



# 67ª RBras e 20º SEAGRO

"A (Bio)estatística e a Biometria na era da revolução digital"

De 24 a 28 de julho de 2023

Londrina/Paraná

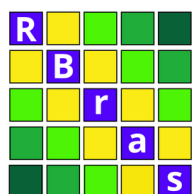
## LIVRO DE RESUMOS

### Patrocínio



**syngenta**

### Realização



**Departamento de Estatística**



UNIVERSIDADE  
ESTADUAL DE LONDRINA

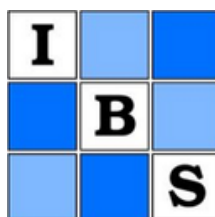
### Apoio



Associação Brasileira de Estatística



Conselho Regional de Estatística da 3ª Região (SP)



FUNDAÇÃO DE AMPARO À PESQUISA  
DO ESTADO DE SÃO PAULO

## Ensemble de modelos de transferência de aprendizado: Uma nova abordagem para detecção de câncer de mama

Fernando Humberto de Almeida Moraes Neto <sup>1</sup>

Adriano Kamimura Suzuki <sup>2</sup>

Francisco Louzada Neto <sup>3</sup>

Ricardo Rocha <sup>4</sup>

### Resumo

Recentemente os modelos de Deep Learning têm ganhado muito destaque na área de análise de imagem, porém necessitam de muitos dados para terem bons desempenhos. Isto dificulta a aplicação em bases de dados médicas visto que existe uma dificuldade em se obter muitas observações equiparadamente, entre os diferentes casos como, por exemplo, patológicos e normal. Para pequenas bases de dados pode-se utilizar modelos de transferência de aprendizado, ensemble e aumento de dados para se ter melhores desempenhos na tarefa de classificar imagens, já para base de dados desbalanceadas pode-se utilizar técnicas de reamostragem como undersampling e oversampling. Neste trabalho propomos uma nova abordagem baseada no ensemble de modelos de transferência de aprendizado com diferentes pesos para melhorar a predição de observações da classe minoritária em pequenas bases de dados desbalanceadas considerando reamostragem e aumento de dados. Ao final, um experimento é realizado com a base de dados do Hospital A.C. Camargo Cancer Center, que é um hospital oncológico especializado no diagnóstico, tratamento e pesquisa de câncer. Esta base de dados nunca foi utilizada para nenhum tipo de análise e tem-se como objetivo classificar imagens de mamografia como malignas ou benignas. A partir dos resultados obtidos nota-se que as combinações com melhores métricas são obtidas com a utilização de reamostragem e as piores sem a utilização desta técnica. O que comprova que a utilização da metodologia de reamostragem (undersampling e oversampling) melhora o desempenho da combinação em prever a classe minoritária. Ainda, com a utilização de aumento de dados, as combinações propostas têm um melhor desempenho do que a não utilização desta técnica.

**Palavras-chave:** Aumento de Dados; Ensemble; Rede Neural Convolucional; Transferência de aprendizado; Undersampling e Oversampling.

---

<sup>1</sup>Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos - fernandomoraes@usp.br

<sup>2</sup>Departamento de Matemática Aplicada e Estatística, Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, Brasil - suzuki@icmc.usp.br

<sup>3</sup>Departamento de Matemática Aplicada e Estatística, Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, Brasil - louzadaneto@gmail.com

<sup>4</sup>Departamento de Estatística, Instituto de Matemáticas e Estatística, Universidade Federal da Bahia, Bahia, Brasil - ricardo8610@gmail.com