

# Detecção de Câncer Cervical com o Exame de Papanicolau com Classificadores Rasos e Profundos

Arthur S. Quadros<sup>1</sup>, Quêzia P. Silva<sup>1</sup>, Sarah S. Magalhães<sup>1</sup>

<sup>1</sup>Instituto de Ciências Exatas e Informática  
Pontifícia Universidade Católica de Minas Gerais (PUC Minas)  
30.535-901 – Belo Horizonte – MG – Brasil

{aquadros, qsilva, sarah.magalhaes.1280966}@sga.pucminas.br

**Abstract.** *Este artigo investiga a eficácia de técnicas de aprendizagem rasa e profunda para detecção de câncer cervical usando imagens de exame de Papanicolau. Utilizamos os descritores Haralick de seis matrizes de coocorrência como features para um Support Vector Machine (SVM) e utilizamos a imagem inteira para o modelo de aprendizado profundo EfficientNetB1. Os resultados demonstram a superioridade do EfficientNetB1, alcançando uma acurácia de 85% e 65% nas tarefas de classificação binária e multiclasse (6 classes), respectivamente. Por outro lado, os classificadores SVM alcançaram precisões inferiores de 64% e 47% para os problemas binários e multiclasse. Estas descobertas sugerem que as abordagens de aprendizagem profunda são promissoras para melhorar a detecção de câncer cervical.*

## 1. Introdução

O exame de Papanicolau é um procedimento crítico na detecção precoce do câncer cervical. Ao coletar células do colo do útero e examiná-las ao microscópio, este teste pode identificar processos potencialmente pré-cancerosos e cancerosos [Koss 1989, Naucler et al. 2007]. A integração de Inteligência Artificial (IA) para esse fim potencializa a análise de imagens médicas, aumentando a acurácia e a rapidez na detecção do câncer. Os algoritmos de IA podem aprender com conjuntos de dados a reconhecer padrões indicativos de câncer, tornando o exame de Papanicolau uma ferramenta ainda mais potente no combate a esta doença.

Nesse estudo, empregamos duas abordagens para classificar imagens de Papanicolau do dataset em [Rezende et al. 2021]: um classificador raso usando descritores de Haralick - entropia, homogeneidade, contraste - derivados de múltiplas matrizes de coocorrência, e um classificador profundo analisando matrizes de pixels de imagens coloridas via EfficientNetB1, especificado na seção 3. Esses métodos foram aplicados a contextos binários e multiclasse. O aumento de dados quadruplicou nosso conjunto de dados com imagens giradas em ângulos de 0, 90, 180 e 270 graus. Avaliamos o desempenho de nossos modelos usando métricas de acurácia e confusão, bem como perdas especificamente para EfficientNetB1.

As implementações propriamente ditas foram feitas em Python com as bibliotecas Scikit-Learn para o uso do SVM, Tensorflow para uso do EfficientNetB1, OpenCV e Pillow para pré-processamento e manipulação de imagens, numpy para cálculo dos Descritores Haralick e Tkinter para criação de uma interface gráfica simples.

## 2. Trabalhos Relacionados

Estudos semelhantes utilizam técnicas de aprendizagem superficial e profunda para reconhecer indicadores de câncer. [Chen et al. 2023] alcançou mais de 90% de acurácia com 3 classes usando classificadores profundos EfficientNetB0, [Rastogi et al. 2023] alcançou 94% de acurácia em conjuntos de dados de benchmarking para classificação binária de células “normais” e “cancerígenas” usando classificadores profundos EfficientNetB7. [Zhang and Liu 2004] usa SVM para detectar câncer em classificações em nível de pixel, segmentando regiões da célula potencialmente identificando câncer, [Amole and Osalusi 2018] usa classificadores rasos SVM e k-NN para prever câncer em configurações binárias alcançando mais de 88% de acurácia para ambos os modelos. Com isso em mente, neste estudo propomos o uso de SVM e EfficientNetB1 para detectar câncer em exames de Papanicolau em classificação binária e de 6 classes, comparando os resultados de cada abordagem.

## 3. Metodologia

### 3.1. Pré-processamento das Imagens

As imagens utilizadas para o exame Papanicolau foram agrupadas por classe, sendo elas (0) “Negativas para lesão intraepitelial”, (1) “ASC-US”, (2) “ASC-H”, (3) “LSIL”, (4) “HSIL”, ou (5) “SCC”. O comprimento total do conjunto de dados inicial é de 5.581 imagens separadas de forma desigual entre as classes mencionadas. A configuração inicial das imagens está disposta na Tabela 1.

**Table 1. Configuração inicial das imagens no dataset utilizado**

Nome de classe	ID de classe	Tamanho
Negativas para lesão intraepitelial	0	4244
ASC-US	1	84
ASC-H	2	391
LSIL	3	287
HSIL	4	493
SCC	5	82

Todas essas imagens foram recortadas com dimensões 100x100 ao redor do núcleo das células e separadas cada uma em sua própria pasta com seus respectivos identificadores no banco de dados original.

### 3.2. Classificadores

Os classificadores implementados tiveram uma combinação de diferentes abordagens, totalizando 4 combinações de testes. Classificamos os exames de Papanicolau usando os Descritores de Haralick de entropia, homogeneidade e contraste, a partir das matrizes de co-ocorrência de  $C_{1,1}$ ,  $C_{2,2}$ ,  $C_{4,4}$ ,  $C_{8,8}$ ,  $C_{16,16}$  e  $C_{32,32}$ , (um total de 18 features) para um classificador raso; e os classificamos usando as matrizes que representam os pixels das imagens coloridas, com 3 canais e 10.000 pixels (um total de 30.000 features) para um classificador profundo. Ambas as abordagens foram aplicadas em um contexto binário, em que, com base na Tabela 1, temos a classe 0 como sendo negativa, e todas as outras como sendo positivas, assim como na versão completa, com todas as 6 classes. E

por fim, quanto aos classificadores em si, utilizamos um classificador raso, SVM, e um classificador profundo, EfficientNetB1.

O treinamento de todos os classificadores, exceto o classificador binário EfficientNetB1, incluiu os dados aumentados da tabela 1 na qual o conjunto de dados original acabou sendo 4 vezes maior. Todas as imagens originais foram giradas no sentido horário em quatro ângulos: 0 (imagem original), 90, 180 e 270 graus. Como o corte da imagem original mantém intencionalmente o núcleo da célula no centro, nenhum corte adicional ou edição de imagem é aplicado no conjunto de dados original.

Comparamos todas as métricas combinadas em termos de acurácia, perda (para EfficientNetB1) e confusão, por meio das matrizes de confusão da predição. O SVM foi implementado com as combinações de hiperparâmetros dos kernels *rbf* e *poly*, parâmetro de regularização (C) de 0,1, 1 e 10. Para o kernel *poly*, testamos juntamente com essas combinações o grau de 1 (linear), 2, 3, 4, 5, 6 e 7; e o EfficientNetB1 foi implementado com as combinações de hiperparâmetros de batch size 4, 8, 16, 32, 64 e 128, taxa de aprendizado de 0,001, 0,0001, 0,00005 e 0,00002 e 0,00001, 0, 1, 2, 3 e 4 camadas ocultas com 0, 4, 8, 16, 32, 64, 128, 256 e 512 neurônios cada um com dropout de 50%, otimizadores Adam e SGD (Stochastic Gradient Descent), todos com parada antecipada e redutores de taxa de aprendizagem.

#### 4. Resultados e Discussão

O cálculo das 18 features para todo o conjunto de dados incluiu o cálculo das 6 matrizes de coocorrência para cada imagem, e o resultado é usado para calcular sua entropia, homogeneidade e contraste. Este cálculo levou mais de 1 hora para ser executado no Google Colab. O tempo de treinamento para o melhor modelo raso foi de cerca de 30 minutos, enquanto o treinamento do classificador profundo levou bem mais de 3 horas (sem aceleração de GPU) para ser concluído.

A figura 1 foi usada como exemplo para demonstrar os resultados obtidos por todas as quatro implementações do modelo. A célula possui classe 1 nas classificações binária e multiclasse (ASC-H), com o SVM binário respondendo 1 (indicando corretamente câncer); o SVM multiclasse respondendo 1 (indicando corretamente ASC-H); o EfficientNetB1 binário respondendo 1 (indicando corretamente câncer); e o EfficientNetB1 multiclasse respondendo 0 (indicando incorretamente que não há câncer). Para esta imagem em particular, 3 em 4 modelos responderam corretamente.

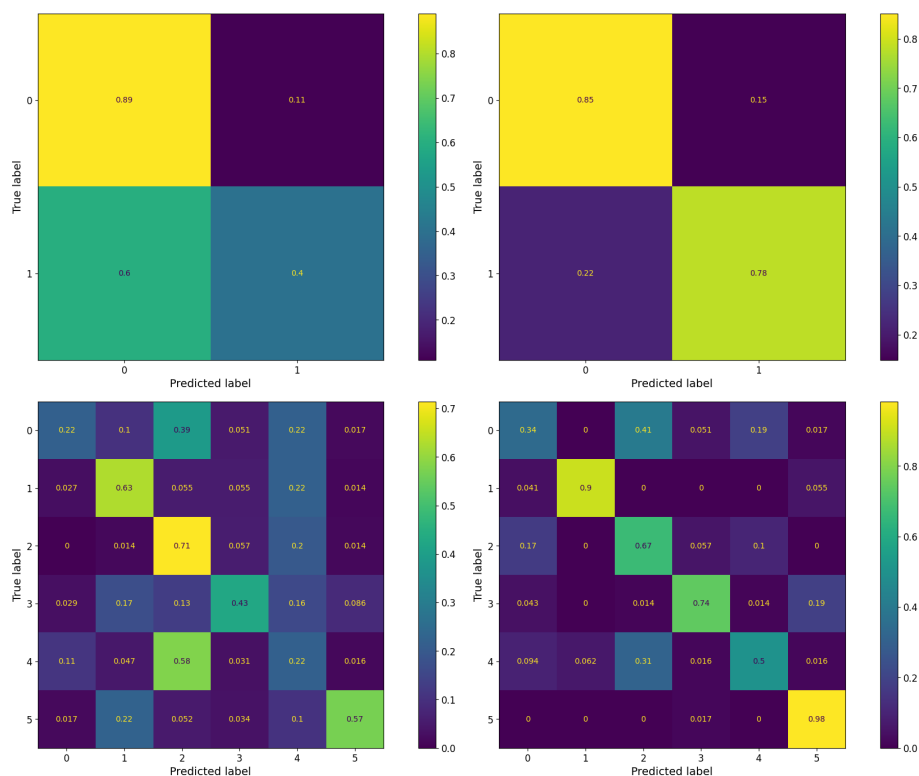


**Figure 1. Exemplo de imagem de ASC-H sendo classificada por todos os quatro modelos**

Os melhores resultados para SVM foram os parâmetros de kernel *poly*, parâmetro

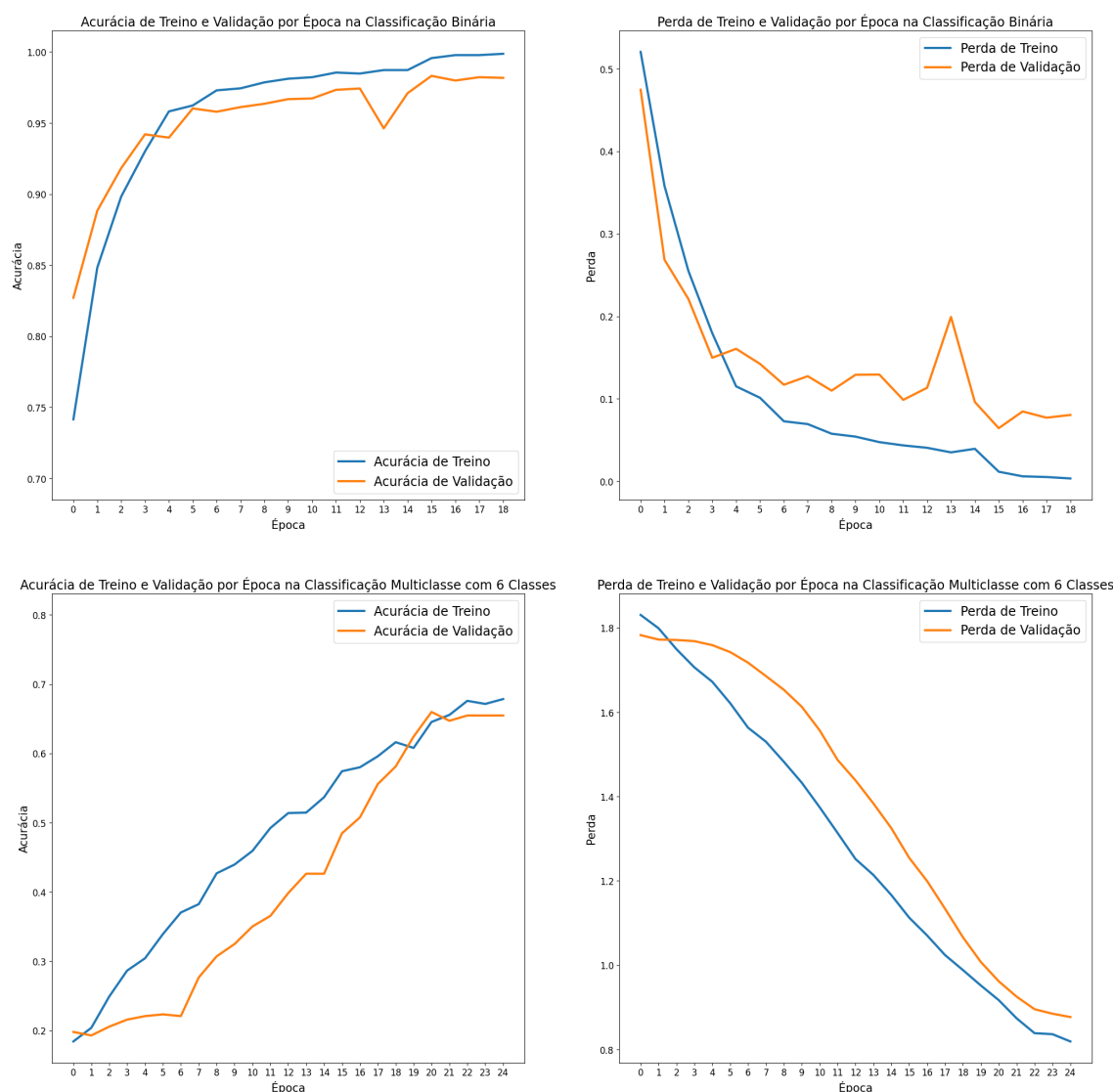
de regularização 10 e grau 7, tanto para a versão binária quanto para a versão multiclasse, com acurácia de 64% e 47% respectivamente. Enquanto isso, para os parâmetros EfficientNetB1, os melhores resultados foram obtidos para a versão binária com batch size de 16, taxa de aprendizado inicial de 0,00001, 3 camadas ocultas com 8 neurônios sem dropout e otimizador Adam; e as mesmas configurações para a versão multiclasse exceto o uso de batch size de 128, 2 camadas ocultas com 512 neurônios com dropout de 50%, com acurácias de 82% e 65%. A matriz de confusão normalizada para as melhores combinações de SVM e EfficientNetB1 é exibida na Figura 2.

Quanto aos resultados do EfficientNetB1, temos um classificador multiclasse que por si só é melhor que o SVM binário, com 65% de acurácia, e uma versão binária com 82% de acurácia. E simplesmente adaptando o EfficientNetB1 multiclasse para considerar qualquer previsão das classes 1, 2, 3, 4 ou 5 como sendo 1 e tratando o problema como binário com o classificador multiclasse conseguimos chegar a uma média de 85% de acurácia em uma média de 10 execuções em uma amostra de dados aleatória. A acurácia e perda de treinamento e validação são exibidas na Figura 3.



**Figure 2. Matrizes de confusão dos melhores modelos SVM binário (superior esquerdo) e EfficientNetB1 binário (superior direito), e SVM multiclasse (inferior esquerdo) e EfficientNetB1 multiclasse (inferior direito)**

No geral, os resultados para o classificador binário EfficientNetB1 foram fracos, uma vez que a adaptação binária do EfficientNetB1 multiclasse foi capaz de obter uma melhor acurácia média. O treinamento do classificador binário EfficientNetB1 não apresentou resultados muito confiáveis, com as métricas de treinamento da Figura 3 não sendo mantidas após o carregamento do modelo (caindo de 97% para 82%).



**Figure 3. Métricas de treinamento e validação para o EfficientNetB1 binário e multiclasse**

Nossos resultados indicam que as abordagens de Deep Learning com EfficientNetB1 são mais adequadas para exames de Papanicolau do que os classificadores rasos com o SVM. A versão multiclasse do nosso classificador profundo alcançou uma acurácia maior do que a versão binária do nosso classificador superficial, conforme observado na Figura 2.

## 5. Conclusão

Concluimos que as abordagens de Deep Learning são adequadas para classificações de imagens de exames de Papanicolau, possivelmente atingindo mais de 85% de acurácia na configuração binária para identificação do câncer cervical. Enquanto isso, a classificação superficial com SVM atingiu no máximo 65% de acurácia na configuração binária. Possíveis trabalhos futuros e continuações desse estudo incluem equilibrar o conjunto de dados com um downsampler com estimador de probabilidades para evitar a perda de informações vitais em imagens removidas devido à redução aleatória da amostragem na

tentativa de melhorar a acurácia final.

## References

- Amole, A. and Osalusi, B. S. (2018). Textural analysis of pap smears images for k-nn and svm based cervical cancer classification system. *Advances in Science, Technology and Engineering Systems Journal*.
- Chen, X., Pu, X., Chen, Z., Li, L., Zhao, K.-N., Liu, H., and Zhu, H. (2023). Application of efficientnet-b0 and gru-based deep learning on classifying the colposcopy diagnosis of precancerous cervical lesions. *Cancer Medicine*, 12(7):8690–8699.
- Koss, L. G. (1989). The Papanicolaou Test for Cervical Cancer Detection: A Triumph and a Tragedy. *JAMA*, 261(5):737–743.
- Naucler, P., Ryd, W., Törnberg, S., Strand, A., Wadell, G., Elfgrén, K., Rådborg, T., Strander, B., Johansson, B., Forslund, O., Hansson, B.-G., Rylander, E., and Dillner, J. (2007). Human papillomavirus and papanicolaou tests to screen for cervical cancer. *New England Journal of Medicine*, 357(16):1589–1597.
- Rastogi, P., Khanna, K., and Singh, V. (2023). Classification of single-cell cervical pap smear images using efficientnet. *Expert Systems*, 40(10):e13418.
- Rezende, M. T., Silva, R., Bernardo, F. d. O., Tobias, A. H., Oliveira, P. H., Machado, T. M., Costa, C. S., Medeiros, F. N., Ushizima, D. M., Carneiro, C. M., et al. (2021). Cric searchable image database as a public platform for conventional pap smear cytology data. *Scientific data*, 8(1):151.
- Zhang, J. and Liu, Y. (2004). Cervical cancer detection using svm based feature screening. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 873–880. Springer.