

## ***Deep learning* como sistema de auxílio diagnóstico e classificação do câncer de mama**

**Daniel G Tiezzi<sup>1,2</sup>, Lucas B Figueira<sup>2</sup>, Rodrigo Plotze<sup>3</sup>**

<sup>1,2,3</sup>Faculdade de Tecnologia de Ribeirão Preto (FATEC)  
Ribeirão Preto, SP - Brasil

<sup>2</sup>Faculdade de Medicina de Ribeirão Preto – Departamento de Ginecologia e Obstetrícia  
– Setor de Oncologia Ginecológica e Mastologia – Universidade de São Paulo –  
Ribeirão Preto, SP – Brasil

<sup>1</sup>dtiezzi@usp.br, <sup>2</sup>lucas.figueira@fatec.sp.gov.br,  
<sup>3</sup>rodrigoplotze@gmail.com

**Resumo.** *A classificação histológica do cancer de mama é baseada no grau de diferenciação. Vários estudos demonstraram uma baixa concordância intra e inter observador. Redes neurais convolucionais como método de aprendizado de máquina têm sido utilizadas com sucesso na classificação de imagens médicas, em especial na área da Radiologia. Estudos recentes demonstraram que a aplicação deste método de inteligência artificial é útil na classificação histológica de neoplasias malignas. Este projeto tem como objetivo desenvolver um sistema automatizado de auxílio à classificação histológica do câncer de mama utilizando redes neurais artificiais.*

**Abstract.** *The histological classification of breast cancer is based on the degree of tissue differentiation. Several studies have shown the intra- and inter-observer agreement is low. Convolutional neural networks as a machine learning method have been used successfully in the classification of medical images, especially in the area of Radiology. Recent studies have shown that the application of this artificial intelligence method is useful in the histological classification of malignant neoplasms. This project aims to develop a decision support system for histological classification of breast cancer using artificial neural networks.*

### **1. Introdução**

O câncer de mama é a neoplasia maligna que mais frequentemente acomete a população feminina em todo o mundo. No Brasil não é diferente e a estimativa para 2020 é da ordem de 60 mil casos novos. Em números totais, também é a principal causa de morte por doença maligna em nossa população (INCA, 2018). A morte pelo câncer de mama é quase que exclusivamente decorrente da disseminação metastática da doença. O carcinoma da mama é uma doença heterogênea, e os fatores determinantes da disseminação de células neoplásicas para órgãos distantes ainda não são completamente conhecidos (CIRIELLO

et al., 2015).

Uma série de fatores têm sido descritos como métodos de predição de prognóstico e são atualmente utilizados na prática clínica para definição de tratamento, em especial na definição de tratamento com drogas antineoplásicas (CURIGLIANO et al., 2017). Dentre esses fatores se destacam critérios clínicos, histológicos e a utilização de marcadores tumorais. O estadiamento da doença, em especial o estado axilar, é um dos principais determinantes de prognóstico no câncer de mama. Do ponto de vista histológico, o grau de diferenciação tumoral, que é baseado no sistema de escore de Nottingham (NGS), é descrito por alguns autores como um potente método de predição, principalmente quando associado ao status axilar (WARWICK et al., 2004). Embora o NGS seja tão importante para a definição de tratamento e como um fator de prognóstico, ele vem recebendo críticas associadas à sua baixa reprodutibilidade.

A reprodutibilidade do método de classificação do grau histológico em microscopia de luz foi estudada em várias ocasiões. Os estudos demonstram que, de uma forma geral, existe uma baixa concordância inter e intra observador (RAKHA et al., 2010). O motivo da baixa reprodutibilidade parece ser decorrente da subjetividade do método e do processamento pré-analítico da amostra. A grande repercussão da falha de classificação é o seu impacto na predição do prognóstico, que influencia diretamente na decisão clínica em administrar ou não quimioterapia sistêmica para o paciente. O desvio da acurácia em predizer o risco de recorrência da doença, utilizando parâmetros clínicos e histopatológicos, tem incentivado a investigação e desenvolvimento de métodos de biologia molecular para inferir uma maior ou menor agressividade da doença.

Uma grande limitação aos métodos atualmente disponíveis de caracterização molecular do câncer de mama é que são produtos comerciais e de alto custo, o que inviabiliza a sua utilização, em especial em países de baixo poder econômico e nos sistemas públicos de saúde.

Os métodos de aprendizado de máquina vêm sendo aplicados em larga escala em diferentes áreas do conhecimento. Na Medicina, hoje é um importante método de auxílio diagnóstico de imagens radiológicas (HU et al., 2018). Estudos recentes têm demonstrado que redes neurais convolucionais, uma metodologia computacional conhecida como *deep learning*, pode ser útil na classificação de imagens histológicas. Redes neurais

convolucionais (CNN, do inglês *Convolutional Neural Network*) podem ser utilizadas para classificação de imagens, identificação de objetos / padrões na imagem e como forma de segmentação automatizada da imagem de forma eficiente (COUDRAY et al., 2018).

### **Justificativa**

Investir em um modelo de inteligência artificial que possa auxiliar na classificação de carcinomas de mama de forma eficiente pode ser determinante para melhorias na assistência à portadores da doença com otimização do tratamento e redução de efeitos adversos. Além de ser uma ferramenta com alto potencial de aplicação clínica, ela pode ser desenvolvida com custo relativamente baixo e, após o treinamento do modelo, que exige alto poder computacional, o modelo é altamente portátil e pode ser implementado em sistemas móveis que podem ser utilizados por profissionais nas mais remotas áreas do globo terrestre.

### **Objetivos**

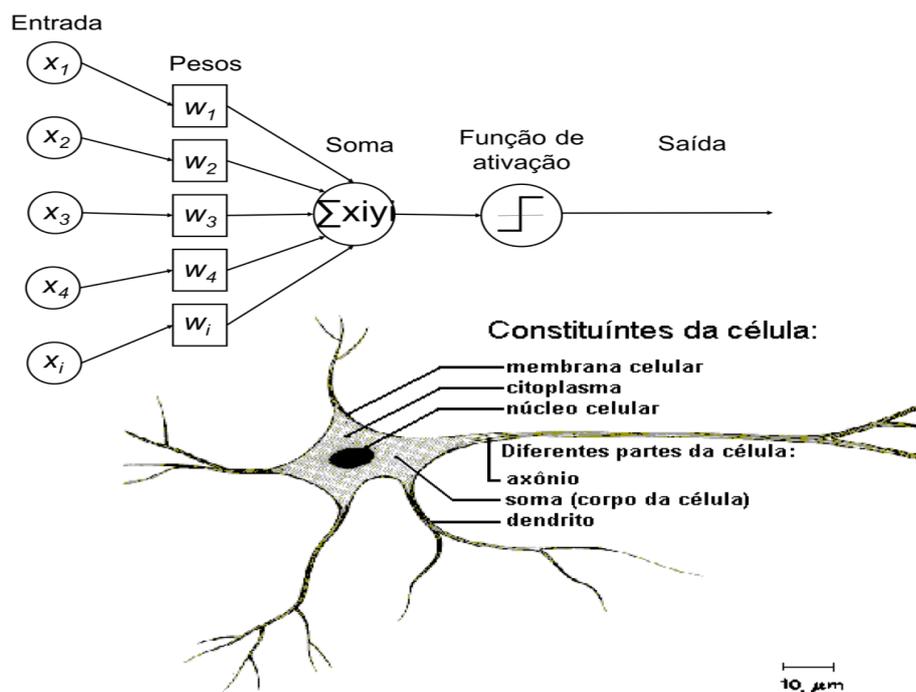
O objetivo do estudo é a utilização de métodos computacionais de aprendizado de máquina para o desenvolvimento de tecnologia de auxílio a classificação histológica do câncer de mama.

#### **Objetivos específicos:**

- a) Desenvolvimento de sistema computacional para seleção de áreas de interesse (ROI, do inglês *Regions Of interest*) da imagem histológica;
- b) Criação de um banco de imagens de câncer de mama que possa ser utilizado como base de treinamento para uma CNN;
- c) Utilização de métodos computacionais para segmentação nuclear das imagens histológicas;
- d) Desenvolvimento de uma CNN para treinamento e aprendizado de máquina como método de classificação histológica do câncer de mama;
- e) Análise da acurácia do CNN na predição de grau nuclear e histológico no câncer de mama;
- f) Impacto na sobrevida da classificação histológica do câncer de mama por inteligência artificial.

## 2. Revisão da Literatura

As redes neurais têm sua origem em um algoritmo idealizado por Frank Rosenblatt, em 1958, denominado de perceptron. O perceptron foi baseado na operação de uma unidade de processamento proposta por McCulloch e Pitts, em 1943 (ROSENBLATT 1958; MCCULLOCK e PITTS 1943). Ele é um modelo de classificação binária que tem como base a estrutura de um neurônio. Ele recebe os valores de entrada como se fossem os sinais recebidos pelos dendritos de um neurônio. No “corpo celular”, o algoritmo computa a soma do produto dos pesos com os valores de entrada e aplica uma função de ativação para fazer a previsão. A função tem uma saída binária (o neurônio foi ativado ou não) que é transmitida pelo axônio. Na Figura 1 está apresentada a estrutura de um perceptron, simulando a estrutura de um neurônio.



**Figura 1. Modelo esquemático do perceptron e sua similaridade com a estrutura de um neurônio**

Fonte: <http://conteudo.icmc.usp.br/pessoas/andre/research/neural/>

A função de ativação do perceptron determina que os valores de entrada  $x_i$ , modulados pelos pesos  $w_i$ , serão mapeados para uma saída  $f(x)$  ou  $y_j$ , sendo  $j$  o número de elementos em sua amostra. A função de ativação pode ser definida por:

$$f(x) = 1 \text{ se } w \cdot x + b > 0; \text{ ou}$$

$$f(x) = 0$$

em que  $w \cdot x$  é o produto escalar de  $\sum_{i=1}^m w_i \cdot x_i$ ;  $m$  = número de entradas  $x$

A base para o aprendizado desta classificação binária é a atualização dos pesos  $w_i$  em cada uma das épocas (cada vez que o classificador recebe os valores e calcula a saída de todos os elementos do ensaio). Ao final de cada época, podemos calcular o erro, também denominado de função de perda, que é definido pela subtração do valor de saída pelo valor real ( $d_j - y_j$ ). Ou seja, se o erro for igual a 0, a classificação está correta e os pesos não precisam ser atualizados. Caso contrário, é necessário atualizar os pesos baseados no erro e completar outra época. Este cálculo matemático é conhecido como a regra *delta*, e é definido como:

$$w : w_i(t + 1) = w_i(t) + \alpha(d_j - y_j)x_{ji}$$

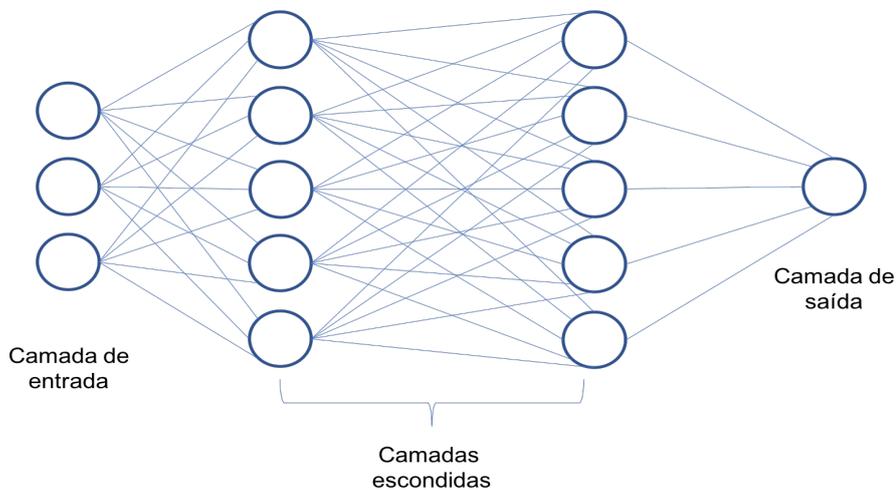
em que  $0 \leq i \leq n$  e  $\alpha$  é a taxa de aprendizado.

Este processo deve ser repetido quantas vezes forem necessárias até que os pesos tenham valores ótimos para minimizar ao máximo o erro. O valor de  $\alpha$  é a magnitude do ajuste dos pesos. Um valor pequeno de  $\alpha$  determina que os ajustes sejam feitos por pequenos “passos” e, se for muito pequeno, pode fazer com que o modelo demore uma eternidade para ajustar os pesos ideais. Por outro lado, um alto valor de  $\alpha$  pode levar a um sobre ajuste dos pesos e o modelo nunca atingirá uma predição ótima.

A utilização deste algoritmo permite a classificação precisa de objetos linearmente separáveis. Na vida real, a maioria dos dados amostrais não possuem esta característica. Desta forma, a utilização de uma rede complexa de neurônios, que chamamos de rede neural, foi elaborada para solucionar problemas de classificação de amostras não linearmente separáveis. O modelo do perceptron pode ser considerado a unidade funcional das redes neurais artificiais.

A primeira rede neural desenvolvida foi o perceptron multicamada (MLP, do inglês *MultiLayer Perceptron*). O modelo foi uma evolução do perceptron e permitiu a

resolução de problemas não linearmente separáveis (RUMELHART, DE et al., 1986). Esta nova arquitetura é composta por mais de uma camada de neurônios: (i) a camada de entrada; (ii) pelo menos uma camada intermediária ou escondida; (iii) a camada de saída. As camadas são conectadas em modo sequencial. A camada de entrada recebe os valores da amostra, a camada intermediária é responsável pela maior parte do processamento e onde é realizada a extração de características da amostra e a camada de saída apresenta a conclusão do resultado final. Na Figura 2 é apresentada um exemplo arquitetura de um perceptron multicamadas.



**Figura 2. Arquitetura de uma rede neural artificial (MLP). Cada uma das camadas escondidas representam o modelo de um perceptron, o que cria a arquitetura denominada de perceptron de múltiplas camadas. Este modelo é o que denominamos de rede neural completamente conectada (ou do inglês fully connected). Note que cada neurônio é conectado com todos os outros neurônios das camadas sequenciais. QUE NOME PARA LEGENDA GRANDE!!!!**

A grande evolução do perceptron para o MLP é o mecanismo de aprendizado, denominado de *Backpropagation* ou de propagação retrógrada. Como no MLP, cada camada possui um peso específico para cada neurônio. O ajuste do erro estimado, definido por  $\alpha(dj - yj)$ , deve ser aplicado a todos os pesos de todas as camadas. Desta forma, o sistema de propagação retrógrada, no qual os pesos de uma camada são ajustados conforme o erro da camada que está a sua frente, foi implementado para ajuste da modelo multicamada. Daí o nome de *Backpropagation*. Este algoritmo é o pilar que permite o aprendizado das redes neurais modernas.

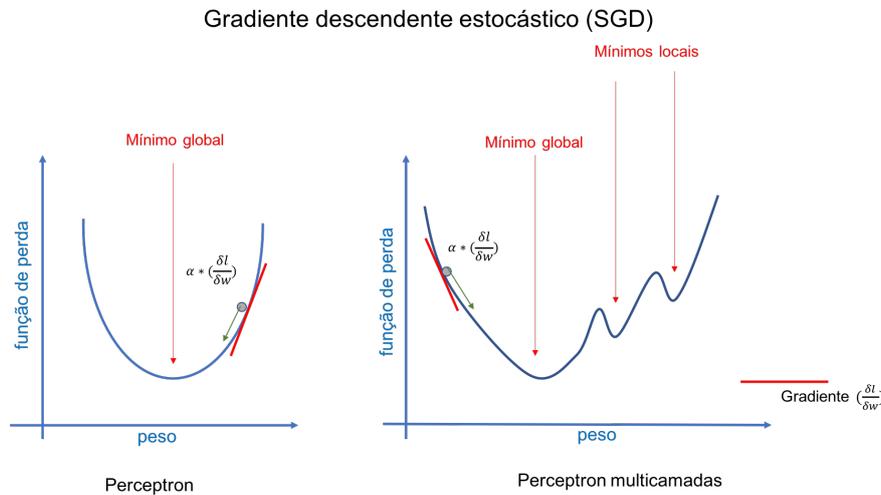
O algoritmo de *Backpropagation* consiste em duas fases: (i) a fase de ida (*forward pass*), na qual os valores de entrada passam por toda a rede e a predição final é obtida;

(ii) a fase de volta (*backward pass*), na qual o gradiente da função de perda é computado na camada final e é recursivamente aplicado para atualizar os pesos de toda a rede.

O gradiente da função de perda neste modelo de multicamadas nada mais é que uma variação da regra delta aplicada ao perceptron: uma regra delta generalizada. No caso do perceptron, a superfície de erro tem o formato de uma parábola com apenas um valor mínimo. No caso da função generalizada, a superfície de erro pode ser bastante irregular e a função pode estar relacionada à vários valores mínimos locais.

As redes neurais mais complexas utilizam modelos mais aperfeiçoados para minimizar este problema. O gradiente da função, que é a derivada da função em um determinado ponto, é utilizado para ajustar os valores dos pesos. Na prática, o gradiente da função de perda é computado utilizando parte das amostras (um lote) que denominamos de *mini-batch* (do inglês, mini-lote). Este método é denominado de *mini-batch gradient descent*, também conhecido como gradiente descendente estocástico (SGD, do inglês *Stochastic Gradient Descent*). O gradiente descendente é a derivada parcial da função de perda pelo peso (Figura 3).

Este conhecimento do aprendizado de máquina, utilizando redes neurais, é a base para a implementação de redes neurais convolucionais (CNN, do inglês *Convolutional Neural Networks*). As CNNs podem ser aplicadas como método de classificação da imagem sem que seja necessário realizar a extração de atributos da imagem. No MLP, cada neurônio é conectado a todos os neurônios da camada seguinte (camadas *fully connected* ou FC). CNNs utilizam inicialmente camadas de convolução para extração de atributos da imagem e a camada FC é aplicada no final da rede neural para inferência da classificação final.



**Figura 3. Representação esquemática do gradiente descendente estocástico. Note que a derivada parcial orienta o sentido para a minimização da perda.**

A convolução da imagem é a multiplicação por elementos de duas matrizes, seguido da soma de todos os elementos da correlação cruzada (*cross correlation*) de uma sub imagem  $w(x, y)$  (denominada *Kernel* ou máscara) de tamanho  $K \times L$  sobre uma imagem  $f(x, y)$  de tamanho  $M \times N$ , em que  $K \leq M$  e  $L \leq N$ . A correlação entre  $w(x, y)$  e  $f(x, y)$  em um ponto  $(i, j)$  pode ser definida pela fórmula:

$$C(i, j) = \sum_{x=0}^{L-1} \sum_{y=0}^{K-1} w(x, y)f(x + i, y + i)$$

em que  $i = 0, 1, \dots, M - 1$ ;  $j = 0, 1, \dots, N - 1$ , e a soma é relacionada a área da imagem onde  $w$  sobrepõe  $f$ .

A convolução pode utilizar diferentes padrões de *kernels* para gerar filtros lineares, não lineares ou morfológicos na imagem. Esses filtros realçam características específicas da imagem, como em um realce não direcional de bordas. Desta forma, a aplicação de convoluções utilizando uma variedade de *kernels* amplia a dimensão da imagem e realça atributos em cada um dos novos canais da imagem. CNNs utilizam o padrão de realces decorrentes de cada filtro em camadas de nível inferior para detectar objetos de alto nível na imagem nas camadas mais profundas da rede.

Esta metodologia permite que o modelo extraia padrões que estão distribuídos na imagem e é a base do reconhecimento de padrões de imagem na atualidade. Esta aplicação

como método de classificação tem se mostrado uma ferramenta útil ao auxílio diagnóstico na interpretação de imagens médicas. Este modelo pode auxiliar patologistas menos experientes e pode ser utilizado como uma ferramenta de treinamento e ensino de novos profissionais.

### **3. Métodos**

#### **3.1 Base de dados**

Iremos utilizar dois grandes bancos de dados publicamente disponíveis: TCGA (<https://portal.gdc.cancer.gov/projects/TCGA-BRCA>) e METABRIC (<https://ega-archive.org/datasets/EGAD00010000270>) e um banco de dados de pacientes com câncer de mama tratadas entre 2000 a 2013 no Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto – USP (HCRP). O projeto para levantamento de dados clínicos e aquisição de imagem do coorte HCRP já foi previamente aprovado pelo comitê de ética local (2.638.453/2018). O TCGA possui um acervo com informações clínicas e histopatológicas de 1098 pacientes com câncer de mama, incluído a imagem histológica de todas elas. O METABRIC é um banco de dados com informações clínicas e histopatológicas de 1992 casos de câncer de mama. As imagens histológicas disponíveis são 564. Os dados do TCGA já foram adquiridos e estão sendo utilizados para o desenvolvimento inicial do sistema. Os dados clínicos de 1967 pacientes do coorte HCRP já foram levantados e pretendemos adquirir digitalmente as imagens histológicas no futuro próximo. As imagens histológicas do METABRIC e HCRP serão utilizadas para validação do modelo e impacto na predição de sobrevida.

#### **3.2 Desenvolvimento do sistema computacional**

O sistema computacional será composto por quatro módulos: (i) módulo de leitura de imagens digitais; (ii) módulo de pré-processamento, normalização e seleção de imagens; (iii) módulo de segmentação nuclear; (vi) módulo de treinamento de máquina. Todo o sistema foi desenvolvido em linguagem de programação *Python* e a perspectiva é oferecer a ferramenta em um sistema web:

- i) Módulo de leitura de imagens digitais:* as imagens dos bancos de dados TCGA e METABRIC foram adquiridas com o scanner Aperio®, o qual gera um arquivo SVS que contém quatro imagens em formato TIFF com diferentes tamanhos. O

o sistema utilizará a biblioteca *OpenSlide* para leitura das imagens, e a biblioteca *OpenCV* será utilizada para visualização da imagem e determinação do ROI corresponde a região tumoral de interesse. O sistema de pré-processamento e normalização fará a leitura das imagens em PNG do banco de dados, permitindo que o usuário determine modificações na dimensão da imagem. Um módulo de normalização da imagem será implementado. Ele utilizará uma ferramenta desenvolvida por Macenko M et al. (2009) que faz a normalização da intensidade de coloração das lâminas.

ii) **Módulo de segmentação nuclear:** as estratégias de segmentação nuclear serão implementadas utilizando *Mask R-CNN* após a construção do banco de imagens definitivo, com a exclusão de imagens de baixa qualidade;

iii) **Módulo de treinamento de máquina:** este módulo é a base do aprendizado da máquina. Ele será implementado após a segmentação nuclear. Pretendemos testar diferentes arquiteturas clássicas incluindo a AlexNet, VGGnet, Lenet e a GoogLeNet. Toda a CNN será desenvolvida com a biblioteca *Keras*, utilizando o *Tensorflow* como ferramenta de aprendizado.

Após o processo de treinamento e estabelecido o modelo de predição, iremos fazer a aplicação e testar o modelo de classificação histológica como preditor de sobrevida no câncer de mama. Esta fase do projeto será desenvolvida em conjunto com os colaboradores da Universidade de Cambridge no Reino Unido.

Toda a fase de treinamento da rede neural será realizada com o banco de dados do TCGA e o banco de dados de imagens do METABRIC será utilizado para testar a classificação e analisar o impacto dela na predição de recorrência e morte pela doença. O METABRIC é um consórcio entre a Universidade de Cambridge e a British Columbia Cancer Agency, em Vancouver, Canadá. As pacientes deste estudo têm um seguimento de longo prazo e todas foram submetidas a um tratamento padronizado e a classificação histológica foi realizada nos laboratórios das duas instituições, o que torna a análise de sobrevida e a classificação histológica bastante consistente. Iremos aplicar o modelo treinado com as imagens no banco de dados do TCGA para a classificação histológica das amostras do estudo METABRIC e HCRP. Posteriormente, iremos analisar o impacto da classificação automatizada na sobrevida, em comparação com a classificação tradicional.

#### 4. Resultados

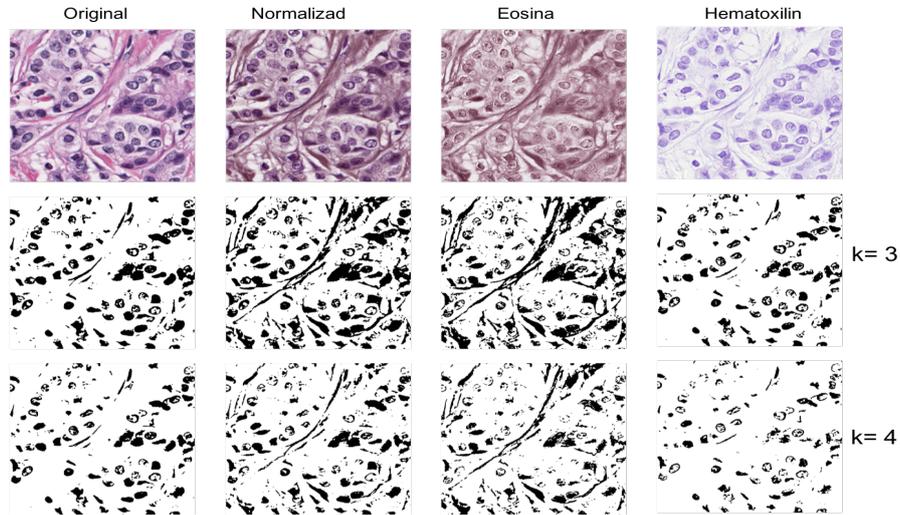
Após a seleção automática dos ROIs, identificamos que uma proporção substancial das sub imagens (cerca de 30%) não apresentam amostragem de células tumorais suficientes ou representam áreas de tecido normal, áreas fora de foco ou artefatos da imagem. O total de sub imagens geradas é da ordem de milhões de imagens. Assim, antes da aplicação do treinamento final, se faz necessário classificar as imagens de acordo com a qualidade. Para tal fim, estamos desenvolvendo um modelo de classificação binário (imagem com qualidade boa x ruim) utilizando uma CNN aplicada após uma filtragem de imagens classificadas de forma não supervisionada baseada no algoritmo K-means. No K-means, cada pixel da imagem recebe uma classificação conforme a sua similaridade com os outros pixels. Desta forma, os pixels de tonalidade azul da hematoxilina ficam todos com a mesma classificação. Com isto, é possível criar uma máscara e converter todos os outros pixels para branco. Baseado na proporção de pixels não brancos, podemos inferir qual é a representatividade de células neoplásicas em cada imagem.

Seguindo a metodologia de normalização de imagens em H&E, é possível separar cada um dos canais H e E. Com esta abordagem, nós pudemos extrair as características da imagem mais representativas do núcleo celular (o canal H) após a subtração da coloração da eosina. O algoritmo de clusterização não supervisionado K-means foi aplicado a imagens originais, imagens normalizadas e na imagem com subtração da eosina (canal H). Este algoritmo realiza a clusterização de cada pixel baseado em similaridade e no número de clusters (Ks). Nós aplicamos o método com  $k = 3$  e  $k = 4$  (Figura 4).

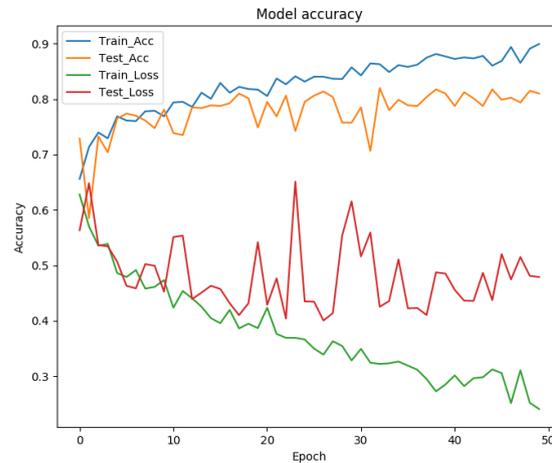
O segundo sistema de filtragem está em desenvolvimento. Um total de 3597 imagens foram classificadas como de qualidade boa ou ruim por um Médico. Estamos testando a performance de CNNs com diferentes profundidades. As imagens são separadas em subgrupos para treinamento e teste (40% para teste) e as imagens foram reduzidas para  $256^2 \times 128^2$  pixels. O sistema utiliza a classe ImageDataGenerator() para aumento de dados e aplicamos como padrão um treinamento com 50 e 100 épocas.

Diferentes funções de otimização foram testadas e a de melhor performance foi o Adam com uma taxa de aprendizado inicial de 0.001. Após alguns ajustes nas camadas de *Dropout*, o modelo atinge um acurácia de cerca de 80%. No entanto, ainda existe um

evidente *overfitting* que pode ser observado em torno da 20 época (Figura 5).



**Figura 4.** Segmentação das utilizando o algoritmo K-means com  $k=3$  e  $k=4$ . Note que, visualmente, a segmentação do núcleo celular é mais precisa quando aplicada à imagem com subtração da eosina (canal H).



**Figura 5.** Acurácia e erro durante o treinamento da CNN para prever imagens com qualidade boa e ruim. Note que por volta da 20ª época a acurácia do modelo no grupo em teste continua a subir e na validação ela estabiliza. A curva de erro também não acompanha o declínio.

## 5. Conclusões

Baseado nos resultados parciais obtidos podemos afirmar que (i) a utilização de métodos de normalização da imagem seguido de técnicas de segmentação por agrupamento

permite a identificação de ROIs e filtragem de imagens com baixa representação de células tumorais; (ii) a utilização de redes convolucionais para seleção de imagens histológicas de baixa qualidade é viável; (iii) existe a necessidade de sistemas de processamento de dados mais robustos para a otimização e treinamento de imagens histológicas.

## REFERÊNCIAS

- CIRIELLO, G.; GATZA, M. L.; BECK, A. H. et al. Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. **Cell**, v. 163, n. 2, 2015.
- COUDRAY, N.; OCAMPO, P. S.; SAKELLAROPOULOS, T. et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. **Nature Medicine**, v. 24, n. 10, p. 1559–1567, 2018.
- CURIGLIANO, G.; BURSTEIN, H. J.; WINER, E. P. et al. De-escalating and escalating treatments for early-stage breast cancer: the St. Gallen International Expert Consensus Conference on the Primary Therapy of Early Breast Cancer 2017. **Annals of Oncology**, v. 28, p. 1700–1712, 2017.
- HU, Z.; TANG, J.; WANG, Z. et al. Deep learning for image-based cancer detection and diagnosis – A survey. **Pattern Recognition**, v. 83, p. 134–149, 2018.
- INCA. **Estimativa 2018: incidência de câncer no Brasil**. 2018.
- MACENKO M.; NIETHAMMER M.; MARRON J. S.; BORLAND D.; WOOSLEY J. T.; XIAOJUN G.; SCHMITT C.; THOMAS N. E. A method for normalizing histology slides for quantitative analysis. IEEE ISBI, 2009.  
[dx.doi.org/10.1109/ISBI.2009.5193250](https://doi.org/10.1109/ISBI.2009.5193250)
- RAKHA, E. A.; REIS-FILHO, J. S.; BAEHNER, F. et al. Breast cancer prognostic classification in the molecular era: The role of histological grade. **Breast Cancer Research**, 2010.
- ROSENBLATT, F. The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain, Cornell Aeronautical Laboratory, Psychological Review, v65, No. 6, pp. 386–408 - 1958.
- RUMELHART DE, H. G. E; WILLIAMS, R. J. “Neurocomputing: Foundations of Research”. In: edited by James A. Anderson and Edward Rosenfeld. Cambridge, MA, USA: MIT Press, 1988. Chapter Learning Representations by Back-propagating Errors, pages 696–699. ISBN: 0-262-01097-6.
- TOLKACH, Y.; DOHMGÖRGEN, T.; TOMA, M. et al. High-accuracy prostate cancer pathology using deep learning. **Nat Mach Intell** 2, 411–418, 2020.  
<https://doi.org/10.1038/s42256-020-0200-7> **NÃO FOI CITADA.**

WARWICK, J.; TABÀR, L.; VITAK, B.; DUFFY, S. W. Time-Dependent Effects on Survival in Breast Carcinoma: Results of 20 Years of Follow-Up from the Swedish Two-County Study. **Cancer**, v. 100, n. 7, p. 1331–1336, 2004.