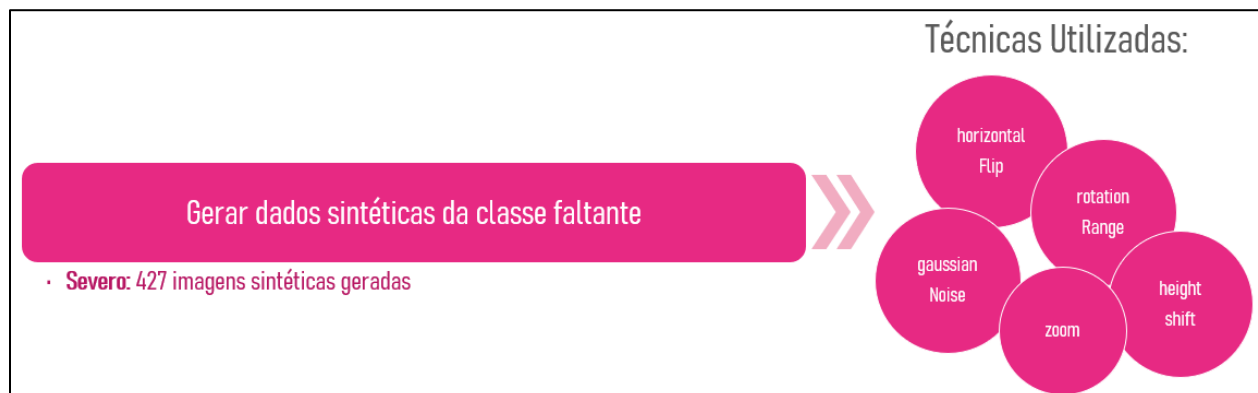


Figura 4: Data Augmentation

Amostras de dados sintéticos gerados:

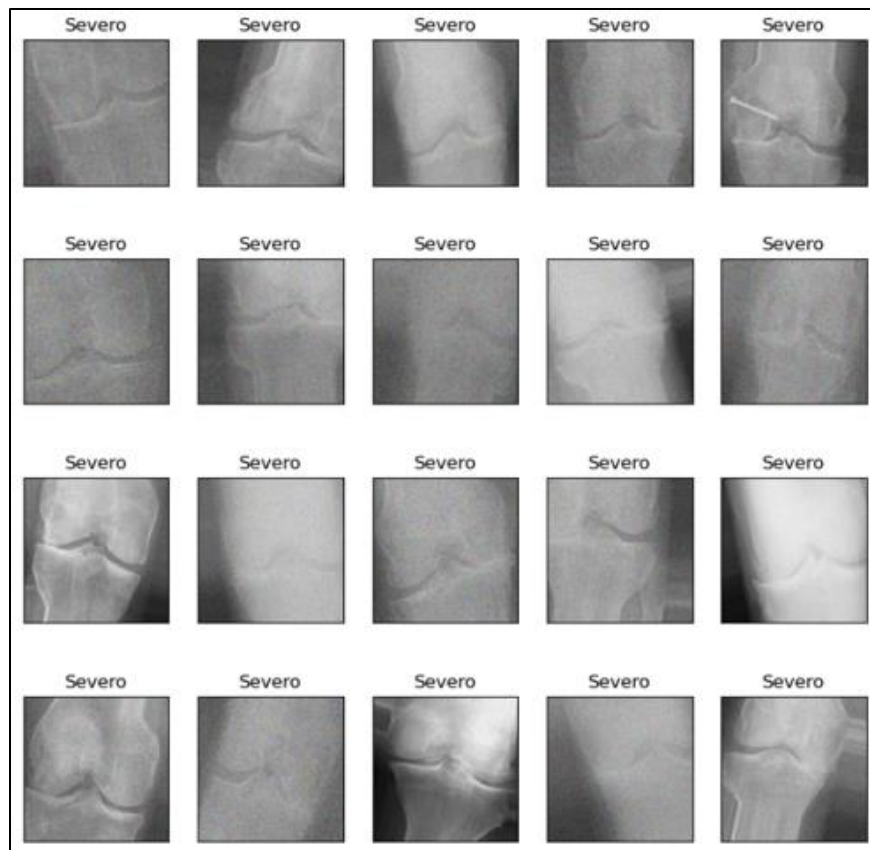


Figura 5: Dados Sintéticos Gerados

## Hiper-Parâmetros

Para cada experimento, foram utilizados os hiper-parâmetros presentes na Tabela 2:

Hiper-Parâmetro	Valor
Épocas	50
Taxa de Aprendizado	0.001
Tamanho do mini-lote	32

Tabela 2: Hiper-parâmetros

## Criação de Um Baseline

A baseline determinará o ponto de partida para o qual novos modelos serão incorporados.

## Rede Convolucional Simples

Neste projeto, a baseline escolhida foi uma rede convolucional simples.

Model: "sequential"		
Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 222, 222, 32)	896
max_pooling2d (MaxPooling2D)	(None, 111, 111, 32)	0
flatten (Flatten)	(None, 394272)	0
dense (Dense)	(None, 5)	1971365
Total params: 1,972,261		
Trainable params: 1,972,261		
Non-trainable params: 0		

Figura 6: Baseline (Camadas e Parâmetros)

## Resultado nos Dados de Validação

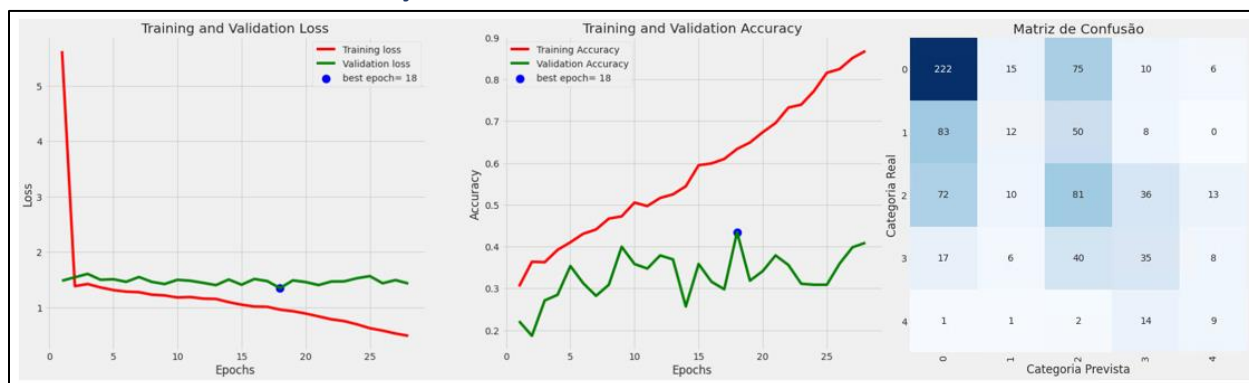


Figura 7: Resultados Baseline

	precision	recall	f1-score	support
0	0.56	0.68	0.61	328
1	0.27	0.08	0.12	153
2	0.33	0.38	0.35	212
3	0.34	0.33	0.33	106
4	0.25	0.33	0.29	27
accuracy			0.43	826
macro avg	0.35	0.36	0.34	826
weighted avg	0.41	0.43	0.41	826

Figura 8: Métricas Baseline

## Benchmark com Outros Modelos

### ResNet50 + Dense Layers + Fine Tuning

Resultado nos Dados de Validação:

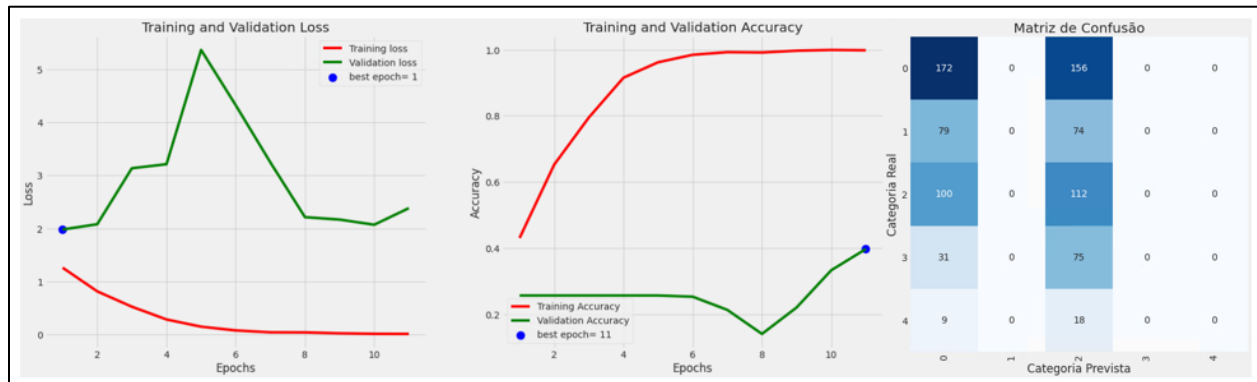


Figura 9: Resultado ResNet50 + Dense Layers + Fine Tuning

	precision	recall	f1-score	support
0	0.44	0.52	0.48	328
1	0.00	0.00	0.00	153
2	0.26	0.53	0.35	212
3	0.00	0.00	0.00	106
4	0.00	0.00	0.00	27
accuracy			0.34	826
macro avg	0.14	0.21	0.16	826
weighted avg	0.24	0.34	0.28	826

Figura 10: Métricas ResNet50 + Dense Layers + Fine Tuning

### VGG16 + Dense Layers

Resultado nos Dados de Validação:

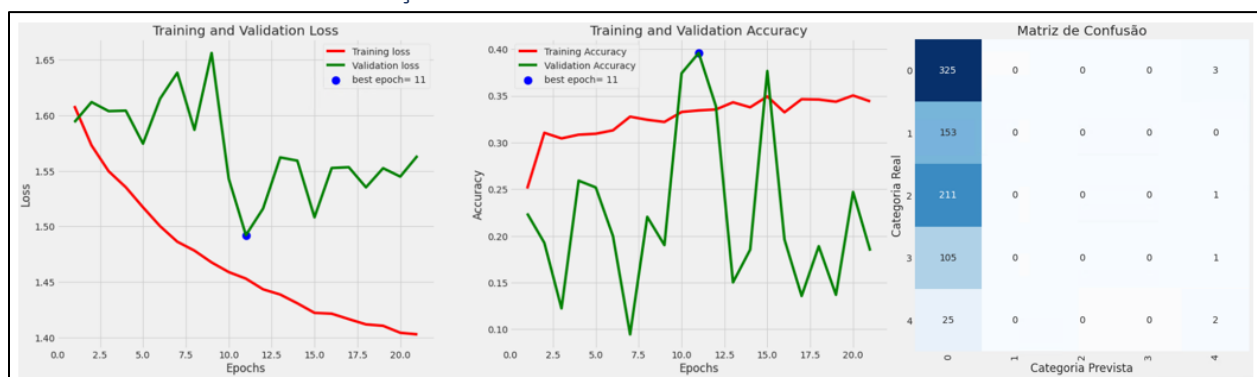


Figura 11: Resultado VGG16 + Dense Layers

	precision	recall	f1-score	support
0	0.40	0.99	0.57	328
1	0.00	0.00	0.00	153
2	0.00	0.00	0.00	212
3	0.00	0.00	0.00	106
4	0.29	0.07	0.12	27
accuracy			0.40	826
macro avg	0.14	0.21	0.14	826
weighted avg	0.17	0.40	0.23	826

Figura 12: Métricas VGG16 + Dense Layers

## VGG16 + Dense Layers + Fine Tuning

Resultado nos Dados de Validação:

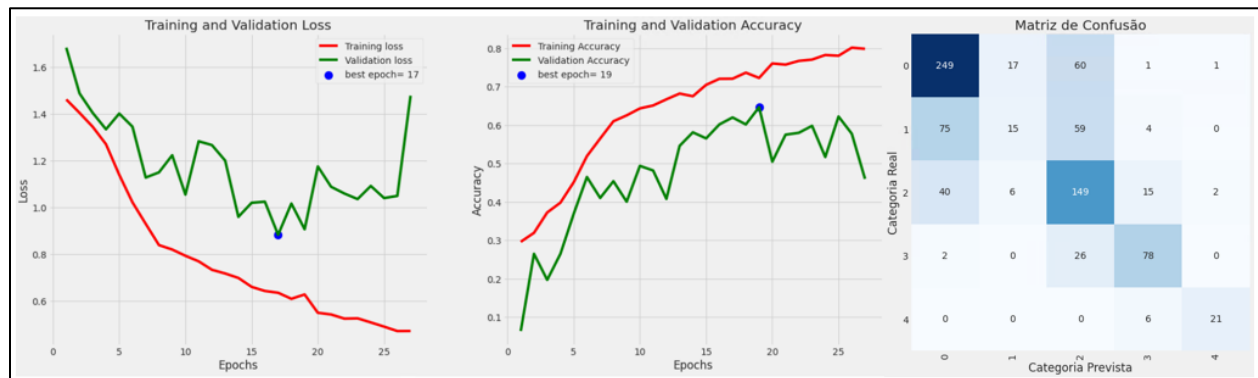


Figura 13: Resultado VGG16 + Dense Layers + Fine Tuning

	precision	recall	f1-score	support
0	0.68	0.76	0.72	328
1	0.39	0.10	0.16	153
2	0.51	0.70	0.59	212
3	0.75	0.74	0.74	106
4	0.88	0.78	0.82	27
accuracy			0.62	826
macro avg	0.64	0.61	0.61	826
weighted avg	0.60	0.62	0.59	826

Figura 14: Métricas VGG16 + Dense Layers + Fine Tuning

### Sumário dos Experimentos:

Modelo	Acurácia Balanceada nos Dados de Validação	Comentários
ResNet50 + SVM	12.71%	Resultado não expressivo
ResNet50 + Dense Layers	18.52%	Resultado não expressivo
ResNet50 + Dense Layers + Fine Tuning	34%	Top 3
VGG16 + SVM	20.21%	Resultado não expressivo
VGG16 + Dense Layers	40%	Top 2
VGG16 + Dense Layers + Fine Tuning	62%	Top 1

Tabela 3: Sumário dos Experimentos

Conforme Tabela 3, o experimento que obteve melhor resultado nos dados de validação foi o modelo “VGG16 + Dense Layers + Fine Tuning”.

### Resultado Predição do Melhor Modelo nos Dados de Teste

VGG16 + Dense Layers + Fine Tuning

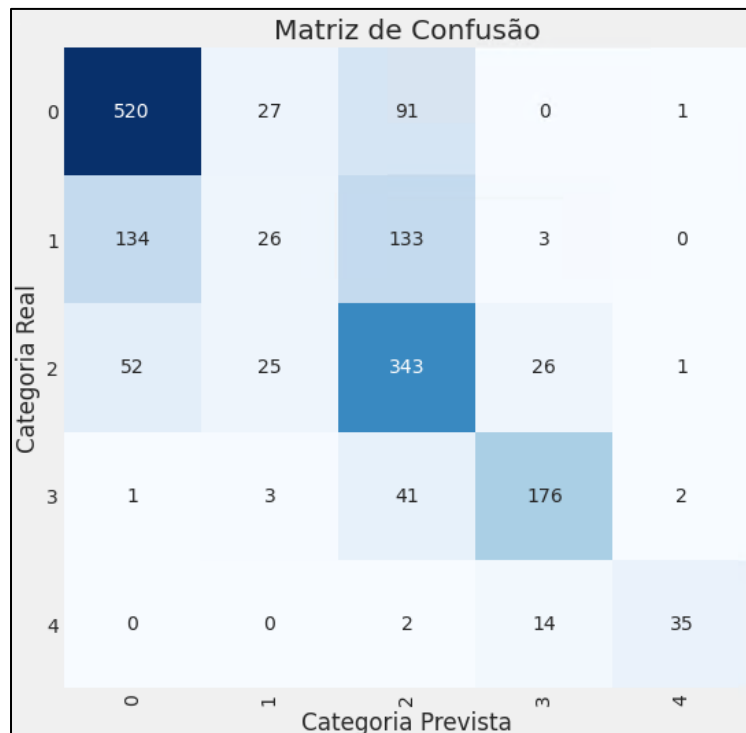


Figura 15: Tabela de Confusão nos Dados de Teste



	precision	recall	f1-score	support
0	0.74	0.81	0.77	639
1	0.32	0.09	0.14	296
2	0.56	0.77	0.65	447
3	0.80	0.79	0.80	223
4	0.90	0.69	0.78	51
accuracy			0.66	1656
macro avg	0.66	0.63	0.63	1656
weighted avg	0.63	0.66	0.63	1656

Figura 16: Métricas nos Dados de Teste

## Conclusão

Transferência de aprendizado (fine-tuning) com o modelo VGG16 foi o que obteve melhor resultado, atingindo bons resultados em toda escala de KELLGREN-LAWRENCE.

Tratando-se de dados médicos, é importante considerar a métrica de revocação (recall) e, mesmo na categoria SEVERA (4), onde foi utilizado mais de 50% de dados sintéticos para treinamento, o modelo obteve resultado satisfatório.

A categoria DUVIDOSO (1) que teve pior desempenho foi a que obteve pior desempenho, o que é justificável pelo próprio nome da categoria.

VGGNet é considerada a rede preferida pela comunidade para aprendizado por transferência pois sua arquitetura uniforme é boa para extrair características de imagens.

Isso se comprovou na competição Large Scale Visual Recognition Challenge (ILSVRC) de 2014 onde o este modelo conquistou o 1º e 2º lugar nos desafios de detecção e categorização de imagens deste ano.

## Referências Bibliográficas

CAMANHO, G. Tratamento da osteoartrose do joelho. **Revista Brasileira de Ortopedia**, v. 36, n. 5, 2001.

MARCIOR4. **Classificação de Kelgreen-Lawrence**. Disponível em:

<<https://traumatologiaeortopedia.com.br/informe/classificacao-de-kelgreen-lawrence/>>.

AZANK, F. **Dados Desbalanceados — O que são e como evitá-los**. Disponível em:

<<https://medium.com/turing-talks/dados-desbalanceados-o-que-s%C3%A3o-e-como-evit%C3%A1-los-43df4f49732b>>.

Disponível em: <<http://cursos.leg.ufpr.br/ML4all/apoio/Final.html>>. Acesso em: 15 dez. 2022.

**Deep Learning::Reconhecimento de Imagens.** Disponível em: <<https://lapix.ufsc.br/ensino/visao/visao-computacionaldeep-learning/deep-learningreconhecimento-de-imagens/>>. Acesso em: 15 dez. 2022.

**VGG-16 | CNN model.** Disponível em: <<https://www.geeksforgeeks.org/vgg-16-cnn-model/>>.

**Balance Data with Augment Images F1score=88%.** Disponível em: <<https://www.kaggle.com/code/gpiosenska/balance-data-with-augment-images-f1score-88>>. Acesso em: 15 dez. 2022.

**Knee Osteoarthritis Dataset with Severity Grading.** Disponível em: <<https://www.kaggle.com/datasets/shashwatwork/knee-osteoarthritis-dataset-with-severity>>. Acesso em: 15 dez. 2022.