Revista Eletrônica
Acervo Saúde
Electronic Journal Collection Health ISSN 2178-2091

Uso de redes neurais convolucionais para classificação do câncer cervical com imagens citológicas do papanicolau

Using convolutional neural networks for cervical cancer classification with pap smear cytological images

Uso de redes neuronales convolucionales para la clasificación del cáncer cervical con imágenes citológicas del papanicolaou

Iumy Pimentel Farias[1], Pâmela Cristina Magalhães dos Santos[1], Gabriella Marreiros Staudt[1], Agatha Sousa Teixeira[1], João Vitor Jaques de Sousa[1], Élcio Leal[2], Marta de Oliveira Barreiros[1].

RESUMO

Objetivo: Implementar uma arquitetura de rede convolucional para identificar padrões de doenças em diferentes estágios da lesão cervical, utilizando dados de esfregaço de Papanicolau, que se referem às imagens obtidas a partir do exame, o qual é utilizado para detectar alterações nas células do colo do útero que podem ser indicativas do câncer cervical. **Métodos:** Trata-se de um estudo experimental, utilizando o conjunto de dados de imagens citológicas do Papanicolau em base de dados públicas SIPaKMeD, disponíveis de forma online do repositório Kaggle. O aprendizado de transferência proposto foi empregado usando modelos de rede neural convolucional (CNN) pré-treinados, como InceptionV3, Xception, VGG e uma abordagem modificada da VGG, para classificar imagens de esfregaço de Papanicolau de acordo com categorias de células: Disceratótica, Coilocitótica, Metaplásica, Parabasal e Superficial-Intermediária. **Resultados:** Os melhores resultados da classificação das categorias das células foram para as arquiteturas InceptionV3 e VGG com uma modificação nas camadas da arquitetura original, que apresentaram 91% de acurácia. Já a arquitetura Xception apresentou 88% de acurácia e VGG sem modificação nas camadas apresentou 86% no modelo treinado. **Conclusão:** Conclui-se que as presentes descobertas destacam a capacidade do modelo de discernir padrões e diferenciar células com precisão.

Palavras-chave: Câncer de colo do útero, Aprendizado de máquina, Redes neurais convolucionais, Diagnóstico precoce.

ABSTRACT

Objective: To implement a convolutional network architecture to identify disease patterns in different cervical surgery options using Pap smear data, which refer to the images obtained from the exam, which is used to detect changes in cervical cells that may be indicative of cervical cancer. **Methods**: This is an experimental study, using the Pap smear cytology image dataset in the public SIPaKMeD database, available online in the Kaggle repository. The proposed transfer learning was developed using pre-trained convolutional neural network (CNN) models, such as InceptionV3, Xception, VGG and a modified VGG approach, to classify Pap smear images according to cell categories: Dyskeratotic, Koilocytotic, Metaplastic, Parabasal and Superficial-Intermediate. **Results:** The best results in classifying cell categories were for the InceptionV3 and VGG architectures with a modification in the layers of the original architecture, which presented 91% accuracy. The Xception architecture presented 88% accuracy and VGG without modification in the layers presented 86% in the trained model. **Conclusion:** It is concluded that the findings highlight the model's ability to discern patterns and differentiate cells accurately.

Keywords: Cervical cancer, Machine learning, Convolutional neural networks, Early diagnosis.

[1] Universidade do Estado do Pará (UEPA), Castanhal - PA.
[2] Universidade Federal do Pará (UFPA), Belém - PA.

SUBMETIDO EM: 6/2025 | **ACEITO EM: 7/2025** | **PUBLICADO EM: 9/2025**

RESUMEN

Objetivo: Implementar una arquitectura de red convolucional para identificar patrones de enfermedad en diferentes opciones de cirugía cervical utilizando datos de Papanicolaou, que se refieren a imágenes obtenidas del examen, que se utiliza para detectar cambios en las células cervicales que pueden ser indicativos de cáncer cervical. **Métodos:** Se trata de un estudio experimental, en el que se utilizó el conjunto de datos de imágenes citológicas de frotis de Papanicolaou de la base de datos pública SIPaKMeD, disponible en línea en el repositorio Kaggle. El aprendizaje de transferencia propuesto se desarrolló utilizando modelos de redes neuronales convolucionales (CNN) entrenados previamente, como InceptionV3, Xception, VGG y un enfoque VGG modificado, para clasificar las imágenes de frotis de Papanicolaou según categorías de células: disqueratósicas, coilocíticas, metaplásicas, parabasales y superficiales-intermedias. **Resultados:** Los mejores resultados para la clasificación de categorías celulares fueron para las arquitecturas InceptionV3 y VGG con una modificación en las capas de la arquitectura original, las cuales presentaron una precisión del 91%. La arquitectura Xception presentó un 88% de precisión y VGG sin modificaciones en las capas presentó un 86% en el modelo entrenado. **Conclusión:** Se concluye que los hallazgos resaltan la capacidad del modelo para discernir patrones y diferenciar células con precisión.

Palabras clave: Cáncer de cuello uterino, Aprendizaje automático, Redes neuronales convolucionales, Diagnóstico temprano.

INTRODUÇÃO

O câncer do colo do útero é o terceiro tipo de câncer mais incidente entre mulheres, sendo a quarta causa mais comum de morte em todo o mundo. No Brasil, a taxa de mortalidade por câncer do colo do útero, ajustada pela população mundial, foi de 4,51 óbitos/100 mil mulheres, em 2021. Segundo o INCA - Instituto Nacional de Câncer (2023), a causa principal por malignidade cervical é dada pela contaminação pelo papilomavírus humano (HPV), transmitida sexualmente. O cancro do colo do útero é uma doença altamente evitável, e a estratégia mais eficaz para minimizar o fardo global do cancro do colo do útero é o rastreio precoce. Contudo, devido à falta de conscientização, à ausência de acesso a centros médicos e aos procedimentos altamente dispendiosos nos países em desenvolvimento, as populações de pacientes vulneráveis não conseguem arcar com os custos de realizar exames regularmente.

A detecção precoce do câncer cervical é essencial para aumentar significativamente as chances de cura. Quando os sinais pré-cancerígenos são identificados e tratados nas fases iniciais, é possível prevenir completamente o desenvolvimento do câncer do colo do útero. O exame de Papanicolau é amplamente reconhecido por sua eficácia na identificação de lesões precursoras, sendo uma ferramenta fundamental no monitoramento do câncer cervical e de outras condições relacionadas (SANKARANARAYANAN R, et al., 2015).

A classificação das lesões escamosas cervicais, segundo o sistema de Neoplasia Intraepitelial Cervical (NIC), é categorizada da seguinte forma: NIC1 – displasia leve; NIC2 – displasia moderada; NIC3 – displasia grave ou carcinoma espinocelular in situ; e SQCC (carcinoma de células escamosas invasivo) – estágio avançado da doença.

Nesse contexto, a automação dos processos de diagnóstico e classificação das lesões cervicais representa um avanço promissor, contribuindo para maior precisão, agilidade e padronização dos resultados. Essa abordagem está em consonância com os achados de diversos estudos na literatura (ALLAHQOLI L, et al., 2022; WANG CW, et al., 2021; SIMMS KT, et al., 2019; SCHIFFMAN M, et al., 2007).

Dessa forma, as mortes causadas por malignidades cervicais podem ser reduzidas através da implementação de estratégias eficazes de rastreamento e diagnóstico por meio de algoritmos computacionais inteligentes. Com base nisso, muitos estudos estão aplicando algoritmos de aprendizado de máquina para detectar o câncer do colo do útero precocemente.

No estudo de Allahqoli L, et al. (2022) estabeleceu-se a classificação do câncer cervical usando redes neurais convolucionais e máquinas de aprendizagem extrema. Além disso, foi desenvolvido um estudo comparativo de modelos de aprendizado de máquina para modelagem preditiva e ferramenta baseada na web para avaliação de risco de câncer cervical. Em estudos recentes, Ramahan, et al. (2021) desenvolveu o

framework DeepCervix, utilizando imagens do conjunto SIPaKMeD para classificar células cervicais com alta precisão. Esse trabalho evidenciou a superioridade das redes neurais convolucionais (CNNs) em relação a métodos tradicionais para a detecção de padrões celulares complexos. Com o avanço da tecnologia, o uso de sistemas inteligentes e processamento de imagens médicas tem se tornado essencial na análise de células malignas. Essas ferramentas são mais rápidas, acessíveis e vêm ganhando popularidade em comparação aos métodos tradicionais, como o exame de Papanicolau, a colposcopia e a cervicografia. Essas abordagens tecnológicas oferecem análises imparciais e independentes da experiência humana (XUE P, et al., 2020).

Portanto, o presente estudo visa implementar um algoritmo baseado em redes convolucionais, a fim de classificar imagens de esfregaços de Papanicolau de célula única (citologia) cervical em diferentes estágios. Além disso, o estudo inclui um comparativo de modelos de aprendizado de máquina com diferentes arquiteturas da literatura.

MÉTODOS

A base de dados utilizada no presente estudo é a SIPaKMeD, disponível de maneira online na plataforma Kaggle, contendo imagens de exame citológico de Papanicolau com cinco tipos diferentes de células: Disceratótica, Coilocitótica, Metaplásica, Parabasal e Superficial-Intermediária, que são categorizadas com base em suas características citomorfológicas (PLISSITI ME, et al., 2018). O conjunto de dados SIPaKMeD compreende um total 4068 imagens, porém foram usadas 2.257 cortadas por células. Segundo o repositório SIPaKMeD, as características de cada classe são descritas como:

Células Superficiais-Intermediárias (Superficial-Intermediate cells) são as células mais comuns nos esfregaços cervicovaginais, variando em predominância conforme a fase do ciclo menstrual, a gravidez ou outras influências hormonais. Apresentam formato plano, com citoplasma abundante, de coloração eosinofílica ou cianofílica, podendo ser redondo, oval ou poligonal. O núcleo pode ser vesicular ou picnótico, dependendo do grau de maturação celular, sendo esta característica essencial para diferenciá-las entre si (LIMA DNO, 2012).

As Células Parabasais (Parabasal cells) são células escamosas imaturas e representam as menores células epiteliais observadas em um esfregaço vaginal típico. Elas predominam em condições de deficiência estrogênica, como na infância, lactação e menopausa, quando o epitélio se torna atrófico. Além disso, possuem formato arredondado, seu citoplasma é geralmente cianofílico, e apresentam um núcleo grande, vesicular, de formato redondo ou oval, com grânulos de cromatina ou cromocentros, características que auxiliam na sua identificação (LIMA DNO, 2012).

As Células Coilocitóticas (Koilocytotic cells) são células escamosas alteradas, frequentemente associadas à infecção pelo papilomavírus humano (HPV), e desempenham um papel fundamental na identificação de lesões cervicais precursoras de neoplasia. Sua principal característica morfológica é a presença de uma grande cavidade perinuclear clara, resultante da condensação do citoplasma, que pode apresentar coloração rosa, azul ou mista com a coloração de Papanicolaou, além disso, essas células exibem núcleos aumentados, hipercromáticos, com contornos irregulares e, em alguns casos, binucleação. A presença de coilócitos é um critério importante no diagnóstico citológico da infecção pelo HPV, contudo, sua interpretação pode variar entre observadores, tornando necessária a associação com outros métodos diagnósticos para maior precisão (KRAWCZYK E, et al., 2008).

As células Discetáróticas (Dysketarotic cells) são células escamosas que apresentam uma ceratinização anormal e prematura, podendo ser observadas em lesões de pele e mucosas, frequentemente associadas ao HPV. Essas células podem ser encontradas em condições benignas e também em lesões precursoras de câncer, como displasia escamosa e carcinoma cervical. Sua presença em esfregaços cervicais pode ser um indicativo de alterações epiteliais associadas a lesões precoces no colo do útero (MYPATHOLOGYREPORT, 2023).

Células Metaplásicas (Metaplastic Cells) são células que passaram por um processo de transformação denominado metaplasia, no qual um tipo de célula adulta é substituído por outro tipo celular. Esse fenômeno ocorre em resposta a fatores ambientais, como irritação crônica, inflamação ou alterações hormonais. No caso do colo do útero, a metaplasia escamosa é uma das transformações mais comuns, caracterizando-se pela substituição do epitélio glandular do endocérvix por epitélio escamoso.

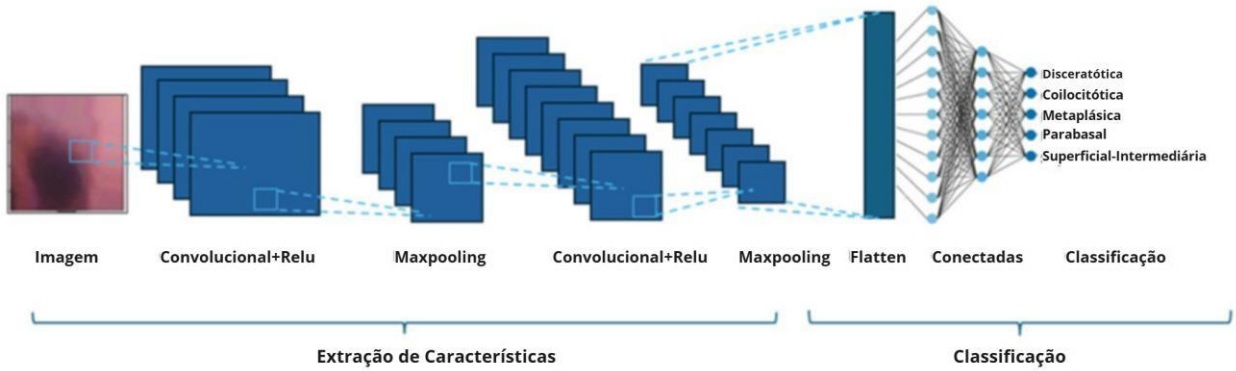
As células Metaplásicas apresentam morfologia escamosa, citoplasma espesso e mais eosinofílico que o normal, além de núcleos frequentemente aumentados ou discretamente deformados. São comumente observadas em exames citológicos, como o teste de Papanicolau, e sua presença pode estar relacionada a condições fisiológicas, como a adolescência e o uso de anticoncepcionais, ou ainda a situações patológicas, como a exposição crônica a infecções ou processos inflamatórios (VAL R, 2021).

As respectivas distribuições das categorias são: Discetáróticas: 27 imagens para teste e 406 para treinamento. Coilocitóticas: 29 imagens para teste e 413 para treinamento. Metaplásicas: 33 imagens para teste e 382 para treinamento. Parabasal: 13 imagens para teste e 394 para treinamento. Superficial-Intermediária: 15 imagens para teste e 416 para treinamento.

Algoritmos de redes convolucionais

As redes convolucionais são modelos projetados para processar dados como imagens, extraindo automaticamente características locais por meio de camadas de convolução e pooling. Elas usam conexões locais, compartilhamento de pesos e funções de ativação para capturar padrões simples a complexos. No final, os dados são transformados em 1D para servir de entrada nas camadas totalmente conectadas, e assim, realizam a classificação ou predição, como mostrada na **Figura 1**.

Figura 1 - Diagrama representativo de uma arquitetura de redes convolucionais.



Nota: Elaboração própria realizada com o auxílio do Microsoft PowerPoint (versão Microsoft 365), sob licença institucional.

Fonte: Farias IP, et al., 2025

Para o tratamento dos dados, visando melhorar a generalização e reduzir o risco de risco de sobreajuste (overfitting), foram definidas configurações específicas para as imagens de treinamento. O congelamento das 10 primeiras camadas foi implementado para preservar o conhecimento previamente adquirido, permitindo que apenas as camadas finais fossem ajustadas para a nova tarefa. Além disso, ao reduzir o número de parâmetros a serem atualizados, pode acelerar o processo de treinamento.

Todos os hiperparâmetros de treinamento foram mantidos idênticos em todos os experimentos. As arquiteturas de redes foram treinadas por 80 épocas usando uma perda de entropia cruzada categórica, um tamanho de lote de 16 e otimizador Adagrad. Todos os modelos foram treinados com uma taxa de aprendizado inicial de 0,001, que altera seu valor por um fator de 0,5 se não houver incremento na precisão da validação ao longo de 10 épocas consecutivas até atingir um valor de 0,00001.

Inception V3

InceptionV3 é uma arquitetura de rede neural convolucional (CNN) profunda, uma evolução do modelo Inception proposto por Szegedy C, et al. (2016), projetada para tarefas de visão computacional, como classificação de imagens, detecção de objetos e segmentação. A principal inovação do InceptionV3 são os módulos inception, que realizam múltiplas operações de convolução com diferentes tamanhos de filtro (1x1, 3x3, 5x5) simultaneamente, combinando os resultados em uma única camada. Isso permite que a rede capture informações de diferentes escalas e contextos de forma mais eficiente. Para lidar com características

visuais em diferentes escalas, o modelo combina pooling com essas camadas de convolução, sejam elas assimétricas e de concatenação. O InceptionV3 utiliza convoluções 1x1, que desempenham um papel importante na redução da dimensionalidade das ativações, ou seja, diminuem o número de parâmetros na rede, tendo 48 camadas de profundidade. Além disso, após cada convolução, a técnica de normalização em lotes (BatchNormalization) introduzida por Sergey I e Szegedy C (2014) também é aplicada para reduzir as flutuações nas ativações, estabilizando o treinamento e prevenindo o sobreajuste. Essa técnica torna o modelo mais leve e eficiente, sem comprometer a capacidade de aprender representações complexas das imagens (CHRISTIAN S, et al., 2015).

Xception

A Xception (Extreme Inception) é uma arquitetura de rede neural convolucional (CNN) profunda proposta por François Chollet, o criador do Keras, François C (2017). Ela é uma variação da arquitetura Inception, com uma inovação no uso de convoluções separáveis em profundidade, com pooling integrado em diferentes camadas para reduzir dimensionalidade espacial e preservar características de alto nível ao mesmo tempo em que prepara elas para serem extraídas para camadas finais de decisão, com um conceito previamente introduzido em redes como o MobileNet. Diferente das convoluções tradicionais, em que um único filtro é aplicado a todos os canais de entrada simultaneamente, o Xception organiza a rede em módulos sequenciais compostos por convoluções que são separáveis em profundidade. Assim, esse design permite que a rede aprenda representações mais complexas e detalhadas de maneira mais eficiente.

A convolução separável em profundidade é composta por duas etapas: Convolução Depthwise, onde cada filtro é aplicado de forma independente a um único canal de entrada, sem interação entre os canais; Convolução Pontual (1x1), após a convolução depthwise, é aplicada uma convolução 1x1 para combinar as saídas de diferentes canais, permitindo que a rede aprenda as relações entre eles. Todo esse processo reduz significativamente o número de parâmetros e o custo computacional ao mesmo tempo em que mantém ou até melhora a capacidade da rede neural convolucional captar informações complexas.

VGG16

O modelo VGG16, proposto por Simonyan K e Zisserman A (2015), é estruturado em blocos sequenciais que combinam diferentes camadas para extração, estabilização e compactação de características, totalizando 13 camadas convolucionais e 3 camadas totalmente conectadas, perfazendo 16 camadas treináveis. As camadas de convolução 2D utilizam filtros 3x3 com passo 1 para extrair padrões como bordas, texturas e formas simples nas etapas iniciais, enquanto nas posteriores capturam representações mais complexas, utilizando um número crescente de filtros. Após conjuntos de camadas convolucionais, são aplicadas as camadas de MaxPooling2D com filtros de 2x2 e passo 2, que compactam as características extraídas ao selecionar os valores máximos de regiões da imagem, reduzindo a dimensionalidade e aumentando a eficiência computacional, além de capturar invariância e pequenas translações (SCHERER D, et al., 2010).

VGG Modificada

Para esta pesquisa foi desenvolvido um algoritmo de rede neural convolucional baseado na família de redes VGG. A arquitetura é composta por 7 camadas de convolução, 4 camadas MaxPooling e duas BatchNormalization.

A entrada da camada consiste em imagens do exame de Papanicolau que foram redimensionadas na dimensão 244x244 com 3 canais básicos de cores nas imagens digitais coloridas (RGB), sendo direcionados para o primeiro bloco do modelo que se inicia com duas camadas de convolução (Conv2D) com 16 filtros para a extração inicial de características como bordas, texturas e formas básicas nas camadas iniciais, seguindo por uma camada de normalização(BatchNormalization) em lotes para a estabilização do processo de treino e permitindo o uso de aprendizado mais altas, MaxPooling2D que compacta as características extraidas e preserva as informações mais importantes e finalizando com Dropout (0.2) que desativa aleatoriamente 20% dos neurônios durante o treinamento buscando evitar o sobreajuste (overfitting), o segundo bloco segue executando duas camadas seguidas de convolução (Conv2D) agora com 32 filtros para aprender as diferentes características das imagens, com a mesma sequência de blocos (BatchNormalization, MaxPooling2D e Dropout).

O terceiro bloco dispõe de três camadas de Conv2D, com um aumento de número de filtros, contando com 64 filtros cada e finalizando com uma camada de MaxPooling2D. Após a execução das camadas e extração das características, é utilizado uma camada de Global Max Pooling (GlobalMaxiPooling2), que reduz toda a matriz de características para um vetor unidimensional por canal, preservando os valores máximos mais representativos. O fluxo de saída consiste na aplicação de uma camada densa (Dense, 64, ReLU) com 64 neurônios e ativação da ReLU para introdução da não-linealidade, sendo aplicada novamente uma camada de normalização de lote (BatchNormalization) e por fim uma camada Dropout de 50%. Finalizando o modelo com a camada de saída (Dense, output) que é uma camada densa com o número de neurônios (nbout) sendo o mesmo valor do número de classes e ativação de softmax.

Essas mudanças se fizeram necessárias a fim de aumentar as métricas de avaliação em comparação com o modelo Vgg_16 proposto por Simonyan K e Zisserman A (2015), buscando uma otimização e refinamento da arquitetura para os dados disponíveis, o que apresentou resultados muito positivos.

Métricas de Avaliação

Para avaliar o desempenho dos modelos, utilizou-se a matriz de confusão, uma ferramenta que permite verificar a eficácia do modelo na identificação correta das diferentes classes. A matriz compara as predições realizadas com os valores reais, sendo que cada linha e coluna representam as classes Disceratótica, Coilocitótica, Metaplásica, Parabasal e Superficial-Intermediária. Cada valor registrado indica a quantidade de vezes que uma classe foi corretamente ou incorretamente classificada em relação às demais, categorizando os resultados em: Verdadeiros Positivos (VP), quando a predição correta é positiva; Verdadeiros Negativos (VN), quando a predição correta é negativa; Falsos Positivos (FP), quando o modelo prediz incorretamente uma classe positiva; e Falsos Negativos (FN), quando falha em identificar uma classe positiva.

A avaliação dos modelos também considerou as seguintes métricas:

- **Acurácia:** mede a proporção de previsões corretas em relação ao total de previsões realizadas.
- **Precisão:** indica a proporção de previsões positivas corretas em relação ao total de predições positivas feitas pelo modelo.
- **Recall:** avalia a capacidade do modelo de identificar corretamente os casos positivos, ou seja, a taxa de detecção dos verdadeiros positivos.
- **F1-score**: representa a média harmônica entre precisão e recall, equilibrando essas duas métricas. É calculado como o dobro do produto entre precisão e recall dividido pela soma dos dois valores.

RESULTADOS E DISCUSSÃO

Os modelos Vgg_Mod e InceptionV3 apresentaram desempenho satisfatório na classificação de células em exames de Papanicolau, alcançando uma acurácia de 91% (**Tabela 1**). O modelo Vgg_Mod destacou-se especialmente nas métricas associadas a células de alto risco, com os seguintes resultados: Disceratótica: Precisão de 90%, Recall de 93% e F1-Score de 91%; Coilocitóticas: Precisão de 84%, Recall de 85% e F1-Score de 85%; Metaplásicas: Precisão de 90%, Recall de 88% e F1-Score de 89%; Parabasal: Precisão de 93%, Recall de 99% e F1-Score de 96%; Superficial-Intermediária: Precisão de 89%, Recall de 93% e F1-Score de 93%.

De forma semelhante, o modelo baseado na arquitetura InceptionV3 também obteve resultados expressivos: Disceratótica: Precisão de 93%, Recall de 87% e F1-Score de 90%; Coilocitóticas: Precisão de 85%, Recall de 86% e F1-Score de 86%; Metaplásicas: Precisão de 85%, Recall de 95% e F1-Score de 89%; Parabasal: Precisão de 96%, Recall de 91% e F1-Score de 93%; Superficial-Intermediária: Precisão de 97%, Recall de 96% e F1-Score de 97%.

A matriz de confusão, ilustrada na **Figura 2**, demonstra a capacidade dos modelos em distinguir eficientemente células de risco em comparação a células normais ou de baixo risco, como Parabasal e Superficial-Intermediária, reforçando o seu potencial de aplicação clínica. Apesar do desempenho positivo,

algumas classes com características semelhantes geraram erros na classificação das classes, problema recorrente na classificação de células cervicais, como relatado na literatura recente publicada por Tan SL, et al. (2024). Esse desafio é atribuído ao desbalanceamento de dados e à similaridade entre o rótulo e o que foi evidenciado com o modelo Vgg_Mod entre as células Metaplásica e Superficial-Intermediária. Da mesma forma, o modelo InceptionV3, entre Metaplásica e Parabasal, ambos podendo ser analisados pela métrica Recall, na **Tabela 1**, reflete a habilidade do modelo em identificar verdadeiramente as células de cada classe.

Tabela 1 - Tabela das métricas de avaliação dos modelos.

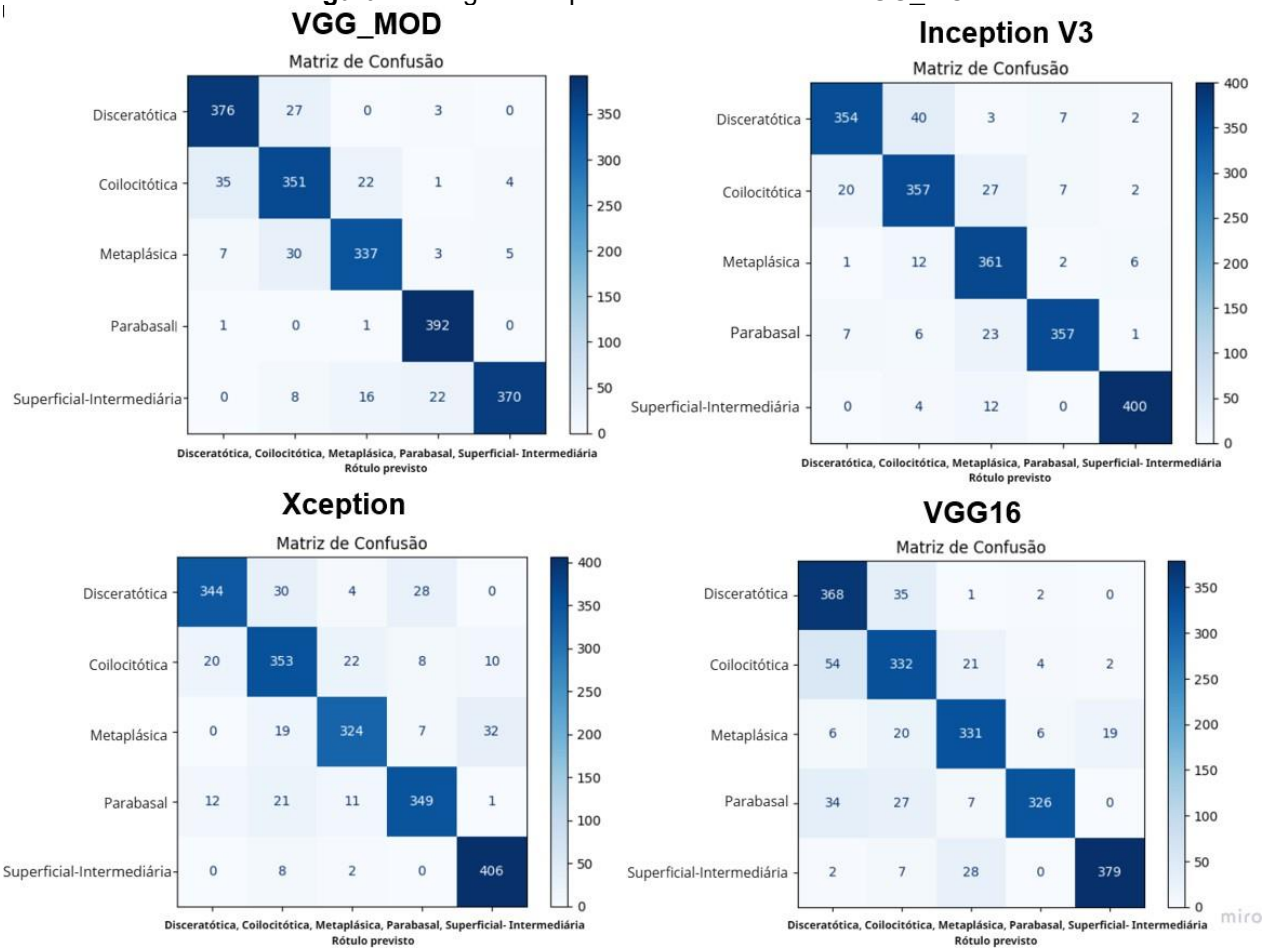
Modelo/Classes	Precisão (%)	Recall (%)	F1-Score (%)	Acurácia (%)
VGG16				
Disceratótica	79%	91%	85%	
Coilocitótica	79%	80%	80%	
Metaplásica	85%	87%	86%	86%
Parabasal	96%	83%	89%	
Superficial- Intermediária	95%	91%	93%	
InceptionV3				
Disceratótica	93%	87%	90%	
Coilocitótica	85%	86%	86%	
Metaplásica	85%	95%	89%	91%
Parabasal	96%	91%	93%	
Superficial- Intermediária	97%	96%	97%	
Xception				
Disceratótica	91%	85%		
Coilocitótica	82%	85%	88%	
Metaplásica	89%	85%	84%	88%
Parabasal	89%	89%	87%	
Superficial- Intermediária	91%	98%	94%	
VGG_MOD				
Disceratótica	90%	93%	91%	
Coilocitótica	84%	85%	85%	
Metaplásica	90%	88%	89%	91%
Parabasal	93%	99%	96%	
Superficial-Intermediária	98%	89%	93%	

Fonte: Farias IP, et al., 2025.

Observa-se uma diferença nos resultados obtidos pelos modelos Vgg_16 e Vgg_Mod. Uma possível explicação para esse comportamento são as alterações realizadas nas camadas do Vgg_Mod, que indicam um refinamento em sua arquitetura. Essas modificações contribuíram para um melhor desempenho do modelo, mesmo diante de adversidades.

Para melhorar o desempenho do modelo, o aumento de dados e a diversificação podem ser úteis para aumentar a assertividade do modelo, espera-se que, ao utilizar dados de diferentes fontes, seja possível aumentar a variabilidade das amostras, o que pode ajudar a reduzir ainda mais os erros de classificação e melhorar a capacidade do modelo em ambientes mais diversos. Ademais, a utilização de técnicas de aumento de dados e refinamento de arquiteturas podem contribuir significativamente para superar os desafios observados. Esses resultados estão alinhados com a literatura recente que destaca o potencial da inteligência artificial na medicina, como o DeepCervix (RAHMAN, et al., 2021).

Figura 2 - Diagrama representativo do modelo VGG_MOD.



Nota: Imagem elaborada por meio do software Miro (www.miro.com), sob licença gratuita para uso individual.
Fonte: Farias IP, et al., 2025.

Recentemente, um estudo de Mathivanan S, et al. (2024) também usou a base de dados SIPaKMeD apresentou uma metodologia que aproveita modelos de rede neural profunda pré-treinadas, como Alexnet, Resnet-101, Resnet-152 e InceptionV3, para extração de recursos, combinando com outros algoritmos de aprendizado de máquina, como árvore de decisão, Floresta aleatória e outros, a fim de melhorar a classificação. Em uma abordagem binária alcançou até 98%. No entanto, a classificação multiclasse, como a proposta neste trabalho, é mais desafiadora pois depende de um grande volume de dados para generalização do problema multiclasse.

Este trabalho reforça a relevância do uso de ferramentas baseadas em aprendizado profundo para o diagnóstico precoce do câncer cervical, demonstrando sua eficácia e potencial de aplicação assertiva em contextos clínicos, conforme evidenciado na literatura (CROFFIL B, et al., 2024). Além disso, destaca-se a importância de explorar uma variedade de abordagens de aprendizado de máquina, a fim de identificar os padrões mais eficazes para o diagnóstico da doença, fortalecendo tanto a precisão quanto a flexibilidade necessária para enfrentar essa desafiadora tarefa de classificação.

CONCLUSÃO

O presente estudo desenvolveu e implementou uma arquitetura de rede neural convolucional voltada para identificação de padrões celulares em diferentes estágios da lesão cervical, utilizando imagens de esfregaço de Papanicolau. Com base nos resultados obtidos, conclui-se que os modelos Vgg_Mod e InceptionV3 se destacaram com altas taxas de acurácia, demonstrando eficácia na classificação de células, especialmente as de alto risco. As modificações na arquitetura do Vgg_Mod contribuíram para um desempenho superior,

evidenciando a importância da personalização dos modelos. Apesar de confusões entre classes semelhantes, os resultados reforçam o potencial da inteligência artificial como ferramenta auxiliar no diagnóstico precoce do câncer cervical. Assim, futuros estudos podem explorar estratégias de aumento de dados e novos ajustes nas arquiteturas para alcançar ainda maior robustez e generalização em contextos clínicos reais.

REFERÊNCIAS

1. ALLAHQOLI L, et al. Diagnosis of Cervical Cancer and Pre-Cancerous Lesions by Artificial Intelligence: A Systematic Review. Diagnostics (Basel), 2022; 12(11): 2771.
2. CHRISTIAN S, et al. Rethinking the Inception Architecture for Computer Vision. CVPR, 2015.
3. CROFFIL. B, et al. Inteligência artificial aplicada ao diagnóstico de câncer por exames de imagem. Revista Eletrônica Acervo Científico, 2024; 47: e16193.
4. FRANÇOIS C. Xception: Deep Learning with Depthwise Separable Convolutions. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017; 1800-1807.
5. INCA.2022. In: Estimativa 2023: incidência de câncer no Brasil. Instituto Nacional de Câncer. Disponível em: https://www.inca.gov.br/sites/ufu.sti.inca.local/files//media/document//estimativa-2023.pdf. Acesso em: 10 dez. 2024.
6. KRAWCZYK E, et al. Koilocytosis: A Cooperative Interaction between the Human Papillomavirus E5 and E6 Oncoproteins. The American Journal of Pathology, 2008; 173(3): 682 – 688.
7. LIMA DNO. Atlas de citopatologia ginecológica, Ministério da Saúde, 2012. Disponível em: https://bvsms.saude.gov.br/bvs/publicacoes/atlas_citopatologia_ginecologica.pdf. Acesso em: 08 mar. 2025.
8. MATHIVANAN S, et al. Enhancing cervical cancer detection and robust classification through a fusion of deep learning models. Sci Rep, 2024; 14: 10812.
9. MEHANDIRATTA P. 2021. In: Cervical Cancer - Largest Dataset (SIPaKMeD). Kaggle. [Online]. Disponível em: https://www.kaggle.com/datasets/prahladmehandiratta/cervical-cancer-largest-dataset-sipakmed. Acesso em: 10 dez. 2024.
10. MYPATHOLOGYREPORT. Células disqueratóticas. MyPathologyReport.ca, 2023. Disponível em: https://www.mypathologyreport.ca/pt/pathology-dictionary/dyskeratotic-cells/. Acesso em: 08 mar. 2025.
11. PLISSITI ME, et al. Sipakmed: A New Dataset for Feature and Image Based Classification of Normal and Pathological Cervical Cells in Pap Smear Images. In: 2018 25th IEEE International Conference on Image Processing (ICIP), 2018. Acesso em: 08 mar. 2025.
12. RAHAMAN, et al. DeepCervix: A deep learning-based framework for the classification of cervical cells using hybrid deep feature fusion techniques. Computers in Biology and Medicine, 2021; 136: 104649.
13. SANKARANARAYANAN R, et al. Oral Cancer: Prevention, Early Detection, and Treatment. Cancer: Disease Control Priorities, Third Edition (Volume 3). Washington (DC): The International Bank for Reconstruction and Development / The World Bank; 2015. Chapter 5. Disponível em: https://www.ncbi.nlm.nih.gov/books/NBK343649/
14. SCHIFFMAN M, et al. Human Papillomavirus and Cervical Cancer. Lancet, 2007, 370(9590):890–907.
15. SIMMS KT, et al. Impact of Scaled Up Human Papillomavirus Vaccination and Cervical Screening and the Potential for Global Elimination of Cervical Cancer in 181 Countries, 2020-99: A Modelling Study. Lancet Oncol., 2019; 20(3): 394–407.
16. SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for largescale image recognition. In: International Conference on Learning Representations (ICLR), 2015. [Online]. Available: https://arxiv.org/abs/1409.1556.
17. TAN SL, et al. Cervical Cancer Classification From Pap Smear Images Using Deep Convolutional Neural Network Models. Interdiscip Sci Comput Life Sci., 2024; 16: 16–38.
18. VAL R. O que é Metaplasia escamosa?. Clínica Rubens do Val, 2021. Disponível em: https://clinicarubensdoval.com.br/o-que-e-metaplasia-escamosa/. Acesso em: 08 mar. 2025.
19. WANG CW, et al. Artificial intelligence-assisted fast screening cervical high grade squamous intraepithelial lesion and squamous cell carcinoma diagnosis and treatment planning. Sci Rep 11, 16244 (2021).
20. XUE P, et al. Development and validation of an artificial intelligence system for grading colposcopic impressions and guiding biopsies. BMC Med, 2020; 18(1): 406.