

# Classificação multicategórica utilizando aprendizagem profunda aplicada ao diagnóstico de adenocarcinoma gástrico

## *Multi-categorical classification using deep learning applied to the diagnosis of gastric cancer*

Jonas Kloeckner<sup>1</sup>; Tatiana K. Sansonowicz<sup>2</sup>; Áttila L. Rodrigues<sup>1</sup>; Tatiana W. N. Nunes<sup>2</sup>

1. Universidade Federal do Rio Grande do Sul (UFRS), Porto Alegre, Rio Grande do Sul, Brasil.

2. Hospital Nossa Senhora da Conceição, Porto Alegre, Rio Grande do Sul, Brasil.

### RESUMO

**Introdução:** Os patologistas enfrentam atualmente um aumento substancial na carga e na complexidade de seu trabalho diagnóstico em diferentes tipos de câncer. Isso ocorre devido ao aumento da incidência e da detecção de neoplasias, além da subespecialização diagnóstica e do advento da medicina personalizada. Existem inúmeros tratamentos disponíveis para diferentes tipos de câncer, e o diagnóstico deve ser dado com celeridade e precisão para cada caso. A aprendizagem profunda é uma ferramenta que vem sendo empregada no dia a dia, inclusive na detecção de imagens, e há crescente interesse em sua aplicação na Medicina, especialmente na Patologia, área em que apresenta potencial revolucionário. **Objetivo:** Neste artigo, apresentamos a aprendizagem profunda, em específico as redes neurais convolucionais, como uma potencial técnica para a análise de imagens digitalizadas de lâminas histopatológicas, detectando padrões identificáveis de forma automatizada, introduzindo a possibilidade de aplicação dessa tecnologia como ferramenta auxiliar no diagnóstico de neoplasias, principalmente no adenocarcinoma gástrico, objeto deste estudo preliminar. **Métodos:** A partir de um banco de dados de imagens digitalizadas de lâminas histopatológicas representativas de adenocarcinoma gástrico, identificamos três padrões morfológicos da neoplasia, bem como padrões de tecidos não neoplásicos, com os quais treinamos um algoritmo de rede neural convolucional, criado com a finalidade de identificar e categorizar imagens similares dentro desses padrões, de forma automatizada. **Resultados:** Os resultados de identificação e classificação automática nas categorias definidas mostraram-se satisfatórios, com curvas ROC acima de 0,9. **Conclusão:** Os resultados evidenciam o potencial de aplicação das redes neurais convolucionais em lâminas digitalizadas de adenocarcinoma gástrico, consoantes com a literatura internacional.

**Unitermos:** redes neurais (computação); neoplasias gástricas; aprendizagem profunda/modelo de aprendizagem de máquina.

### ABSTRACT

**Introduction:** Pathologists currently face a substantial increase in workload and complexity of their diagnosis work on different types of cancer. This is due to the increased incidence and detection of neoplasms, associated with diagnostic subspecialization and the advent of personalized medicine. There are numerous treatments available for different types of cancer, and the diagnosis must be dispensed quickly and accurately for each case. Deep learning is a tool that has been used in daily life, including image detection, and there is growing interest in its application in Medicine and especially in Pathology, where it has a revolutionary potential. **Objective:** In this article, we present deep learning, in particular convolutional neural networks, as a potential technique for the analysis of digitized images of histopathological slides, detecting identifiable patterns in an automated manner, introducing the possibility of applying this technology as an auxiliary tool in the diagnosis of neoplasms, especially in gastric cancer, the object of this preliminary study. **Method:** From a database of digitized images of histopathological slides representative of gastric cancer, we

*identified three morphological patterns of neoplasia, as well as non-neoplastic tissue patterns, with which we train a convolutional neural network algorithm, designed to identify and categorize similar images within these standards, in an automated manner. Results: The results of identification and automatic classification in the defined categories were satisfactory, with ROC curves above 0.9. Conclusion: The results show the potential application of convolutional neural networks for digitized slides of gastric cancer, in accordance with international literature findings.*

**Key words:** neural networks (computer); gastric neoplasms; deep learning/machine learning model.

---

## RESUMEN

**Introducción:** Los patólogos enfrentan actualmente un aumento sustancial de su trabajo diagnóstico en diferentes tipos de cáncer. Eso ocurre debido al incremento de la incidencia y de la detección de neoplasias, además de la subespecialización diagnóstica y del advenimiento de la medicina personalizada. Hay numerosos tratamientos disponibles para diferentes tipos de cáncer, y el diagnóstico debe ser realizado con celeridad y precisión para cada caso. El aprendizaje profundo es una herramienta que ha sido empleada en el día a día, incluso en la detección de imágenes, y hay creciente interés en su aplicación en Medicina, especialmente en Patología, área en la que presenta potencial revolucionario. **Objetivo:** En este artículo presentamos el aprendizaje profundo, en especial las redes neuronales convolucionales, como una técnica potencial para el análisis de imágenes digitalizadas de portaobjetos histopatológicos, detectando patrones identificables de forma automatizada, introduciendo la posibilidad de empleo de esa tecnología como herramienta auxiliar en el diagnóstico de neoplasias, principalmente en el adenocarcinoma gástrico, objeto de este estudio preliminar. **Métodos:** A partir de una base de datos de imágenes digitalizadas de portaobjetos histopatológicos representativos de adenocarcinoma gástrico, identificamos tres patrones morfológicos de la neoplasia, así como patrones de tejidos no neoplásicos, con los cuales entrenamos un algoritmo de red neuronal convolucional, creado para identificar y categorizar imágenes semejantes dentro de esos patrones, de modo automatizado. **Resultados:** Los resultados de identificación y clasificación automática en las categorías definidas se mostraron satisfactorios, con curvas ROC por encima de 0,9. **Conclusión:** Los resultados muestran el potencial de aplicación de las redes neuronales convolucionales en portaobjetos digitalizados de adenocarcinoma gástrico, en conformidad con la literatura internacional.

**Palabras clave:** redes neurales/neuronales (computación); neoplasias gástricas; aprendizaje profundo/modelo de aprendizaje de máquinas.

---

## INTRODUÇÃO

O diagnóstico anatomopatológico de lesões neoplásicas dá-se substancialmente pela análise de lâminas coradas pelo método de hematoxilina e eosina (HE)<sup>(1)</sup>, avaliadas por meio de um microscópio óptico por um patologista treinado. Tal método vem sendo usado e aperfeiçoado desde a segunda metade do século XIX<sup>(2)</sup>. Devido a alta prevalência do câncer e sua elevada mortalidade na população mundial, responsável por 9 milhões de mortes somente no ano de 2016<sup>(3)</sup>, aliado a uma medicina cada vez mais personalizada, o patologista tem encontrado inúmeros desafios. Entre eles, destacam-se, principalmente, o aumento do volume de trabalho, consequência tanto da maior demanda por conta do crescimento populacional, quanto da intensa subespecialização da área em resposta à subespecialização cirúrgica<sup>(4)</sup>. Da mesma forma, os avanços em relação ao conhecimento sobre diferentes

tipos de câncer e a necessidade de maior precisão e celeridade na definição diagnóstica também contribuem para o aumento da carga de trabalho, sendo essenciais os dados fornecidos pelo exame anatomopatológico para a determinação do melhor tratamento disponível<sup>(4)</sup>.

Nesse contexto, o advento de sistemas de escaneamento de lâminas de alta resolução – uma tecnologia disponível em larga escala –, permite a obtenção de lâminas histológicas virtuais totais [*whole slide image* (WSI)] de excelente qualidade e resolução, com ampla utilização em ensino, pesquisa e consultorias remotas, dando origem à microscopia virtual<sup>(5-10)</sup>. Entretanto, ainda não se dispõe igualmente de programas aplicáveis na rotina clínica de análise de tais imagens ou dados de forma automatizada que auxiliem o patologista a acelerar o processo diagnóstico, o qual ainda depende da análise unitária caso a caso por um indivíduo treinado.

O presente estudo aborda uma técnica que vem se popularizando nos últimos sete anos: a aprendizagem profunda, aplicando-a a imagens histopatológicas, e será conduzido de maneira similar à técnica de Litjens *et al.* (2016)<sup>(11)</sup>. Essa técnica foi utilizada para o diagnóstico de câncer de estômago, neoplasia epitelial maligna que afeta, mundialmente, cerca de 990.000 indivíduos a cada ano, levando cerca de 738.000 pacientes a óbito. É o quarto tipo de câncer mais comum em termos de incidência global e a segunda maior causa de mortalidade por câncer<sup>(12)</sup>. No Brasil, o Instituto Nacional do Câncer (INCA) estimou 13.540 novos casos de câncer de estômago entre homens e 7.750 em mulheres para o biênio de 2018-2019; é o quarto tipo mais incidente em homens e o sexto entre as mulheres<sup>(13)</sup>.

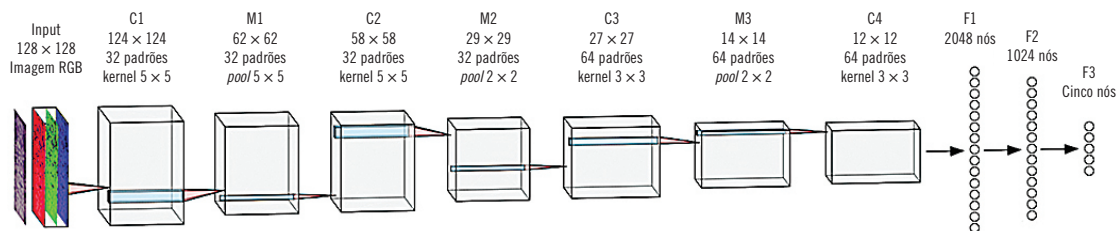
A aprendizagem profunda foi aplicada especificamente para o subtipo adenocarcinoma gástrico. Histologicamente, o diagnóstico desse tumor impõe desafios devido a sua heterogeneidade morfológica, refletida, em parte, na diversidade de esquemas histopatológicos de classificação. A Organização Mundial da Saúde (OMS) adota uma subclassificação histológica estritamente descritiva, a qual reconhece cinco principais tipos de adenocarcinoma gástrico, designados nas categorias: tubular, papilar, mucinoso, pouco coeso (incluindo a variante células em anel de sinete) e misto<sup>(14)</sup>. Outros sistemas classificatórios utilizados pelos patologistas incluem o sistema de Láuren, com as seguintes subclassificações: tipos difuso, intestinal e indeterminado<sup>(15)</sup>, consistindo a descrição microscópica do carcinoma tipo difuso em células pouco coesas com pouca ou nenhuma formação glandular; enquanto o tipo intestinal designa carcinomas com formação glandular com variados graus de diferenciação<sup>(14)</sup>. Há também a classificação de Carneiro, que reconhece quatro categorias: glandular, células isoladas, sólido e misto<sup>(16)</sup>. É preciso citar ainda as lesões pré-malignas, as quais compreendem proliferações epiteliais neoplásicas com atipias celulares e arquiteturais, porém sem evidência de invasão da lâmina própria<sup>(14)</sup>. Uma combinação

dessas classificações foi utilizada no estudo e será descrita na seção metodológica.

A aprendizagem profunda, tecnologia que vem sendo aplicada em diversas áreas do conhecimento<sup>(17-19)</sup>, como dito anteriormente, tem se mostrado uma ferramenta auxiliar promissora na detecção e no diagnóstico histológico de determinados tipos de neoplasias, como adenocarcinoma de próstata e carcinoma mamário<sup>(11)</sup>. Trata-se de uma família de algoritmos que utilizam grandes bases de dados para detectar e aprender a reconhecer padrões relevantes de maneira automática, sem depender de uma extração manual e trabalhosa de dados quantitativos para cada conjunto<sup>(18, 19)</sup>. Divide-se em dois grupos: aprendizado não supervisionado e supervisionado. O primeiro, utilizado neste estudo, é aquele em que, para cada amostra de entrada, existe uma resposta correta que é apresentada ao algoritmo de treinamento.

Para a modelagem, foi utilizada uma técnica específica dentro da família de algoritmos de aprendizado profundo, chamada rede neural convolucional. Seu diferencial é conter uma ou mais camadas convolucionais em sua topologia, indicadas pela letra C na **Figura 1**. O processo de aprendizado consiste na atualização dos pesos das ligações entre os nós das camadas da rede neural para o conjunto de amostras. Uma camada convolucional vai tentar aprender os padrões (*features*) das amostras apresentadas por meio do processo dinâmico de atualização dos seus pesos. O processo é repetido iterativamente, ou seja, se repete diversas vezes para se chegar a um resultado e, a cada vez, gera um resultado parcial que será usado na vez seguinte; o operador do algoritmo estabelece um tempo para que o treinamento seja interrompido. O tempo é denominado época e cada unidade de época corresponde a uma passagem por todas as amostras disponíveis. Na época seguinte, todas as amostras serão novamente revisitadas durante a atualização dos pesos.

Nosso estudo investiga a aplicação da rede neural convolucional na identificação de adenocarcinoma gástrico



**FIGURA 1** – Fluxo de processamento da rede neural convolucional utilizada para a detecção das cinco classes definidas para adenocarcinoma gástrico em lâminas de tecido coradas pelo HE, digitalizadas. As quatro camadas indicadas com C (camadas convolucionais) podem ser consideradas como estágios de extração de feições (*features*), nas quais as características não definidas pelo usuário são consecutivamente extraídas do fragmento de imagem. As camadas indicadas pela letra M são camadas do tipo max-pooling e reduzem o tamanho da imagem, forçando ainda mais iterações entre os fragmentos. As últimas três camadas F são do tipo “classificação” que, com base nas características extraídas, indicam as classes finais de cada amostra

HE: hematoxilina e eosina.

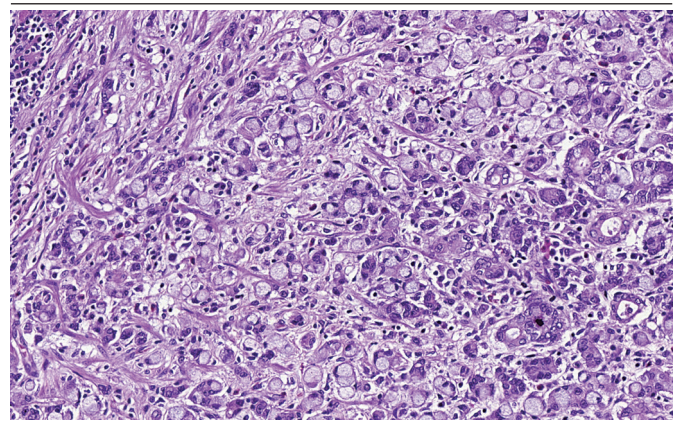
em imagens escaneadas em alta resolução obtidas a partir de lâminas histopatológicas coradas pelo método HE. Investiga, por conseguinte, a possibilidade de associar tal método ao trabalho do patologista que, futuramente, poderá ser útil para enfrentar os desafios da prática atual.

## MATERIAIS

Cinquenta e cinco imagens totais de lâminas histológicas (WSI) provenientes do arquivo do Laboratório de Patologia do Instituto do Câncer do Estado de São Paulo (ICESP), com diagnóstico prévio de adenocarcinoma gástrico, foram obtidas por meio de um digitalizador de lâminas 3D Histech P250, utilizando uma objetiva de 40×, resultando em uma resolução de pixel de 0,19 µm. Após a digitalização, para este estudo inicial, seis lâminas com o mínimo de artefatos de processamento histológico e representativas das diferentes subclassificações de adenocarcinoma gástrico foram selecionadas e visualizadas por meio do *software* Pannoramic Viewer, versão 1.15.4. Destas, 251 áreas representativas foram obtidas, com aumento de 20× (1145 × 707 pixels), correspondentes às diferentes variações morfológicas do adenocarcinoma e também às áreas representativas do epitélio gástrico normal e dos demais tecidos não epiteliais e não neoplásicos, conforme os critérios listados a seguir: 1. tecidos normais não epiteliais (TN) – qualquer área da lâmina onde não há epitélio normal ou neoplásico, contendo ampla representatividade dos tecidos conectivos, musculares, vasculares, adiposo e lâmina própria. Foram selecionadas 50 áreas representativas; 2. epitélio gástrico normal (GN) – epitélio gástrico de tipos fúndico e pilórico sem alterações metaplásicas, displásicas ou neoplásicas. Foram selecionadas 50 áreas representativas; 3. epitélio gástrico neoplásico/adenocarcinoma gástrico tubular (AGT) – epitélio gástrico com displasia moderada e adenocarcinoma gástrico com formação glandular. Foram selecionadas 50 áreas representativas; 4. adenocarcinoma gástrico de tipo sólido (AGS) – adenocarcinoma gástrico de tipo sólido, sem formação de glândulas. Foram selecionadas 41 áreas representativas; 5. adenocarcinoma gástrico de tipo descoeso/difuso (AGD) – adenocarcinoma gástrico com células descoesas e células em anel de sinete. Foram selecionadas 60 áreas representativas.

Essas categorias foram criadas com base nos diferentes sistemas classificatórios do adenocarcinoma gástrico, com o intuito de avaliar o poder discriminatório automatizado entre presença ou não de formação glandular neoplásica e não

neoplásica e presença ou não de células neoplásicas descoesas ou em anel de sinete. Em relação às áreas representativas com tecido neoplásico (designadas AGT, AGS e AGD), 70%, no mínimo, da área total selecionada deveria conter o padrão histológico definido. A **Figura 2** ilustra uma entre as 60 áreas representativas para AGD selecionadas com, no mínimo, 70% da área, contendo AGD. Esse valor foi arbitrário, e o método de categorização por área foi escolhido por ser mais rápido em comparação com o método de separação de estrutura por estrutura por meio da marcação manual das lâminas. Essas áreas representativas deram origem às amostras utilizadas no presente estudo.



**FIGURA 2** – Área representativa AGD selecionada para treinamento da rede neural convolucional com no mínimo 70% da imagem contendo AGD

AGD: adenocarcinoma gástrico de tipo descoeso/difuso.

## MÉTODOS

Após a seleção dessas áreas representativas, elas foram separadas em dois grupos: *training-set* e *testing-set*. Com as imagens do conjunto de *training-set*, procedeu-se o treinamento do algoritmo. A topologia da rede neural convolucional utilizada (Figura 1) é similar à rede neural descrita em Litjens *et al.* (2016)<sup>(11)</sup>, porém há algumas diferenças: a camada F3 é composta por cinco nós, uma vez que a modelagem neste trabalho leva em consideração cinco tipos diferentes de classificações e a utilização de um regularizador do tipo *dropout* entre as camadas F2 e F3. Esse regularizador tem como função impedir que a rede aprenda padrões muito específicos do conjunto de dados apresentado, gerando um modelo com *overfitting*, ou seja, um modelo que se ajusta muito bem ao conjunto de dados observado, mas se mostra pouco eficaz para prever novos resultados. O tensor de entrada dos dados tem a dimensão de uma amostra  $n \times n$  pixels em três camadas correspondentes aos canais

R (*red*), G (*green*) e B (*blue*), onde  $n$  é o número de pixels em linha e coluna. Essa matriz tridimensional de dimensões  $n \times n \times 3$  e suas subsequentes transformações na rede são chamadas de tensores. As camadas C são convolucionais, enquanto as camadas M são do tipo *max-pooling*. As camadas de operação *max-pooling* reduzem efetivamente a área do tensor, dando oportunidade para a próxima camada de convolução aprender um padrão relacionado com uma nova escala da imagem. Uma camada *max-pooling* com tamanho de janela  $2 \times 2$  seleciona um pixel entre quatro para criar um novo tensor de área reduzida. Essa técnica de *pooling* possibilita a criação de topologia de rede com mais camadas, daí o termo conhecido como aprendizagem profunda. A quantidade de *features* é um parâmetro utilizado nas camadas convolucionais e representa a quantidade de padrões diferentes que a rede vai considerar no aprendizado.

O maior consumo de tempo envolvendo redes neurais determina os parâmetros da rede nos quais o aprendizado é melhor de acordo com as métricas escolhidas. Uma rede demasiada complexa para uma determinada modelagem terá resultados insatisfatórios. Outro fator importante para o sucesso do treinamento para uma modelagem com o reconhecimento de padrões em imagens é a quantidade de amostras disponíveis para o aprendizado, pois muitas vezes o número de amostras adequado atinge a ordem de milhões<sup>(19)</sup>.

Na interface entre as camadas F2 e F3, existe ainda a aplicação da função *softmax*, que normaliza as probabilidades de cada classe. O modelo treinado é capaz de receber uma imagem de dimensões  $n \times n \times 3$ , além de entregar um vetor de probabilidades com cinco entradas, uma para cada classe. Esse vetor é normalizado, ou seja, a soma das probabilidades de ocorrência é igual a 1.

As métricas utilizadas para a avaliação do potencial classificatório do algoritmo foram sensibilidade e especificidade. A sensibilidade mede a fração da quantidade de amostras corretamente classificadas na classe escolhida sobre a quantidade total de amostras pertencentes a esta classe. A especificidade mede a fração da quantidade de amostras corretamente classificadas como não pertencentes à classe escolhida sobre a quantidade total de amostras que não pertencem a esta classe. A especificidade relaciona-se com a métrica de falso-negativos de acordo com a equação  $FN = 1 - S$ , onde FN é a taxa de falso-negativos e S é a sensibilidade. A métrica *F1-score* utilizada refere-se a uma média harmônica da sensibilidade e da precisão, correspondendo a uma medida de acurácia do teste, cujos valores variam de 0 a 1, sendo 1 equivalente a precisão e sensibilidade perfeitas. A precisão mede a fração de verdadeiro-positivos sobre a quantidade total de positivos previstos pelo teste (verdadeiro-positivos e falso-positivos). Todos os resultados

das métricas foram calculados a partir do *testing-set* (conjunto de amostras de teste não utilizadas no treinamento).

Por fim, três resultados serão apresentados. A primeira etapa foi determinar o tamanho ideal das amostras a serem geradas a partir das áreas representativas, que serão *input* para a rede neural convolucional. Amostras demasiadamente pequenas não apresentam estruturas que podem ser significativas para a determinação da classe. Por outro lado, amostras grandes demais demandam uma rede neural mais complexa e difícil de parametrizar, além de diminuir sensivelmente a quantidade de amostras disponíveis e, conforme anteriormente comentado, milhares ou até mesmo milhões de amostras são necessárias. O segundo resultado almejado foram as curvas *receiver operator curve* (ROC) satisfatórias para o tamanho de amostra obtida na primeira etapa. Essas curvas são indicativas da capacidade discriminativa do algoritmo. A terceira e última etapa é a classificação das amostras utilizando os melhores parâmetros para a otimização da rede neural convolucional encontrados na etapa anterior.

## RESULTADOS

O tamanho da amostra é um parâmetro relevante no planejamento da modelagem. Para determinar um tamanho ótimo para este estudo, o *F1-score* foi calculado, utilizando uma variedade de dimensões diferentes para as amostras; uma rede neural em cada caso foi treinada. As dimensões (em pixel  $\times$  pixel) utilizadas foram:  $8 \times 8$ ,  $16 \times 16$ ,  $32 \times 32$ ,  $96 \times 96$ ,  $128 \times 128$ ,  $148 \times 148$ ,  $198 \times 198$  e  $256 \times 256$ . A **Figura 3** mostra os resultados do *F1-score* para todas as classes em função do tamanho lateral da amostra. Os resultados mostram que as melhores dimensões foram  $128 \times 128$  e  $148 \times 148$ . Foi escolhido então o tamanho  $128 \times 128$ , pois mesmo havendo uma classe GN com valor maior de *F1-score* para a dimensão  $148 \times 148$  do que para  $128 \times 128$ , as outras classes sofreram penalidade. Utilizando um tamanho de amostra de  $128 \times 128$ , o número total de amostras foi de 10.040, sendo 2.000 para GN, 2.000 para AGT, 1.640 para AGS, 2.400 para AGD e 2.000 para TN.

As curvas ROC, indicativas da capacidade discriminativa do algoritmo, referentes a cinco classes, são apresentadas na **Figura 4**. A classe com menor área sob a curva ROC (AUC) foi a GN – 0,9795 – um valor de performance considerado excelente. Esta conclusão também se estende às outras classes.

Por fim, o terceiro resultado a ser apresentado são as áreas representativas classificadas. Após cada amostra receber um

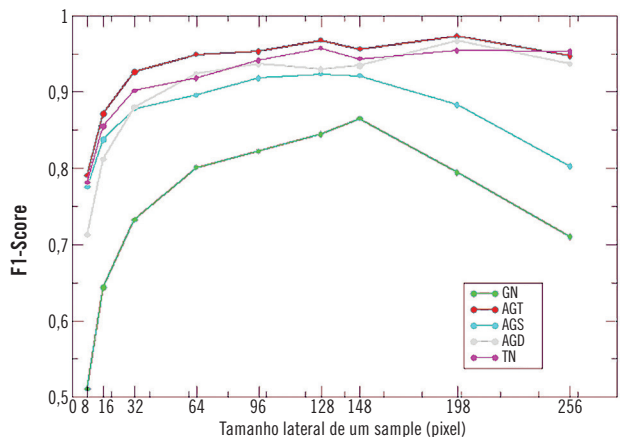


FIGURA 3 – Curvas F1-score para todas as classes em função do tamanho lateral da amostra

GN: epitélio gástrico normal; AGT: epitélio gástrico neoplásico/adenocarcinoma gástrico tubular; AGS: adenocarcinoma gástrico de tipo sólido; AGD: adenocarcinoma gástrico de tipo descoeso/difuso; TN: tecidos normais não epiteliais.

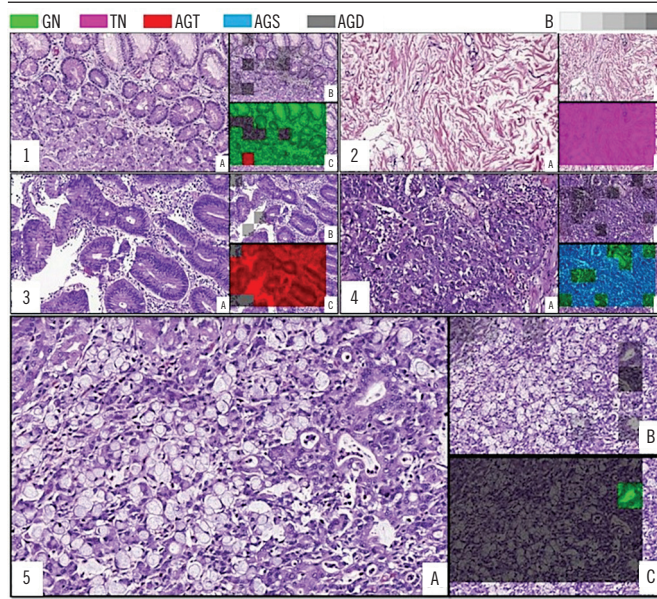


FIGURA 5 – As imagens marcadas de 1 a 5 ilustram os resultados para as cinco diferentes classes. As imagens marcadas pela letra A são as amostras sem classificação; as marcadas com a letra B, as amostras e seus respectivos graus de certeza (conforme a graduação B cinza no código de cores: quanto mais escuro maior a incerteza); e as marcadas com a letra C, as classificações geradas pelo algoritmo, conforme o código de cores na legenda

GN: epitélio gástrico normal; TN: tecidos normais não epiteliais; AGT: epitélio gástrico neoplásico/adenocarcinoma gástrico tubular; AGS: adenocarcinoma gástrico de tipo sólido; AGD: adenocarcinoma gástrico de tipo descoeso/difuso.

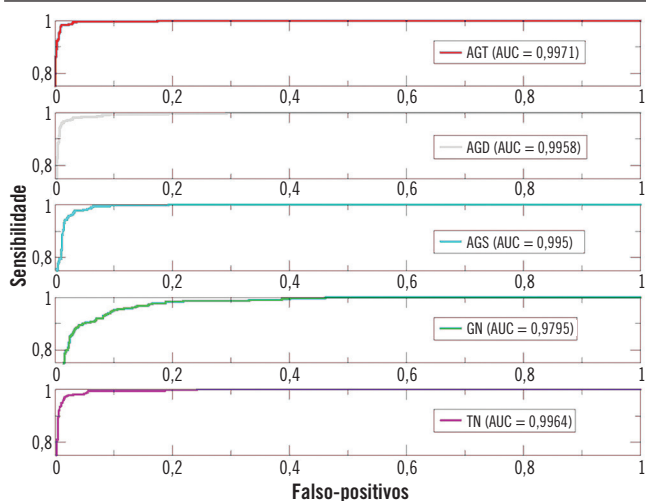


FIGURA 4 – Curvas ROC para as cinco classes definidas e suas respectivas AUC

ROC: receiver operator curve; AUC: área sob a curva ROC; AGT: epitélio gástrico neoplásico/adenocarcinoma gástrico tubular; AGD: adenocarcinoma gástrico de tipo descoeso/difuso; AGS: adenocarcinoma gástrico de tipo sólido; GN: epitélio gástrico normal; TN: tecidos normais não epiteliais.

vetor de probabilidades para cada classe (os cinco vetores de probabilidades somados devem ser igual a 1), o maior vetor será o definidor da classe a que a amostra pertence. É também interessante conhecer o grau de certeza da amostra classificada. A Figura 5 ilustra os resultados para as cinco diferentes classes. As amostras sem classificação são a primeira imagem de cada classe (marcadas pela letra A); a segunda imagem (marcadas com a letra B) apresenta seus graus de certeza; e a terceira imagem de cada classe (marcadas com a letra C), as classificações geradas pelo algoritmo para as amostras do *testing-set*.

## DISCUSSÃO

Neste trabalho, foi estudada a aplicação de redes neurais convolucionais para o reconhecimento de cinco classes presentes em imagens histológicas digitalizadas coradas por HE, representativas de adenocarcinoma gástrico. Os resultados de identificação foram satisfatórios de acordo com as métricas da inspeção visual das curvas ROC e os valores AUC, uma vez que todas as curvas ROC apresentaram AUC acima de 0,97, o que indica uma excelente capacidade discriminativa e classificatória do algoritmo para todas as cinco classes definidas. Comparando esses resultados com outros identificados na literatura<sup>(11)</sup>, três hipóteses não exclusivas foram levantadas para explicar a alta performance das métricas obtidas: 1. características específicas do adenocarcinoma gástrico – o objeto de estudo pode ter um padrão particularmente fácil de identificação em relação à condição normal do tecido, quando analisado para o tamanho ideal de amostra encontrado de 128 × 128 pixels; 2. particularidade das amostras utilizadas, que foram selecionadas buscando representar ao máximo todas as possibilidades morfológicas do adenocarcinoma gástrico, evitando selecionar áreas com muitos artefatos de processamento ou outros possíveis fatores de confusão. Isso implica que as imagens escolhidas podem ter tornado a modelagem muito específica para esse conjunto de casos; 3. característica da metodologia de definição das imagens para amostragem. Para a definição das

amostras de aprendizado, foram pesquisadas imagens com pelo menos 70% da área de interesse correspondente a somente uma das classificações histológicas descritas, evitando capturar áreas com combinações das diferentes classes. É um processo diferente da marcação das classes em uma lâmina total digitalizada e da posterior etapa de pós-processamento, que define as amostras e o vetor de probabilidades a serem utilizados no treinamento. De acordo com Litjens *et al.* (2016)<sup>(11)</sup>, essas duas definições não são equivalentes em relação à modelagem. Esse método de definição das amostras pode então trazer bons resultados usando como teste amostras extraídas com o mesmo método. Isso não quer dizer, necessariamente, que essa modelagem terá os bons resultados para uma previsão em uma imagem de lâmina total (WSI).

Uma generalização mais abrangente da identificação necessita de um número maior de dados para treino com o máximo possível de diversidade. Sabe-se que, na confecção de uma simples lâmina histopatológica de HE, inúmeros fatores influenciam o resultado final do corte histológico em análise, como a variação de intensidade das colorações utilizadas, os artefatos de processamento e o corte, que podem produzir espaços vazios entre os tecidos, as dobras e as sobreposições, além do sentido longitudinal ou transversal do corte em relação à amostra de tecido, que pode gerar padrões de imagens diferentes, facilmente reconhecidas pelo conjunto olho-cérebro humano. Contudo, é um desafio para o algoritmo classificatório se essas variações não forem previamente incluídas no treinamento. Além desses desafios técnicos que geram uma diversidade de padrões, a própria natureza morfológica do adenocarcinoma gástrico é bastante heterogênea, com a possível presença de diferentes padrões na mesma lesão, tornando-se um desafio classificatório também para o patologista treinado<sup>(16)</sup>. Tal limitação se estende para a definição e a seleção das amostras e, conseqüentemente, para o sistema classificatório do algoritmo, predominantemente nos casos de lesões que, quando analisadas em uma lâmina total, se encaixem no subtipo misto. Assim, as atuais definições classificatórias na seleção de amostras deste estudo poderão ser revistas em um estudo posterior, a ser aplicado não apenas em amostras representativas, mas em uma imagem de lâmina total (WSI).

É razoável esperar que um conjunto maior de dados para o treinamento diminua a performance do modelo em relação às métricas apresentadas, porém a modelagem terá maior poder de generalização quanto aos dados que podem ser apresentados.

Dentro do treinamento realizado, verificamos ainda uma robustez dos modelos de acordo com alguns casos nos quais uma amostra é informada ao treinamento como pertencente a uma determinada classe e a predição desse modelo informa outra classe, classificando-a de maneira correta, como pode ser observado na Figura 4. Nela, visualizamos que dentro de uma imagem com classificação predominante de padrão AGD (Figura 4 – 5A), o algoritmo identificou uma área com glândula normal, classificando-a corretamente como GN (Figura 4 – 5C). Isso acontece pois as amostras (128 × 128 pixels) são obtidas a partir de uma área maior de uma imagem representativa (1145 × 707 pixels), na qual somente a classe predominante da imagem maior é informada como presente no treinamento do algoritmo.

## CONCLUSÃO

Este estudo preliminar demonstrou que, para um tamanho de amostra definido de 128 pixels, o algoritmo é capaz de identificar de maneira satisfatória as diferentes estruturas relevantes na imagem de adenocarcinoma gástrico para a classificação dentro das cinco classes determinadas. Ainda mostra que a aplicação de redes neurais convolucionais para a classificação de tecidos em lâminas anatomopatológicas digitalizadas é promissora.

Um conjunto maior de amostras categorizadas e modeladas por um grupo de profissionais como neste trabalho pode apresentar uma predição sem enviesamento individual na seleção das amostras representativas, bem como definir critérios mais estritos para sua inclusão, a fim de gerar uma classificação ainda mais robusta. O próximo passo para este estudo é utilizar uma quantidade maior de amostras e alterar o método de definição delas, marcando as regiões com classes diferentes na imagem total da lâmina (WSI)<sup>(11)</sup>.

## AGRADECIMENTOS

Ao professor doutor Evandro Sobroza de Mello, coordenador do Serviço de Patologia do ICESP, pelo fornecimento das imagens digitalizadas das lâminas histológicas, possibilitando a realização deste trabalho.

## REFERÊNCIAS

1. Fischer AH, Jacobson KA, Rose J, Zeller R. Hematoxylin and eosin staining of tissue and cell sections. *CSH Protocols*; 2008.
2. Van Den Tweel JG, Taylor Cr. A brief history of pathology: preface to a forthcoming series that highlights milestones in the evolution of pathology as a discipline. *Virchows Arch*. 2010; 457(1): 3-10.
3. World Health Statistics 2018: monitoring health for the SDGs. Disponível em: <https://www.who.int/news-room/fact-sheets/detail/cancer> acessado em 01/01/2020.
4. Kamel HM. Trends and challenges in pathology practice: choices and necessities. *Sultan Qaboos Univ Med J*. 2011; 11(1): 38-44.
5. Bashshur RL, Krupinski EA, Weinstein RS, Dunn MR, Bashshur N. The empirical foundations of telepathology: evidence of feasibility and intermediate effects. *Telemed J E Health*. 2017; 23(3): 155-91.

6. Kayser K, Borkenfeld S, Kayser G. How to introduce virtual microscopy (VM) in routine diagnostic pathology: constraints, ideas, and solutions. *Anal Cell Pathol (Amst)*. 2011; 35(1): 3-10.
7. Wienert S, Beil M, Saeger K, Hufnagl P, Schrader T. Integration and acceleration of virtual microscopy as the key to successful implementation into the routine diagnostic process. *Diagn Pathol*. 2009; 4: 3.
8. Kayser K. Introduction of virtual microscopy in routine surgical pathology-a hypothesis and personal view from Europe. *Diagn Pathol*. 2012; 7: 48.
9. Jara-Lazaro AR, Thamboo TP, Teh M, Tan PH. Digital pathology: exploring its applications in diagnostic surgical pathology practice. *Pathology*. 2010; 42(6): 512-8.
10. Mori I, Ozaki T, Muragaki Y, et al. Construction of web-based remote diagnosis system using virtual slide for routine pathology slides of the rural hospital in Japan. *Diagn Pathol*. 2013; 8(Suppl 1): S4.
11. Litjens G, Sánchez CI, Timofeeva N, et al. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. *Sci Rep*. 2016; 6(1): 262-86.
12. Karimi P, Islami F, Anandasabapathy S, Freedman ND, Kamangar F. Gastric cancer: descriptive epidemiology, risk factors, screening, and prevention. *Cancer Epidemiol Biomarkers Prev*. 2014; 23(5): 700-13.
13. Ministério da Saúde. Instituto Nacional do Câncer. Estimativa 2018: incidência de câncer no Brasil. Instituto Nacional de Câncer José Alencar Gomes da Silva. Coordenação de Prevenção e Vigilância. Rio de Janeiro: INCA; 2017.
14. Carneiro F, Curado MP, Franceschi S, et al. Gastric carcinoma. In: Bosman FT, Carneiro F, Hruban RH, Theise N, editores. *WHO classification of tumours of the digestive system*. 4 edição. France: IARC; 2010. p. 48-58.
15. Schaeffer DF, Owen DA. Stomach. In: Sternberg SS, Mills SE, Carter D, editores. *Sternberg's diagnostic surgical pathology*. 6 edição. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2015. p. 1409-46.
16. Carneiro F. Classification of gastric carcinomas. *Curr Diagn Pathol*. 1997; 4(1): 51-9.
17. Hastie T, Tibshirani R, Friedman J. *The elements of statistical learning: data mining, inference, and prediction*. Nova Iorque: Springer-Verlag; 2009.
18. Deo RC. Machine learning in medicine. *Circulation*. 2015; 132(20): 1920-30.
19. Cruz-Roa A. *Data-driven representation learning from histopathology image databases to support digital pathology analysis [tese]*. Universidad Nacional de Colombia; 2015.

#### AUTOR CORRESPONDENTE

---

Jonas Kloeckner  0000-0002-9947-7081  
e-mail: jonaskloeckner@hotmail.com



This is an open-access article distributed under the terms of the Creative Commons Attribution License.