

# Classificação de Leucemia

Bruno Mingoti  
Matrícula: 20204165

EEL7514 / EEL7513 – Tópico Avançado em  
Processamento de Sinais IV (Aprendizado de Máquina)

EEL / CTC / UFSC

# Introdução

- ▶ Existem 4 grupos em que a Leucemia pode ser classificada: *acute lymphoblastic leukemia* (ALL), *acute myeloid leukemia* (AML), *chronic lymphocytic leukemia* (CLL) e *chronic myeloid leukemia* (CML);
- ▶ Grande parte dos diagnósticos ocorre em estágios mais avançados da doença. Porém, quando detectado em fases iniciais, sua taxa de cura é de 80%, o que demonstra a importância do diagnóstico precoce e reafirma a utilidade da construção de um modelo de aprendizado de máquina capaz de realizar esse tipo de classificação;
- ▶ Dificuldade em diferenciar células malignas e saudáveis;
- ▶ **Objetivo:**
  - ▶ Elaborar modelo de aprendizado de máquina capaz de realizar uma classificação binária entre células saudáveis e células com Leucemia do tipo ALL.

# Trabalhos Relacionados

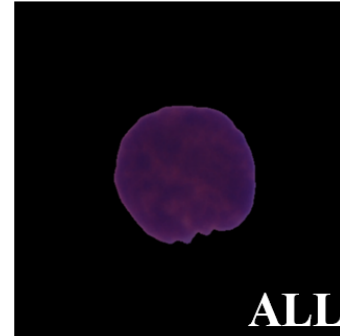
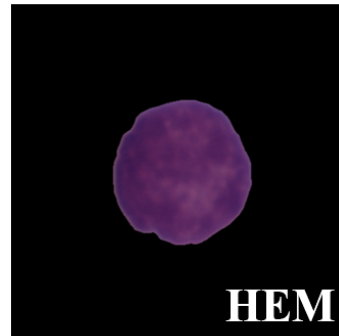
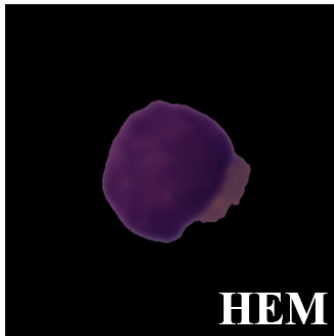
- ▶ Nayak realizou modificações na arquitetura VGG16 com batch normalization no dataset pré-treinado da imagenet. Resultados experimentais mostraram 80% de weighted F1-Score.
- ▶ Marzahl et al. propôs a utilização da ResNet18 com vários métodos de data augmentation e incorporou um mecanismo básico de atenção baseado em uma região da sub-rede proposta para impulsionar o resultado, obtendo um weighted F1-Score de 82,84% no dataset de teste final.
- ▶ Rahul et al. Propôs um método de segmentação de glóbulos brancos imaturos de imagens microscópicas e, utilizando um *stain deconvolutional layer* obteve uma acurácia de 93,2% na tarefa de diferenciar entre células saudáveis e malignas.
- ▶ Ding et al. utilizou uma técnica de ensemble denominada Stacking que combina múltiplos modelos classificadores por meio de um *meta-classifier*, obtendo um weighted F1-Score de 85,52%.
- ▶ Singhal et al. Utilizou um SVM para a detecção automática das células malignas baseada na geometria de suas features e no local binary pattern (LBP) de suas texturas.
- ▶ Mourya et al. Elaborou um modelo denominado LeukoNet, que combina features de domínio Discrete Cosine Transform (DCT) extraído via CNN com recursos de espaço de Optical Density (OD) para a construção de um modelo mais robusto.
- ▶ Kramer aplicou a arquitetura ResNeXt com módulos de Squeeze-and-Extraction, atingindo um weighted F1-Score de 88,91%.

# Trabalhos Relacionados

- ▶ Ramaneswaran et al. Explorou diversos Classifications Head, em vez do softmax comumente utilizado no problema em questão. Foram testadas várias técnicas de data augmentation e aplicado um híbrido entre a arquitetura InceptionV3 e o XGBoost como classification head. Obteve-se 98,6 de F1-Score.
- ▶ Xie et al. Desenvolveu uma arquitetura multi-stream e multi-feature com a adoção da Inception ResNetV2 como rede básica e empregou uma arquitetura híbrida que utiliza o conhecimento de três redes básicas: imagem redimensionada para (299, 299); imagens recortadas para apenas a região de interesse e redimensionadas para (299, 299); e imagem convertida de RGB para HSV e redimensionada para (299, 299). Por fim, as features são extraídas de cada rede básica, são concatenadas e, então, o resultado é usado como input de uma multi-layer perception (MLP) para obter o resultado final da classificação. Com esse método, obteve-se 83,35% de F1-Score.

# Dataset

- ▶ Conjunto de dados do ISBI C-NMC 2019, dataset preparado no Laboratório de Oncologia, AIIMS, New Delhi durante a competição para *Classification of Normal versus Malignant Cells in B-All White Blood Cancer Microscopic Images*, organizado pela CodaLab e disponibiliza um total de 10661 imagens de glóbulos brancos coletados de 60 pacientes diagnosticados com o câncer do tipo ALL e de 41 pessoas saudáveis;
- ▶ Classes:
  - ▶ ALL (*Acute Lymphoblastic Leukemia*)
  - ▶ HEM (*Healthy Cell*)



# Preparação dos Dados

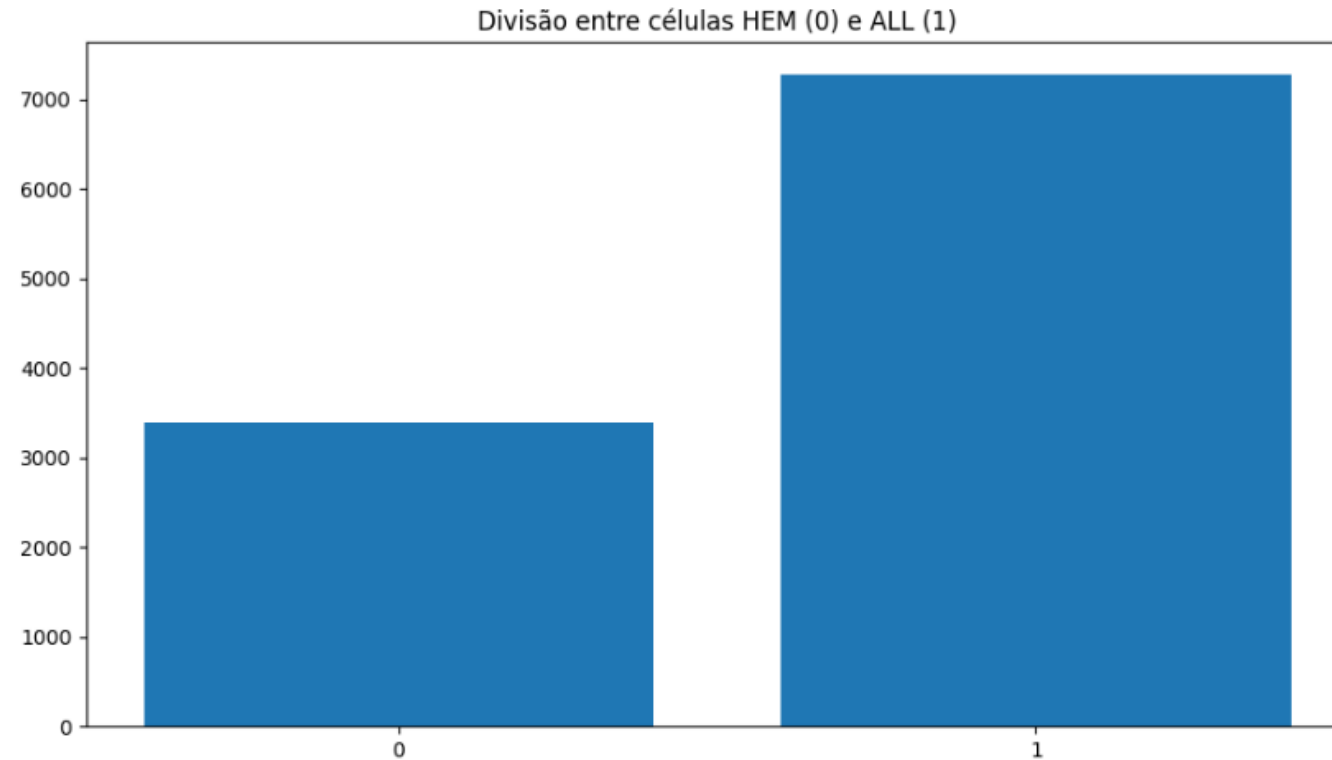
- ▶ Dados obtidos a partir de três folds – fold 0, fold 1 e fold 2, todos contendo as duas classes e com uma distribuição consistente entre os três folds. Para questões de organização, todos os dados serviram para alimentar o dataframe abaixo;
- ▶ ALL passa a se chamar 1;
- ▶ HEM passa a se chamar 0.

full_path	label
/kaggle/input/cnmc-leukemia/fold_0/fold_0/all/...	1
/kaggle/input/cnmc-leukemia/fold_0/fold_0/all/...	1
/kaggle/input/cnmc-leukemia/fold_0/fold_0/all/...	1
/kaggle/input/cnmc-leukemia/fold_0/fold_0/all/...	1
/kaggle/input/cnmc-leukemia/fold_0/fold_0/all/...	1
...	...
/kaggle/input/cnmc-leukemia/fold_2/fold_2/hem/...	0
/kaggle/input/cnmc-leukemia/fold_2/fold_2/hem/...	0
/kaggle/input/cnmc-leukemia/fold_2/fold_2/hem/...	0
/kaggle/input/cnmc-leukemia/fold_2/fold_2/hem/...	0
/kaggle/input/cnmc-leukemia/fold_2/fold_2/hem/...	0

# Análise Exploratória dos Dados (EDA)

- Problema encontrado: desbalanceamento entre as classes, com predominância da classe ALL.

```
1    7272  
0    3389  
Name: label, dtype: int64
```



# Separação dos conjuntos de dados

- ▶ Total de imagens: 10661;
- ▶ Tamanho original: (450, 450), composta por um glóbulo branco centralizado em background preto;
- ▶ Foi realizada uma divisão em 3 conjuntos de maneira estratificada com 75% das imagens para treino e os demais 25% divididos igualmente entre os conjuntos de validação e teste;
- ▶ Todas as execuções intermediárias para a escolha do melhor tipo de pré-processamento, aplicação do data augmentation, aplicação de diferentes pesos entre as classes da arquitetura foi utilizado um dataframe reduzido de 500 imagens.

Divisão dos Datasets	
df_train	7995
df_train_small	500
df_val	1333
df_test	1333



# Criação dos datasets e pré-processamento

- ▶ Biblioteca utilizada: Keras / Tensorflow
- ▶ CenterCrop para o tamanho de (450, 450);
- ▶ Resize para (299, 299) para questões de memória e redução no tempo de treinamento dos modelos intermediários;
- ▶ Utilização de algumas das arquiteturas do Keras Applications com os pesos da ImageNet e com as camadas descongeladas. Por fim, foi adicionada uma camada de ativação Softmax.

Hiperparâmetros	
epochs	10
learning rate	0.0001
img_size	(299, 299)
batch_size	16
Hyperparameter Tuning com Callback	ReduceLROnPlateau(factor=0.2, patience=3, verbose=1)

# Métricas

- ▶ A métrica de avaliação convencionalmente utilizada e adotada no desafio da CodaLab é o F1-Score;
- ▶ Em razão do uso por algumas literaturas do AUC, este foi calculado, mas apenas utilizado como métrica auxiliar.

$$Precision = \frac{VP}{VP + FP}$$

$$Recall = \frac{VP}{VP + FN}$$

$$F1\ Score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

# Escolha da Arquitetura

- ▶ As seguintes arquiteturas abaixo do Keras Applications foram escolhidas, com base em trabalhos relacionados.
- ▶ Todos os treinamentos foram performado em uma GPU T4 x2.

Arquitetura	Train F1 Score	Val F1 Score	Train AUC	Val AUC	Training Time
AlexNet	100.00%	80.07%	100.00%	85.13%	277 s
InceptionV3	100.00%	86.03%	100.00%	90.44%	224 s
InceptionResNetV2	100.00%	84.44%	100.00%	89.83%	479 s
VGG16	87.85%	76.79%	94.98%	85.25%	319 s
ResNet18	100.00%	79.17%	100.00%	83.23%	96 s

# Avaliação do class weight

- ▶ Foi realizado um experimento com a utilização do argumento `class_weight` no treinamento do melhor modelo obtido na etapa anterior;
- ▶ Performance com a aplicação do `class_weight` melhorou em 1,5%.

	Train F1 Score	Val F1 Score	Train AUC	Val AUC
Arquitetura InceptionV3				
Sem <code>class_weights</code>	100.00%	84.53%	100.00%	89.55%
Com <code>class_weights</code>	100.00%	86.03%	100.00%	90.44%

# Data Augmentation

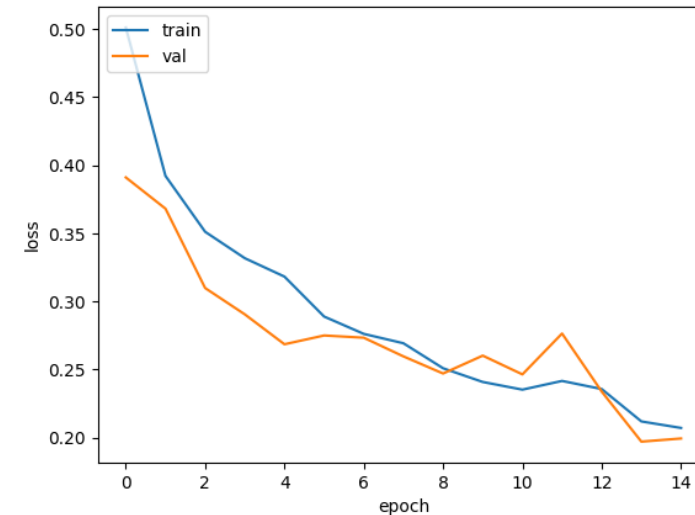
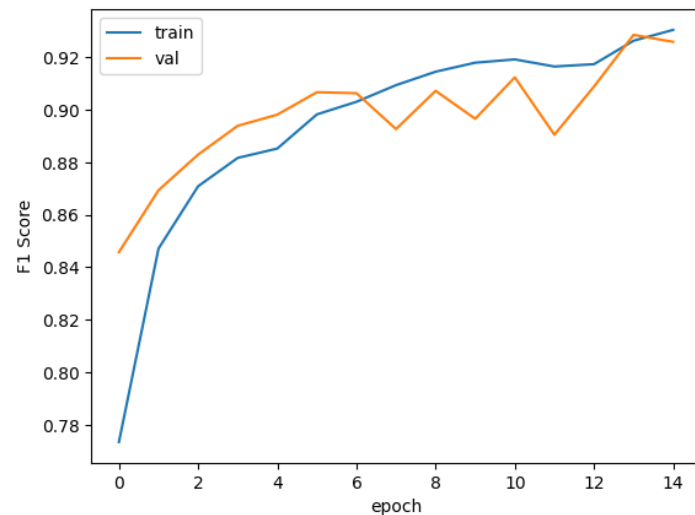
- ▶ Foi aplicado técnicas de Data Augmentation baseadas na literatura, trazendo uma melhoria de 1,75% em relação ao modelo da etapa anterior.
- ▶ Técnicas utilizadas:
  - ▶ *Flips* horizontais e verticais, cada um com 50% de probabilidade de ocorrência;
  - ▶ Rotações aleatórias com fator 0.3, ou seja, as rotações ocorrem no intervalo  $[-30\% * 2\pi, 30\% * 2\pi]$  (função *RandomRotation* do TensorFlow);
  - ▶ Translações aleatórias com fator de 20% na altura e largura da imagem (função *RandomTranslation* do TensorFlow);
  - ▶ Recortes retangulares aleatórios (função *random\_cutout* do TensorFlow Addons);

Modelo	Train F1-Score	Val F1-Score	Train AUC	Val AUC
Sem Data Augmentation	100%	86,03%	100%	90,44%
Com Data Augmentation	93,05%	87,78%	98,45%	93,29%

# Treinamento com o conjunto de treino completo

- ▶ Por limitações de memória, optou-se por realizar o treinamento com 7995 imagens, referente ao conjunto de treino completo;
- ▶ O modelo escolhido para prosseguir com o treinamento foi o InceptionV3 + class\_weight + data\_augmentation;
- ▶ Além disso, as imagens foram reduzidas para o tamanho de (200, 200) e batch size para 4 e o número de épocas foi aumentado para 15;
- ▶ Melhora em 5,01% com o aumento do conjunto de treino.

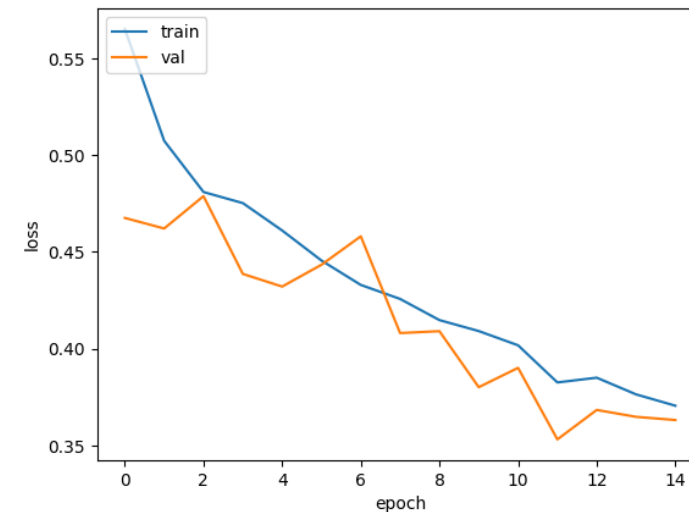
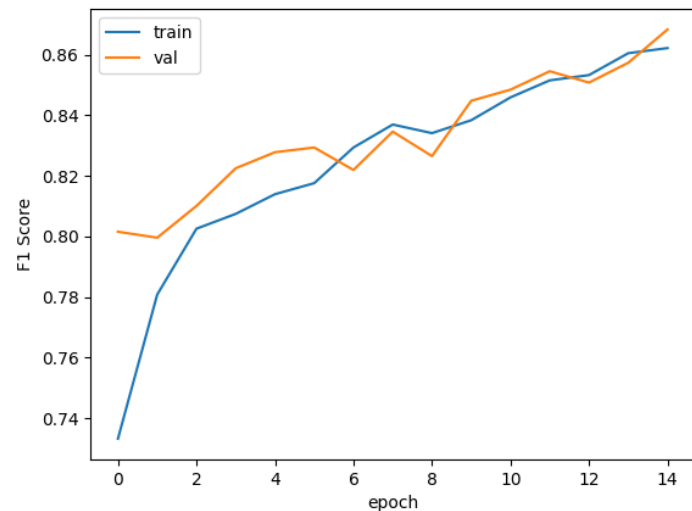
Train Weighted F1-Score	Val Weighted F1-Score	Train AUC	Val AUC
93,04%	92,59%	97,25%	97,27%



# Treinamento com o conjunto de treino completo em imagens HSV

- ▶ A mesma arquitetura foi aplicada para imagens convertidas de RGB para HSV para uma observação mais clara das características das células;
- ▶ Obteve um desempenho inferior ao modelo original com as imagens em RGB.

Train Weighted F1-Score	Val Weighted F1-Score	Train AUC	Val AUC
86,21%	86,03%	92,32%	91,83%



# Ensemble Averaging

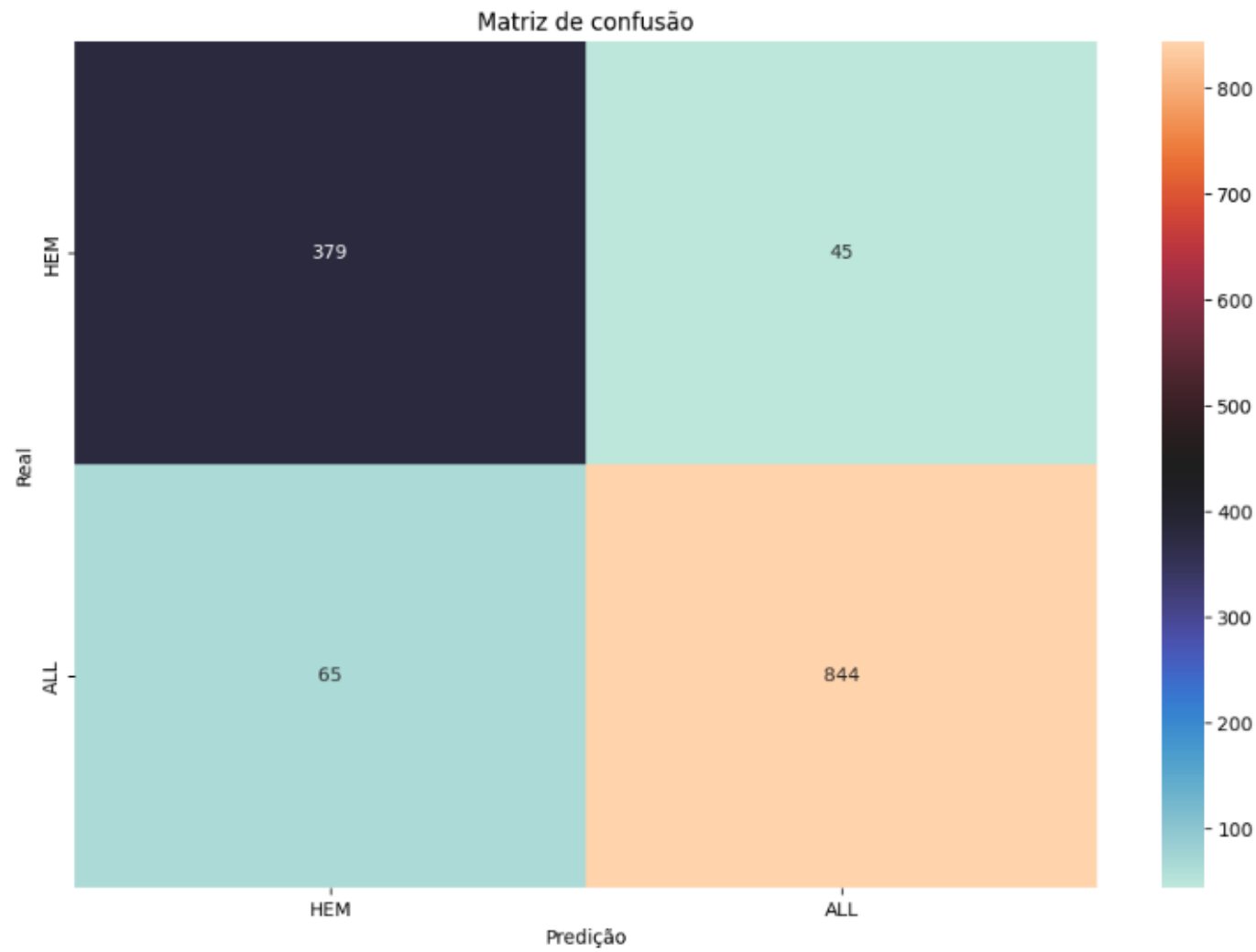
- ▶ Foi realizado um Ensemble Averaging entre os dois modelos propostos (o melhor modelo obtido até o momento com a arquitetura InceptionV3 + Data Augmentation + class\_weight, intitulado “Main Model” com o respectivo modelo com a conversão RGB para HSV, intitulado “HSV Model”);
- ▶ Melhora em 2,56%.

	Test F1 Score	Test AUC
Modelo		
Main Model	0.913151	0.910476
HSV Model	0.860339	0.848142
Ensemble Model	0.938821	0.911180



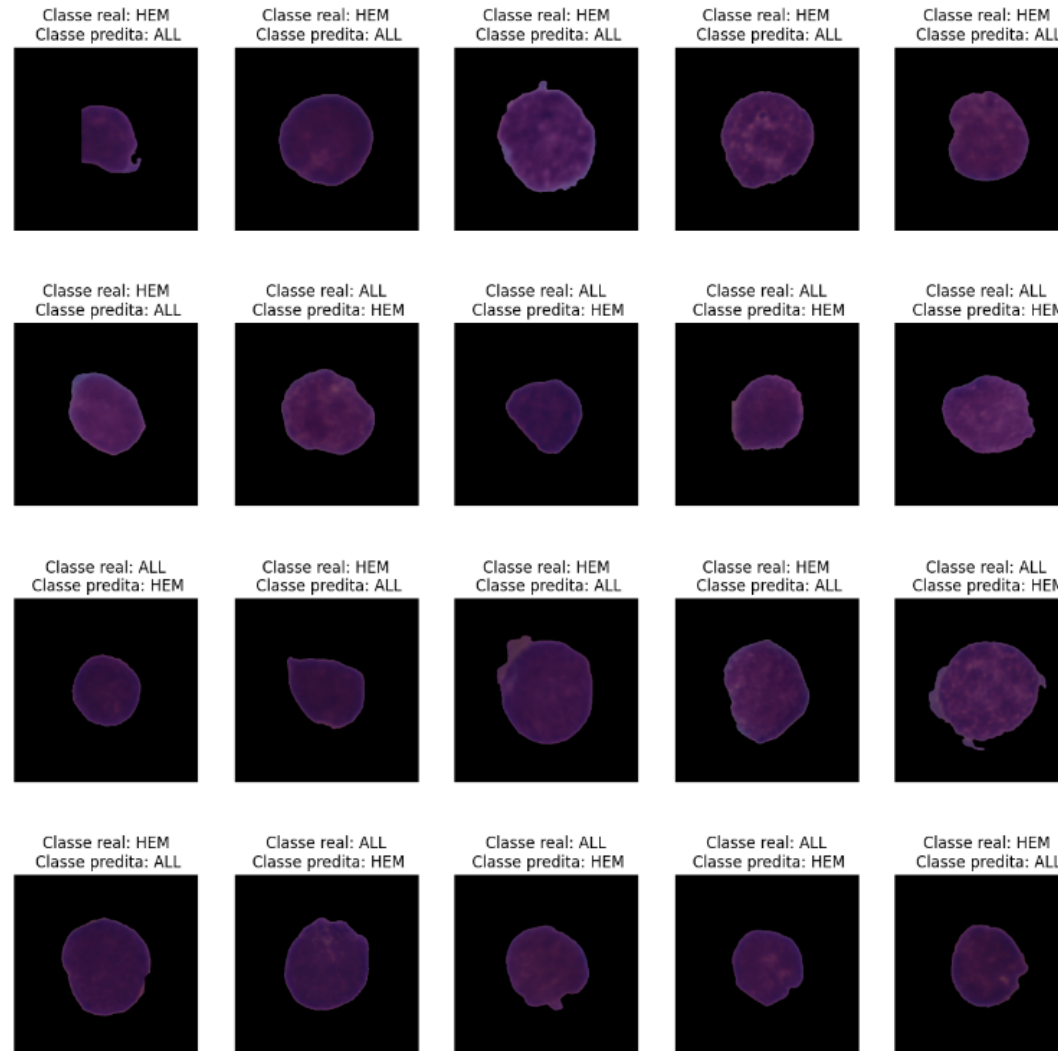
# Avaliação e Resultados

## Matriz de Confusão



# Avaliação e Resultados

## Imagens Erradas



# Avaliação e Resultados

## Comparação com outros modelos

Modelo	Weighted F1-Score
Nayak	80%
Marzahl et al.	82,84%
Rahul et al.	93,2%
Ding et al.	85,52%
Pan et al.	91,04%
Kramer	88,91%
Ramaneswaran	98,60%
<b>Método proposto</b>	<b>93,88%</b>

# Conclusão

Foi proposto um simples e efetivo modelo de Ensemble Neural Network para a classificação binária de glóbulos brancos capaz de diferenciar entre células cancerígenas (ALL) e células saudáveis (HEM) do dataset ISBI C-NMC 2019. Nesse agrupamento foi utilizada a mesma arquitetura InceptionV3 e construção do modelo, sendo que a única diferença aplicada foi a alteração do espaço de cores RGB para HSV.

Além disso, foram explorados diferentes métodos para lidar com o desbalanceamento das classes e variadas técnicas de data augmentation visando aprimorar o desempenho do modelo final, que apresentou um weighted F1-Score de 93,88%, um dos mais altos de acordo com a literatura e que ocuparia o 4º lugar no desafio da CodaLab.

# Conclusão

Results						
#	User	Entries	Date of Last Entry	Prediction score ▲	Duration ▲	Detailed Results
1	shubham14100	1	01/29/20	0.9526 (1)	0.00 (1)	<a href="#">View</a>
2	Kust119	171	06/21/22	0.9494 (2)	0.00 (1)	<a href="#">View</a>
3	shiv_sbilab	1	03/03/20	0.9486 (3)	0.00 (1)	<a href="#">View</a>
4	yspan	1	11/27/19	0.9357 (4)	0.00 (1)	<a href="#">View</a>
5	Able985	447	05/18/22	0.9311 (5)	0.00 (1)	<a href="#">View</a>
6	liuyunfei	13	09/14/20	0.9234 (6)	0.00 (1)	<a href="#">View</a>
7	harsh19466	13	03/03/22	0.9220 (7)	0.00 (1)	<a href="#">View</a>
8	lyhthienan	420	11/26/20	0.9194 (8)	0.00 (1)	<a href="#">View</a>
9	thling	21	02/01/20	0.9173 (9)	0.00 (1)	<a href="#">View</a>
10	Matteo_R	7	07/02/21	0.8965 (10)	0.00 (1)	<a href="#">View</a>

# Referências

- [1] S. Kassani, K. Schneider, P. Kassani, M. Wesolowski, R. Deters, “A Hybrid Deep Learning Architecture for Leukemic B-lymphoblast Classification”. Disponível em: <https://arxiv.org/pdf/1909.11866>
  
- [2] R. Gupta, S. Gehlot, A. Gupta, “C-NMC: B-lineage acute lymphoblastic leukaemia: A blood câncer dataset”. Disponível em: ISBI 2019 C-NMC Challenge: Classification in Cancer Cell Imaging.
  
- [3] J. Prellberg, O. Kramer, “Acute Lymphoblastic Leukemia Classification from Microscopic Images using Convolutional Neural Networks”. Disponível em: <https://arxiv.org/abs/1906.09020>
  
- [4] Y. Liu, F. Long, “Acute lymphoblastic leukemia cells image analysis with deep bagging ensemble learning”. Disponível em: <https://www.biorxiv.org/content/10.1101/580852.full.pdf>
  
- [5] S. Ramaneswaran, K. Srinivasan, P.M. Vincent, C. Chang, “Hybrid Inception v3 XGBoost Model for Acute Lymphoblastic Leukemia Classification”. Disponível em: <https://www.hindawi.com/journals/cmmm/2021/2577375/>

# Referências

- [6] Y. Ding, Y. Cui, “Deep Learning for Classifying os White Blood Cancer”. Disponível em: ISBI 2019 C-NMC Challenge: Classification in Cancer Cell Imaging.
- [7] A. Honnalgere, G. Nayak, “Classification of Normal Versus Malignant Cells in B-ALL White Blood Cancer Microscopic Images”. Disponível em: ISBI 2019 C-NMC Challenge: Classification in Cancer Cell Imaging.
- [8] C. Marzahl, M.Aubreville, J. Maier, “Classification of Leukemic B-Lymphoblast Cells from Blood Smear Microscopic Images with na Attention-Based Deep Learning Method and Advanced Augmentation Techniques”. Disponível em: ISBI 2019 C-NMC Challenge: Classification in Cancer Cell Imaging.
- [9] R. Kulhalli, C. Savadikar, B. Garware, “Toward Automated Classification of B-Acute Lymphoblastic Leukemia”. Disponível em: ISBI 2019 C-NMC Challenge: Classification in Cancer Cell Imaging.
- [10] Y. Pan, M. Liu, X. Shen, D. Shen, “Neighborhood-Correction Algorithm for Classification of Normal and Malignant Cells”. Disponível em: [https://drive.google.com/file/d/1\\_KbpWDWNHbEZSoJuQ026huCkEgtxQxm0/view](https://drive.google.com/file/d/1_KbpWDWNHbEZSoJuQ026huCkEgtxQxm0/view)

# Referências

[11] X. Xie, Y. Li, M. Zhang, Y. Wu, L. Shen, “Multi-streams and Multi-features for Cell Classification”. Disponível em: ISBI 2019 C-NMC Challenge: Classification in Cancer Cell Imaging.

[12] S. Mourya, S. Kant, P. Kumar, A. Gupta, R. Gupta, “Leukonet: DCT-Based CNN Architecture for the Classification of Normal Versus Leukemic Blasts in B-ALL Cancer”. Disponível em: <https://arxiv.org/pdf/1810.07961>

[13] V. Singhal, P. Singh, “Local binary pattern for automatic detection of acute lymphoblastic leukemia”. Disponível em: [https://ieeexplore.ieee.org/iel7/6802611/6811235/06811261.pdf?casa\\_token=E0uPP-o9bosAAAAA:-SnKoN3TUMqCjBBFxIr6TZbO6c5CjG72YhnXHZ6dwzw20LF3M5G\\_kkEnl6sAY9u5i7CrkaHTzOE](https://ieeexplore.ieee.org/iel7/6802611/6811235/06811261.pdf?casa_token=E0uPP-o9bosAAAAA:-SnKoN3TUMqCjBBFxIr6TZbO6c5CjG72YhnXHZ6dwzw20LF3M5G_kkEnl6sAY9u5i7CrkaHTzOE)