

INSTITUTO PRINCIPIA
ESCOLA DE TALENTOS - TURMA DNHÁDRONS

**Reconhecimento de Padrões Morfológicos em Câncer por Redes
Neurais Convolucionais e Transfer Learning**

Gabriel Valério Ventura Alves
Colégio GGE
Caruaru, PE

Prof. Dr. Benilton de Sá Carvalho
Universidade Estadual de Campinas
UNICAMP

Julho, 2025

Resumo

O câncer compreende um conjunto de mais de cem patologias caracterizadas por alterações morfológicas complexas. Diante desse panorama, o presente estudo tem como objetivo investigar a aplicabilidade de redes neurais convolucionais (CNNs), bem como do aprendizado por transferência (*transfer learning*), na tarefa de reconhecimento de padrões morfológicos em lâminas histológicas de diferentes origens tumorais. Para tanto, foi desenvolvido um *pipeline* metodológico que contempla desde a normalização das amostras até a implementação de algoritmos de busca por arquiteturas de redes neurais otimizadas para cada subtarefa. Posteriormente, os modelos foram avaliados por meio de métricas estatísticas, com ênfase na comparação entre a ResNet50 — modelo pré-treinado — e CNNs de arquitetura simples. Os resultados obtidos evidenciam o potencial dessas abordagens como ferramentas de apoio à análise anatomopatológica, com destaque para sua aplicabilidade no suporte ao diagnóstico clínico e na redução da subjetividade inerente à avaliação manual.

Sumário

1	Introdução	1
2	Objetivos	3
2.1	Objetivo Geral	3
2.2	Objetivos Secundários	3
3	Metodologia	4
3.1	Aquisição de Dados	4
3.2	Pré-Processamento	6
3.2.1	Normalização	6
3.2.2	Aumento de Dados e Filtros	7
3.2.3	Conjuntos de Dados de Teste e Validação	8
3.3	Análise Exploratória de Dados	9
3.3.1	LC2500	9
3.3.2	Câncer Gástrico	10
3.3.3	Carcinoma Seroso de Alto Grau (HGSOC)	11
3.4	Modelagem Preditiva	12
3.4.1	Aprendizado de Máquina Supervisionado	12
3.4.2	Redes Neurais Artificiais	13
	Funções de ativação	16
	Determinação dos Pesos	19
3.4.3	Redes Convolucionais (CNNs)	22
	Transfer Learning	24
	Ajuste de Hiperparâmetros	25
3.4.4	Avaliação Estatística	26
4	Resultados	29
4.1	LC25000	29
4.2	Câncer Gástrico	34

4.3	Carcinoma Seroso de Alto Grau - HGSOC	36
5	Discussão	38
6	Conclusão	40
7	Agradecimentos	41

1 Introdução

O câncer permanece situado entre as patologias humanas mais letais e frequentes do globo, tendo, apenas no ano de 2022, uma incidência global de aproximadamente 20 milhões de casos [[International Agency for Research on Cancer, 2022](#)]. Essa condição abriga um conjunto de mais de 100 doenças com características em comum e sua complexidade se manifesta pelo descontrole do ciclo natural de divisão celular do organismo, provocando a replicação celular em ritmo desenfreado sem a necessidade de sinais bioquímicos para a continuação do processo mitótico. Esse desequilíbrio pode resultar na formação de tumores, pedaços de acúmulo de tecido que podem ser benignos ou malignos dependendo da sua capacidade de invasão tecidual (metástase).

Para avaliar o grau de disseminação de um tumor, é seguido um protocolo de estadiamento estabelecido pela União Internacional para o Controle do Câncer (UICC), designado como TNM, que avalia as características do tumor primário (T), as características dos linfonodos circundantes (N) e, por fim, a metástase a distância (M).

Uma forma de estadiamento é o patológico, baseado em fragmentos teciduais retirados por cirurgia em um paciente que, a partir de estruturas microscópicas, podem fornecer informações úteis sobre o tipo específico do câncer, grau de diferenciação celular e invasão de tecidos linfáticos e vasculares. A obtenção desse fragmento, a lâmina histológica, é realizada através de uma metodologia de coleta da amostra, ressecção em pedaços menores, fixação em uma solução que preserve estruturas celulares — como o formol 10% —, inclusão do tecido em parafina para proteção e suporte, corte em fatias finas, coloração e, por fim, preparo da lâmina histológica e observação com microscópio. É possível, também, escanear e digitalizar esses fragmentos teciduais para a análise anatomopatológica computacional através de *scanners* de alta magnificação.

No de histopatologia digital, torna-se necessária a utilização de recursos estatísticos para testes de hipóteses, obtenção de métricas de veracidade de uma suposição e extração de informações contidas nas imagens como intensidades de cor, textura e distribuição espacial do meio celular de modo eficiente e rigoroso.

À luz desse contexto, com base em anotações manualmente realizadas por patologistas, é possível treinar modelos de Aprendizado de Máquina Supervisados, que, com base nessas observações, buscam compreender padrões morfológicas que relacionam imagens com o seu rótulo previamente anotado para que assim o modelo adquira a capacidade de classificar ou prover uma resposta quanti-

tativa em amostras desconhecidas com um respectivo grau de confiança.

Esse advento tecnológico é de suma importância para a redução da variabilidade prognóstica decorrente da subjetividade e inconsistência da análise manual humana além de acelerar a velocidade de diagnósticos. Por se tratar de uma visão computacional, ele também é capaz de detectar características intrínsecas e sutis da morfologia celular que são, muitas vezes, invisíveis ao olho humano, assim permitindo novas descobertas e discernimentos no âmbito da biomedicina. Por fim, essa aplicação pode também servir como suporte para decisões clínicas de um patologista.

2 Objetivos

2.1 Objetivo Geral

Atualmente, observa-se um aumento significativo de dados coletados em diferentes sistemas e processos cotidianos e, conseqüentemente, o volume de dados disponíveis acompanha tal crescimento. Desta maneira, a ciência de dados destaca-se como um conjunto de ferramentas estatísticas e computacionais que nos permite analisar grandes volumes de dados e extrair destas bases informações essenciais para a tomada de decisão baseada em evidência.

O objetivo central deste trabalho é compreender o processo de aquisição de imagens médicas e processá-las empregando redes neurais convolucionais. Para esta fase analítica, serão empregadas técnicas que permitirão contrastar a performance de diferentes arquiteturas. Por fim, deseja-se obter um modelo de capacidade preditiva otimizada. Este modelo otimizado empregará padrões morfológicos identificados nestas imagens para prever tipos e estágios de tumores, podendo se tornar um acessório que utiliza-se da evidência acumulada em um histórico de casos para realizar uma tomada de decisão clínica.

2.2 Objetivos Secundários

- Desenvolver conhecimento técnico-científico sobre pré-processamento de imagens médicas.
- Aprender estratégias cientificamente validadas para a construção de modelos estatísticos otimizados para predição.
- Aprender e executar o ciclo de projeto de ciência de dados.

3 Metodologia

3.1 Aquisição de Dados

A histologia compreende o estudo microscópico de tecidos de organismos animais e vegetais, visando compreender a sua morfologia e as suas funções. Para a realização destas análises, é necessário seguir um protocolo de seccionamento, coloração e inspeção desses fragmentos. As imagens resultantes são, então, utilizadas em análises posteriores. Considerando as dimensões das imagens, é comum que elas sejam recortadas em imagens menores (*patches*), permitindo um processamento de dados mais eficiente

Para este estudo, utilizamos três conjuntos de imagens histológicas. Eles compreendem tecidos de quatro órgãos e, para cada uma dessas bases, há um interesse diferente para modelagem estatística.

1. **LC25000**: Conjunto com 1.350 *slides*, contendo 750 lâminas de tecido pulmonar e 500 lâminas de tecido colônico. Estas lâminas dividem-se em subtipos benignos e tumorais, respectivamente 500 e 750. As amostras em questão foram coletadas no Hospital James A. Haley Veterans’.

[Borkowski et al. \[2019\]](#) descrevem esta base de dados como sendo a fonte de um total de 25.000 *patches* de 5 entidades histológicas diferentes, tendo tamanho 768 x 768 px e 5.000 *patches* em cada uma das categorias a seguir e apresentadas na Figura 1: Adenocarcinoma do cólon (*colon_aca*) (Figura 1a), Tecido colônico benigno (*colon_n*) (Figura 1b), Adenocarcinoma pulmonar (*lung_aca*) (Figura 1c), Carcinoma de células escamosas do pulmão (*lung_ssc*) (Figura 1d) e Tecido pulmonar benigno (*lung_n*) (Figura 1e).

Para este conjunto foram exploradas três tipos de classificação distintas. Uma classificação com diferenciação de três tipos de tecido de origem epitelial do pulmão, sendo dois cancerígenos e um benigno (*lung_ssc*, *lung_aca* e *lung_n*), outra com diferenciação de dois tecidos de origem epitelial do cólon, sendo um cancerígeno e outro benigno (*colon_aca* e *colon_n*) e, por fim, uma classificação com os cinco tecidos, explorando a aptidão de diferenciar domínios histológicos diferentes, nesse caso cólon e pulmão (*lung_ssc*, *lung_aca*, *lung_n*, *colon_aca* e *colon_n*).

[Borkowski et al. \[2019\]](#) utilizaram a técnica de “aumento de dados” implementada na biblioteca Augmentor disponibilizada para Python. Neste processo, de maneira aleatória, as imagens originais foram rotacionadas aleatoriamente em até 25 graus e espelhadas horizontal e verticalmente. Ao fim do processo, obteve-se um total de 5.000 imagens por classe.

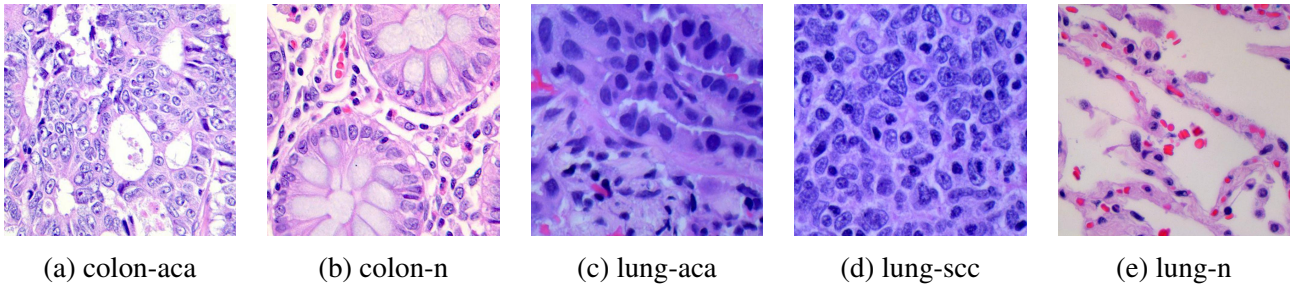


Figura 1: O conjunto de dados LC2500 divide-se em um total de cinco classes, apresentadas acima.

2. **Câncer Gástrico:** Este é um conjunto de dados obtido a partir de imagens de 300 *slides* de também 300 pacientes de câncer gástrico. Cada uma destas imagens foi dividida em *patches* de 224 x 224 px. De acordo com [Lou et al. \[2025\]](#), um total de 31.000 patches foi gerado. Cada *patch* foi classificado em 8 microambientes tumorais a saber, apresentados na Figura 2: tecido adiposo (ADI) (Figura 2a), detritos celulares (DEB) (Figura 2b), muco (MUC) (Figura 2c), tecido muscular (MUS) (Figura 2d), tecido linfóide (LYM) (Figura 2e), tecido estromal (STR) (Figura 2f), mucosa normal (NOR) (Figura 2g) e tecido de epitélio tumoral (TUM) (Figura 2h). Neste conjunto de dados, foi realizada a tarefa de classificação envolvendo as oito classes disponíveis, avaliando a capacidade do modelo preditivo de discernimento entre diferentes microambientes tumorais (*TMEs*), visto que possuem acentuada heterogeneidade.

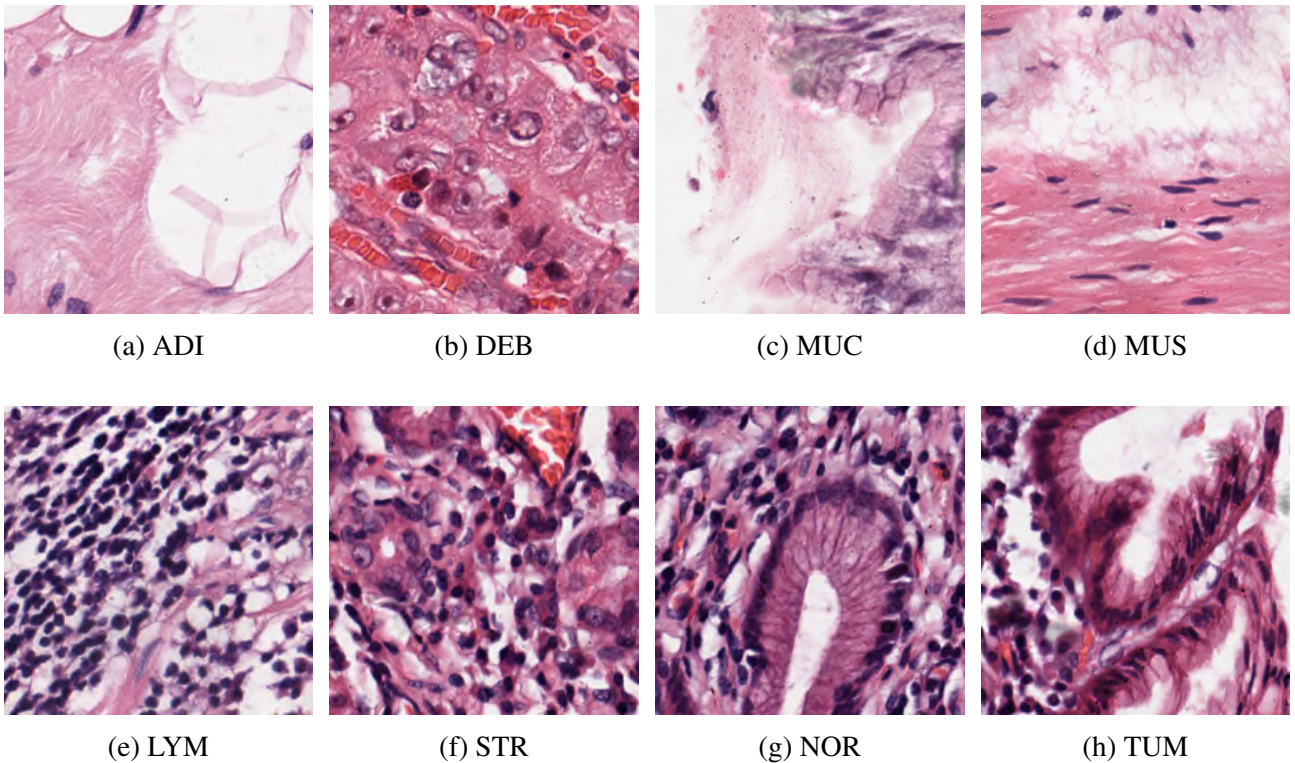


Figura 2: O conjunto de dados Câncer Gástrico disponibiliza 31.000 patches classificados em oito grupos, conforme apresentado acima.

3. **Carcinoma Seroso de Alto Grau (HGSOC)**: Este conjunto de dados abriga lâminas tumorais de carcinoma seroso de alto grau fornecidas pelo Instituto Nacional do Câncer. Ele representa amostras de 194 pacientes, em um delineamento experimental no qual cada paciente é representada minimamente por duas imagens, podendo chegar até quatro imagens por paciente. As amostras foram classificadas em uma de quatro categorias de estagiamento tumoral, representadas na Figura 3: I (13 pacientes) (Figura 3a), II (16 pacientes) (Figura 3b), III (127 pacientes) (Figura 3c) e IV (38 pacientes) (Figura 3d). As imagens originais possuem dimensão média de 5.500 x 5.500 px e foram particionadas em um total de 10.541 *patches* com dimensões 512 x 512 px. Este conjunto de dados foi explorado por meio de uma tarefa de classificação dos quatro estágios do Carcinoma Seroso de Alto Grau.

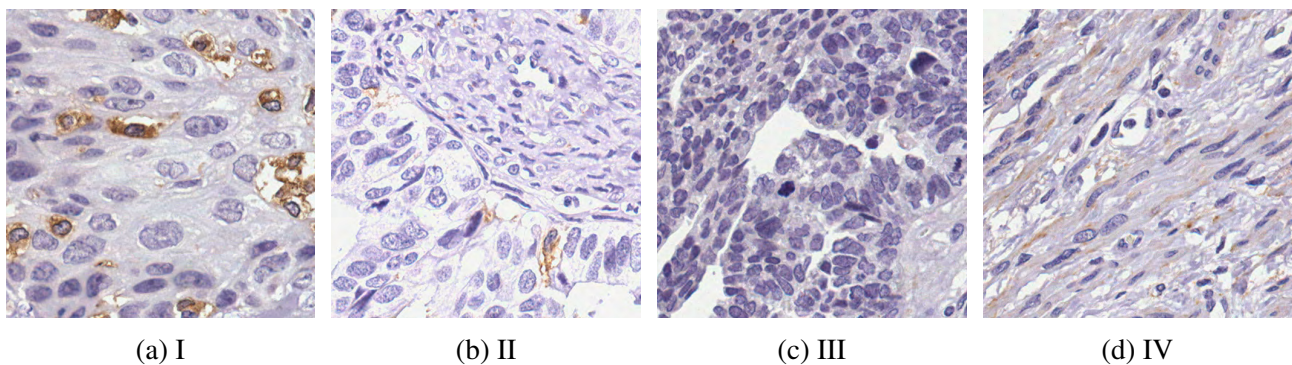


Figura 3: As amostras do conjunto de dados HGSOC são divididas em quatro estágios, conforme apresentado acima.

3.2 Pré-Processamento

3.2.1 Normalização

Considerando a natureza dos dados utilizados neste trabalho (imagens histológicas), decidiu-se pelo emprego sistemático de normalização de imagens. Este procedimento visa padronizar os dados de forma a melhorar o desempenho e a convergência do modelo em questão. Este procedimento também reduz a sensibilidade do modelo a variações que não são relevantes para a classificação, aumentando a precisão do mesmo.

Uma técnica de normalização bem consolidada é descrita por [Macenko et al. \[2009\]](#). Nela, a normalização de lâminas coradas por H&E (hematoxilina e eosina) ocorre pela conversão dos pixels de intensidade I de uma imagem RGB (normalizada para uma escala 0 a 1) em seus respectivos valores de densidade óptica (DO), conforme descreve a Equação 1. A densidade óptica permite a avaliação

da quantidade de luz que cada corante absorve, facilitando a separação dos corantes (hematoxilina e eosina) para análises posteriores. Adicionalmente, pixels de baixa densidade óptica, ou seja, abaixo do limiar β , são removidos para que a ausência de cor não afete o cálculo das etapas subsequentes.

$$DO = -\log_{10}(I) \quad (1)$$

Esta transformação permite a separação das cores da imagem de modo linear. Em seguida, são localizados os pixels com as cores azul-púrpura e vermelho-rosado mais puras, isto é, sem a influência de cores paralelas, para representar a paleta de cores representativa da imagem. Com base nessa paleta, o próximo passo do algoritmo é apurar a proporção de hematoxilina e eosina utilizada para obter a cor de cada pixel. Formam-se, então, duas novas imagens com base na original: uma apenas com as concentrações de hematoxilina e outra apenas com as concentrações de eosina. Essas duas imagens servirão de base para a repintura da lâmina original, indicando a concentração de cores estimada para cada pixel e, também, qual conjunto de cores utilizar. Neste ponto, temos uma imagem com cor e intensidade normalizada.

3.2.2 Aumento de Dados e Filtros

Slides histológicos são imagens de alta resolução e, por isso, o processamento pode requerer grandes recursos computacionais. Uma estratégia para mitigar este alto uso de recursos consiste em dividir a imagem completas em múltiplas imagens menores, denominadas *patches*. Ao criar os *patches*, existe a oportunidade de remover aquelas seções que não são informativas para o problema em questão. Por exemplo, uma região da imagem que não tenha tecido pode ser removida, reduzindo ainda mais o custo computacional. O pseudo-código apresentado no Algoritmo 1 descreve a solução implementada.

Algoritmo 1: Formação e filtragem de *patches*

Input: Imagem, dimensões (x, y) do *patch* desejado tamanho

Output: Conjunto de *patches*

- 1 Iterativamente recortar a imagem em *patches* bidimensionais com dimensões (x, y) e reiterar os passos a seguir para cada *patch* p ;
 - 2 Criar uma cópia de p (p') na escala de cinzas;
 - 3 Binarizar p' com base em um limiar θ para pretos e brancos;
 - 4 Contar o número de pixels brancos (n_b) em p' ;
 - 5 Calcular total de pixels na imagem (n_t);
 - 6 Calcular razão n_b/n_t Se razão for maior que a porcentagem pixels brancos aceita, pular para o próximo p ;
 - 7 Salvar p
-

3.2.3 Conjuntos de Dados de Teste e Validação

É essencial que, antes de treinar a rede neural com um conjunto de dados, haja a separação entre um conjunto de treino e um conjunto de validação. O conjunto de treino consiste das observações que serão utilizadas para a determinação dos parâmetros do modelo (no caso de redes neurais, pesos e vieses). O conjunto de validação é um outro conjunto de dados, que não deve ter sobreposição com o conjunto de treino, que será utilizado para avaliar o desempenho do modelo recém-ajustado. Esta validação em um conjunto de dados independente daquele empregado para ajustar o modelo é essencial para estimarmos a capacidade preditiva do modelo em um conjunto de dados “nunca visto” pelo modelo e, portanto, verificarmos a capacidade de generalização do modelo.

Estas medidas são importantes para que seja possível encontrar um balanço entre *underfitting* e *overfitting*. Em situações de *underfitting*, o modelo não consegue prever a resposta de interesse. No caso de *overfitting*, o modelo prediz com grande acurácia a resposta de interesse, mas esta performance de excelência acontece apenas no conjunto de dados de treino.

Desta forma, para cada conjunto de dados descritos na Seção 3.1, optou-se por selecionar aleatoriamente 80% da base como conjunto de dados de treino. Assim, os 20% remanescentes foram alocados para o conjunto de dados de validação.

3.3 Análise Exploratória de Dados

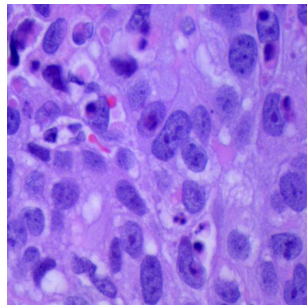
Nesta Seção, serão apresentadas algumas estatísticas sumárias para cada um dos conjuntos de dados. Com estas informações, buscaremos compreender *a priori* características que potencialmente possam afetar o ajuste dos modelos de interesse. Adicionalmente, esta análise exploratória poderá revelar amostras que devam ser removidas da análise, por exemplo, por baixa qualidade da imagem.

3.3.1 LC2500

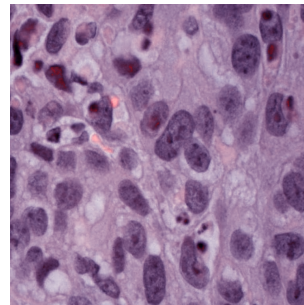
Este conjunto de dados, conforme apresenta a Tabela 1, é perfeitamente balanceado, apresentando 5.000 observações por classe. Essa característica é de grande importância para o ajuste de modelos preditivos, pois não há tendência *a priori* de classificação e, sob a hipótese de que existam diferenças morfológicas entre cada uma das classes, os parâmetros do modelos serão ajustados para identificar tais padrões.

	lung_scc	lung_aca	lung_n	colon_aca	colon_n
n	5000	5000	5000	5000	5000

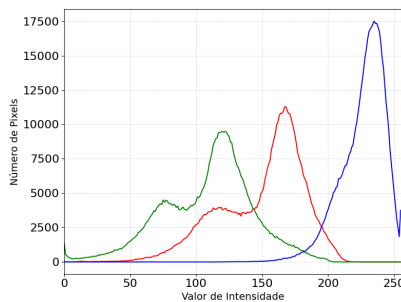
Tabela 1: Quantidade de observações por classe do conjunto LC2500



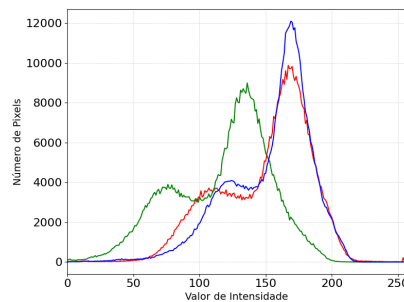
(a) Patch não normalizado



(b) Patch normalizado

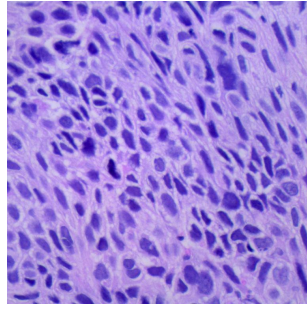


(c) Densidade da intensidade de pixels em cada canal RGB

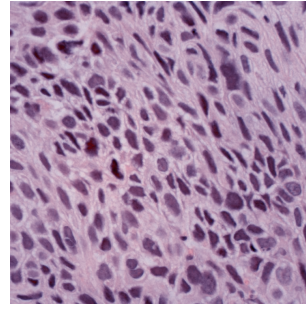


(d) Densidade da intensidade normalizada em cada canal RGB

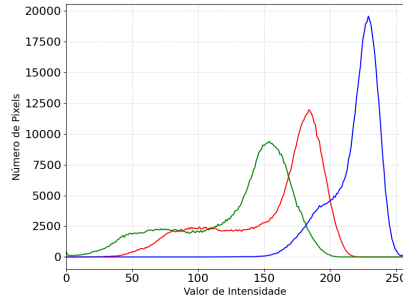
Figura 4: Um patch de uma amostra *lung_scc* do conjunto de dados LC2500 antes e depois da normalização. As imagens 4c e 4d mostram o impacto da normalização na densidade de pixels da imagem.



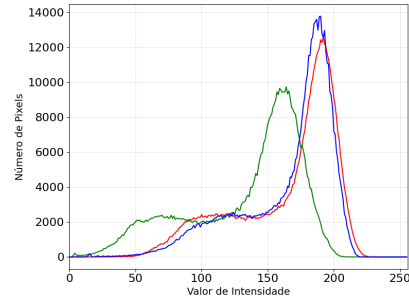
(a) *Patch* não normalizado



(b) *Patch* normalizado



(c) Densidade da intensidade em cada canal RGB.



(d) Densidade da intensidade normalizada em cada canal RGB.

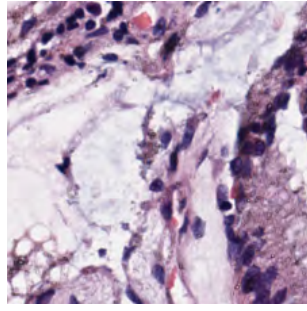
Figura 5: Normalização do *patch* original e normalização do *patch* normalizado, com seus respectivos histogramas RGB - *lung_ssc* LC2500

3.3.2 Câncer Gástrico

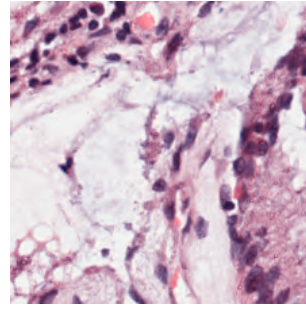
Este conjunto de dados também apresenta um balanceamento perfeito entre as classes. Como representado na Tabela 2, cada uma das classes apresenta 3.887 imagens.

	ADI	DEB	LYM	MUC	MUS	NOR	STR	TUM
n	3887	3887	3887	3887	3887	3887	3887	3887

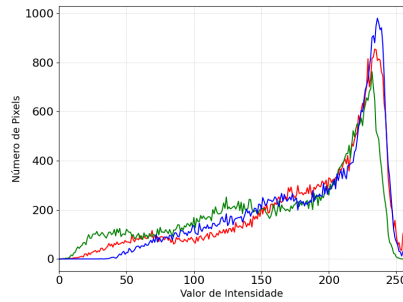
Tabela 2: Quantidade de observações por classe do conjunto de Câncer Gástrico



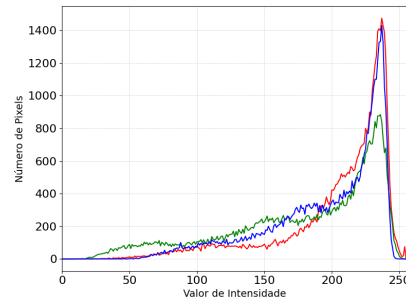
(a) *Patch* não normalizado



(b) *Patch* normalizado



(c) Densidade da intensidade em cada canal RGB.



(d) Densidade da intensidade normalizada em cada canal RGB.

Figura 6: Um patch de uma amostra de classe *MUC* do conjunto de dados de Câncer Gástrico antes e depois da normalização. As imagens 6c e 6d mostram o impacto da normalização na densidade de pixels da imagem.

3.3.3 Carcinoma Seroso de Alto Grau (HGSOC)

A Tabela 3 apresenta as frequências de observações para cada uma das classes de estágio tumoral. Nota-se uma grande diferença na contagem de observações. Essa heterogeneidade se dá pelo fato que o HGSOC ser uma doença de diagnóstico tardio, dessa maneira a grande maioria dos diagnósticos acontecem em Estágios III e IV, de forma que aproximadamente 13% dos casos são diagnosticados nos Estágios I e II [Narod, 2016].

	I	III	III	IV
n	32	38	300	77

Tabela 3: Quantidade de observações por classe do conjunto de HGSOC

Outra particularidade desse conjunto foi a significativa presença de *outliers*, estes descartados manualmente. Conforme a Figura 7a, é possível visualizar amostras descartadas em comparação com uma preservada.

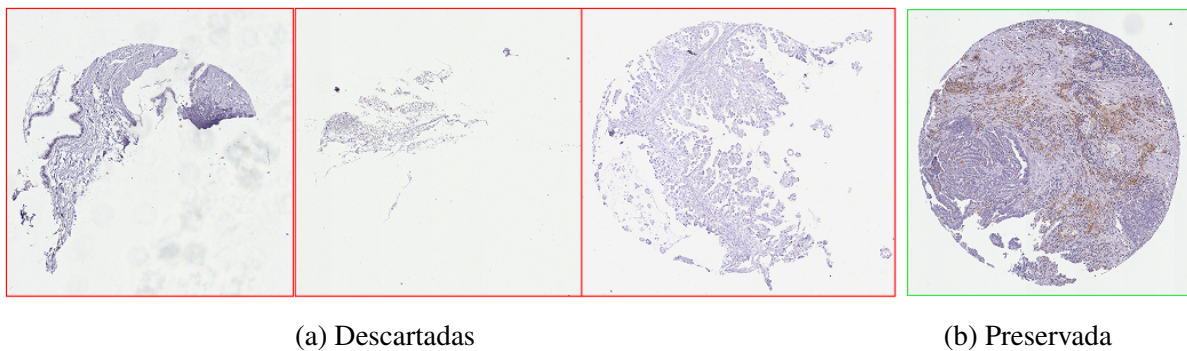
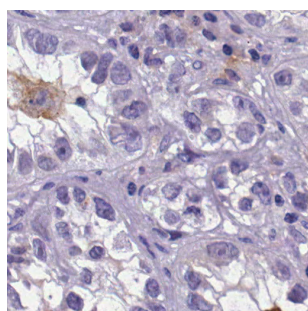
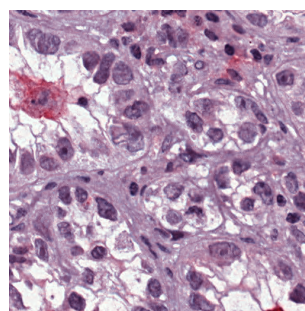


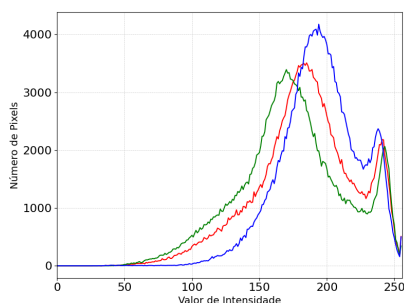
Figura 7: Comparação entre amostras descartadas e preservadas no conjunto. As amostras marcadas como descartadas mostram danos claros ao tecido, mas também incluem imagens que possuem artefatos que afetam a qualidade da imagem. As amostras marcadas como preservadas são imagens correspondentes a tecidos íntegros.



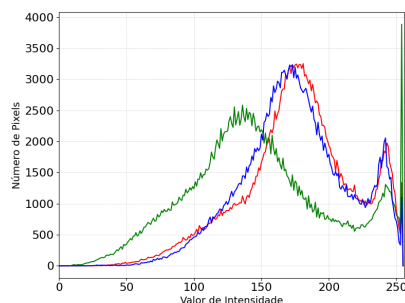
(a) *Patch* não normalizado



(b) *Patch* normalizado



(c) Densidade da intensidade de pixels em cada canal RGB



(d) Densidade da intensidade normalizada em cada canal RGB

Figura 8: Um patch de uma amostra de estágio I do conjunto de dados do HGSOc antes e depois da normalização. As imagens 8c e 8d mostram o impacto da normalização na densidade de pixels da imagem.

3.4 Modelagem Preditiva

3.4.1 Aprendizado de Máquina Supervisionado

O Aprendizado de Máquina é um ramo da inteligência artificial que consiste na criação de modelos capazes de aprender relações entre dados fornecidos e, consequentemente, realizar tarefas sem

programação prévia. Estes tratam-se de modelos analíticos que são capazes de se adaptarem de forma independente quando expostos a dados desconhecidos. Um subconjunto do Aprendizado de Máquina é o Aprendizado Profundo, no qual redes neurais com, pelo menos, três camadas são empregadas para a identificação de padrões e predição de respostas de interesse.

No Aprendizado de Máquina Supervisionado, apresenta-se ao modelo um conjunto de observações. Para cada observação, o modelo tem acesso a um conjunto de variáveis preditoras, também conhecidas como *features* e também a uma variável resposta. As variáveis preditoras descrevem características da respectiva observação e, com elas, o modelo busca identificar associações com a variável resposta de interesse. A variável resposta pode ser uma medida quantitativa (como, por exemplo, peso, estatura, temperatura, etc.) ou uma variável qualitativa (como, por exemplo, *status* com respeito a uma doença, estágio de um certo câncer, tipo de resposta a um certo tratamento, etc.). No caso de respostas quantitativas, dizemos tratar-se de um **problema de regressão**; já, para respostas qualitativas, denominamos como um **problema de classificação**.

Considerando os conjuntos de dados descritos na Seção 3.1, trataremos de problemas de classificação. Neste cenário, é importante observar a existência de uma grande gama de modelos estatísticos que podem ser utilizados no Aprendizado de Máquina para Problemas de Classificação. Podemos destacar modelos de regressão logística e multinomial, florestas aleatórias, *support vector machines* e redes neurais. Neste trabalho, empregaremos redes neurais especializadas em imagens, que buscarão padrões morfológicos capazes de prever a classificação das amostras disponíveis.

3.4.2 Redes Neurais Artificiais

O cérebro humano é um órgão de complexidade notável, resultado de uma evolução da ordem de milhões de anos, que visa garantir as melhores chances de sobrevivência do organismo diante da natureza. Ao longo desse período, a neuroplasticidade possibilitou a formação de novas ligações entre neurônios e o fortalecimento de conexões sinápticas específicas com base em experiências e seus resultados. O neurônio, representado na Figura 9, é a unidade elementar desse sistema e responsável por receber impulsos nervosos, processá-los e conduzir para uma tomada de decisão.

Essa estrutura serviu como inspiração para a criação das redes neurais artificiais, aqui denominadas simplesmente por **redes neurais**. Em resumo, elas tomam sinais de entrada, aplicam transformações matemáticas e retornam resultados de saída que são, posteriormente, empregados em análises ou

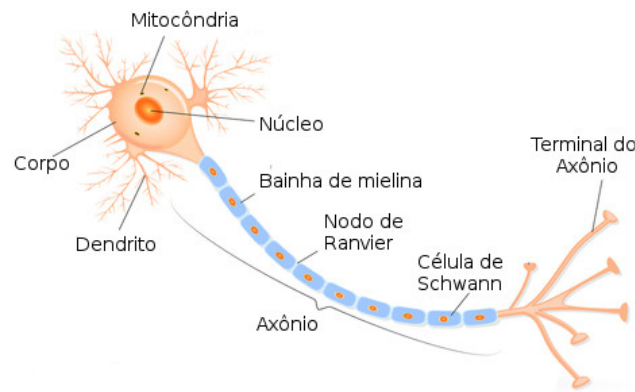


Figura 9: Estrutura de um neurônio. Ilustração: Designua / Shutterstock.com [adaptado]

interpretados em tomada de decisão. A Figura 10 a seguir apresenta a unidade básica de uma rede neural, análoga ao mecanismo biológico citado, também conhecida como rede de camada única ou *Perceptron Rosenblatt*.

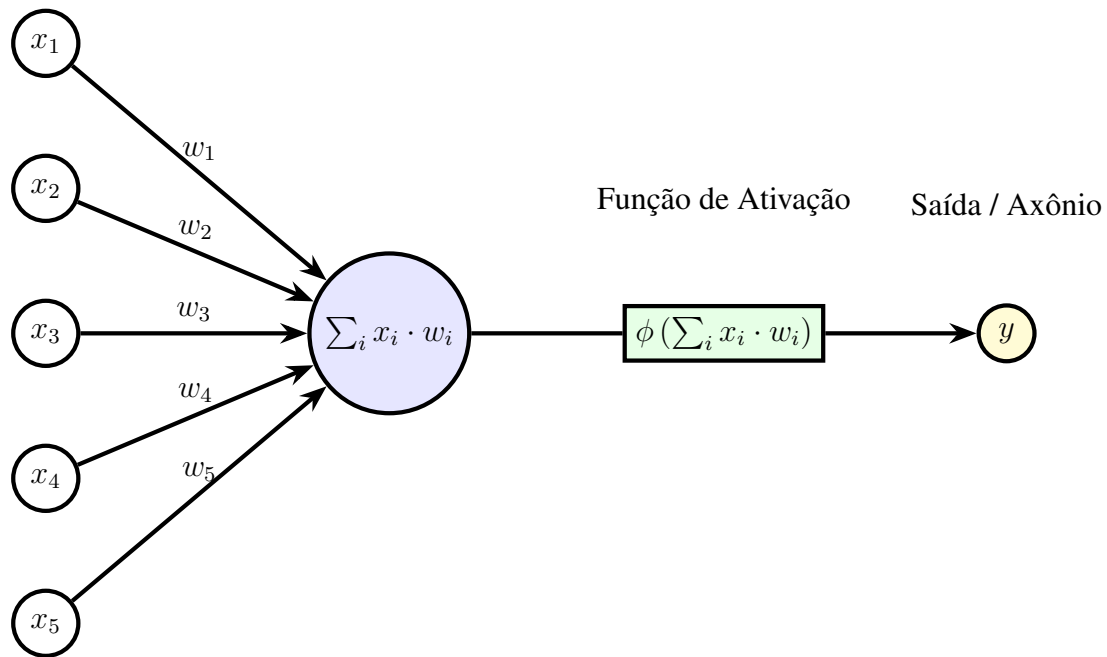


Figura 10: Unidade básica de uma rede neural com cinco entradas.

Nos nós de entrada, como demonstra a Figura 10, dados a serem analisados são inseridos, multiplicados por seus respectivos pesos e transferidos para a função de ativação como resultado do produto interno entre ambos os vetores (x e w). Como se trata de um cenário de aprendizado de máquina supervisionado, i.e. sabemos *a priori* qual resultado o modelo deve obter, procede-se então para a comparação do resultado obtido na saída da função de ativação com este resultado verdadeiro. Desta maneira, é possível determinar os pesos w_i iterativamente, de forma a aproximar-se o máximo possível do resultado verdadeiro.

As redes neurais podem, também, possuir camadas adicionais, entre a camada de entrada e a saída. A Figura 11 apresenta esta configuração. Essas camadas adicionais são chamadas de camadas ocultas e são empregadas na identificação de padrões mais complexos do que aqueles detectados por redes de camada simples.

A identificação de padrões mais complexos em camadas ocultas advém de uma combinação de fatores: a) a capacidade destas camadas detectarem padrões não lineares e; b) as camadas ocultas atuarem sequencialmente, permitindo que padrões cada vez mais complexos sejam detectados. Um modo claro de visualização é no emprego de Redes Neurais Convolucionais, que serão abordadas posteriormente, que no reconhecimento de imagens, as camadas ocultas quebram o contexto geral em menores elementos de compreensão. As primeiras camadas podem identificar elementos mais simples, como bordas e texturas. À medida que a rede torna-se mais profunda, as camadas ocultas passam a detectar padrões mais complexos, como formas, cantos e até formatos de objetos em certos casos.

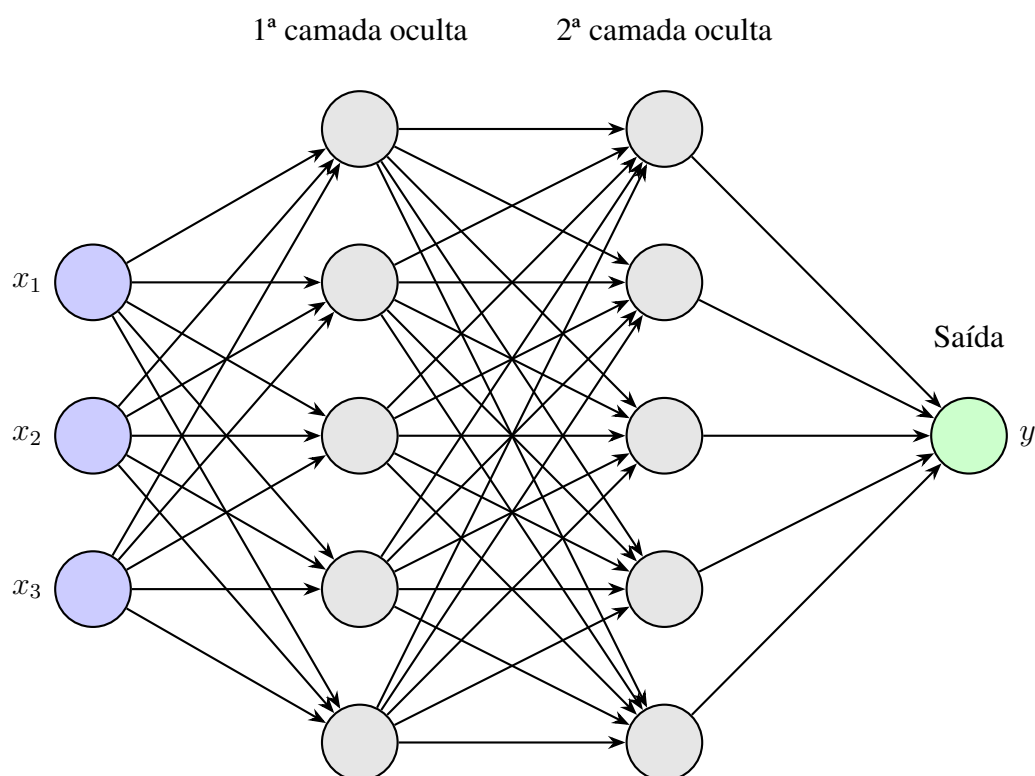


Figura 11: Rede Neural com 2 camadas oculta. Diagrama de autoria própria.

Os neurônios das camadas ocultas, da mesma forma que a camada de saída, têm em sua constituição funções de ativação. Estas funções são responsáveis por determinar o quão intensamente esse nó é ativado baseado no produto interno das medidas em questão e seus respectivos pesos. As funções

de ativação também desempenham um importante papel na detecção de padrões não-lineares que podem existir no conjunto de dados em questão. A Figura 12 apresenta um exemplo de função linear, representada por uma reta de inclinação 2, e outro exemplo de função não-linear, um polinômio de segundo grau.

A exigência de detectar padrões não-lineares provém do fato que regiões de discriminação entre grupos nem sempre são delimitadas por retas. Esta situação pode ser visualizada na Figura 13c. Um cenário embasado na vida real que exemplifica isso é a classificação de zonas de silêncio e zonas de *show* com base nas variáveis: distância e altura do som. É notório que nem toda região com um som acima de 80 dB corresponde a uma área de show, como nem todo local mais distante que 200 metros é uma zona de silêncio. Assim, esse caso não é separável por uma reta e requer uma relação não-linear.

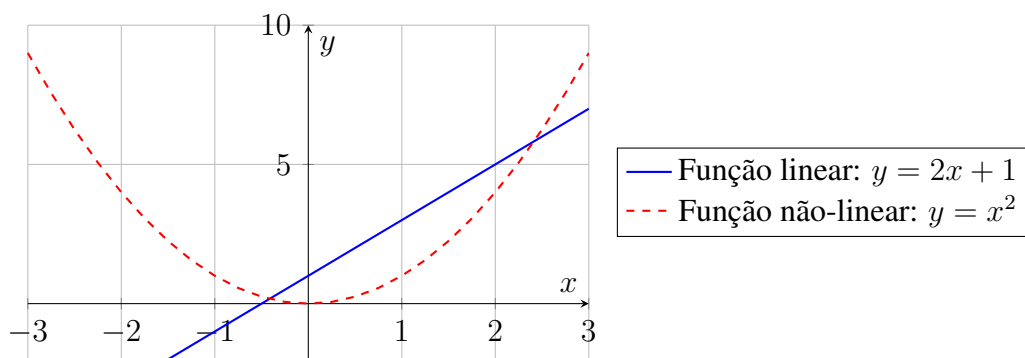
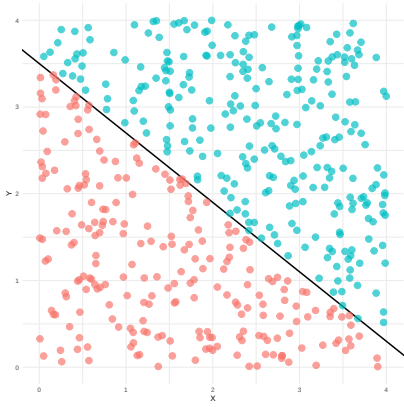


Figura 12: Comparação entre função linear e não-linear

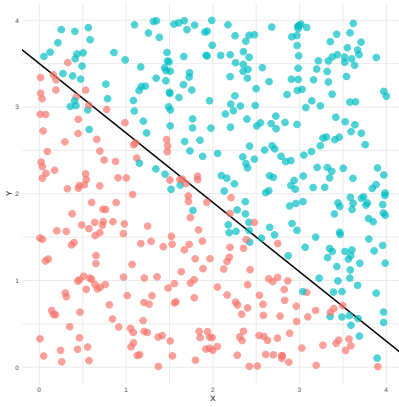
Desta maneira, é possível identificar que diferentes cenários são prováveis de acontecer. Numa primeira situação, é claramente possível empregar uma função linear para traçar uma fronteira de decisão entre dois grupos, como apresenta a Figura 13a. Em muitas ocasiões, a situação mais comum é que a fronteira de decisão não separa perfeitamente os grupos em questão, podendo ocorrer alguns erros de classificação quando utilizando a fronteira de decisão como parâmetro para a escolha entre os grupos, como mostra a Figura 13b.

Assim, à medida que a complexidade da relação entre dados aumenta, torna-se mais difícil a tarefa de identificar uma fronteira de decisão entre grupos. Empregar uma função linear, nos moldes do que apresentam as Figuras 12, 13a e 13b, deixa de ser uma opção. Para melhor atender a esta situação, torna-se necessário o uso de relações não lineares, como a apresentada na Figura 13c.

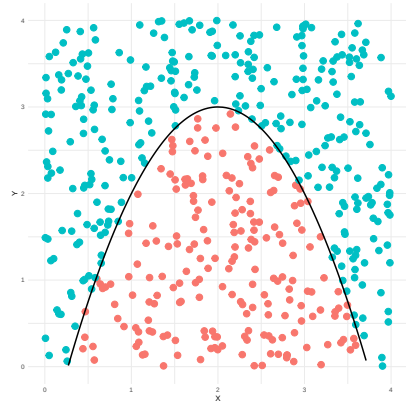
Funções de ativação



(a) Separação de dois grupos em um caso no qual uma reta permite a separação perfeita entre ambos os grupos.



(b) Separação de dois grupos em uma situação em que algumas observações podem ser incorretamente classificadas.



(c) Situação em que a discriminação de dois grupos depende de uma relação não-linear das variáveis disponíveis.

Figura 13: A discriminação de grupos pode ocorrer em diferentes cenários. Em algumas situações, é possível que os grupos sejam separados sem dificuldades com uma reta, como apresenta a Figura 13a. Em outros cenários, que são mais comuns, é possível discriminar corretamente a maioria das observações, mas aquelas mais próximas da barreira de separação dos grupos podem ser incorretamente classificadas, como indica a Figura 13b. Ainda existem ocasiões em que a fronteira de separação dos grupos é não-linear, como mostra a Figura 13c, que também podem incluir dificuldade na classificação das observações próximas à fronteira.

As funções de ativação são estratégias matemáticas empregadas nas saídas de cada neurônio de redes neurais para permitir que elas identifiquem padrões não lineares. Uma rede neural, por mais complexa que seja, sem funções de ativação é apenas uma função linear, que terá aplicações bastante limitadas por não conseguir aprender padrões complexos habitualmente encontrados em problemas atuais. Considerando que, em cada nó, as saídas da camada imediatamente anterior são combinadas com os respectivos pesos de conexão e viés por meio do produto interno, temos que a representação pré-ativação é dada pela Equação 2, apresentada abaixo:

$$= b + \sum_{j=1}^n x_j w_j \quad (2)$$

$$= \langle (1, x_1, x_2, \dots, x_n), (b, w_1, w_2, \dots, w_n) \rangle. \quad (3)$$

Desse modo, em cada nó, aplica-se uma função de ativação $f(\cdot)$ no resultado apresentado pela Equação 2, ou seja, determina-se $f\left(b + \sum_{j=1}^n x_j w_j\right)$. Assim, a rede neural torna-se capaz de identificar padrões não lineares. Faz-se aqui um destaque para a representação dada pela Equação 3, que gerará o mesmo resultado numérico, mas será ordens de magnitude mais rápida em operações

computacionais por conta de suas extensões naturais em forma matricial.

Diferentes funções de ativação suprem demandas específicas de determinadas aplicações. Quando uma arquitetura de rede neural é estabelecida, para cada camada fixa-se uma função de ativação, a qual é reiterada em todos os neurônios. A Figura 14 apresenta funções de ativação e suas respectivas funções gradiente (derivadas) frequentemente utilizadas em redes neurais. As funções gradiente são essenciais para o processo de determinação dos pesos e vieses de uma rede neural, pois indicam a direção em que o processo de otimização deve seguir até que seja atingido o ponto ótimo da função perda.

Ainda com base na Figura 14, observam-se três funções de ativação comumente utilizadas. A ReLU é ideal para casos em que se busca esparsidade, isto é, dispersão na quantidade de ativações dos neurônios, dado que retorna o máximo entre 0 e o valor de entrada, efetivamente desabilitando os nós com valores abaixo de zero. A função sigmoide se encaixa em contextos nos quais é necessária uma interpretação probabilística (como em camadas de saída de modelos de classificação), pois retorna um número real entre 0 e 1, representando a probabilidade de ocorrência de um evento de interesse. Já a *softmax* é uma função semelhante à sigmoide, utilizada em casos de classificação multiclasse, pois retorna um vetor de probabilidades que somam 1. A tangente hiperbólica lida eficientemente com valores negativos pois ela retorna um intervalo de -1 até 1. Adicionalmente, esta função é centrada em zero, preservando a simetria sobre o eixo das ordenadas e contribuindo para o melhor balanceamento dos pesos.

Outro parâmetro importante nas redes neurais são os vieses (bias), valores que são adicionados à soma dos pesos multiplicados por valores da entrada. Partindo de uma analogia simples, uma função linear, temos:

$$y = W \cdot x + b, \quad (4)$$

sendo y o vetor de saída, W denota a matriz de pesos (coeficiente angular), de dimensão dependente da quantidade de ligações de nós de uma camada com a sua sucessora, sendo cada elemento o peso da conexão do j -ésimo nó de uma camada L com o i -ésimo nó da camada seguinte $L + 1$ (W_{ji}), x é o vetor de entradas e b é o vetor de deslocamento do viés (intercepto). Se uma função não tem o termo b explicitamente apresentado, então $b = 0$ e ela passará pela origem cartesiana (0,0) em todos os cenários. A depender da natureza do problema, esta pode ser uma característica desejada. Cada neurônio, quando se utiliza o viés, incorpora esse valor ao produto interno calculado entre a entrada

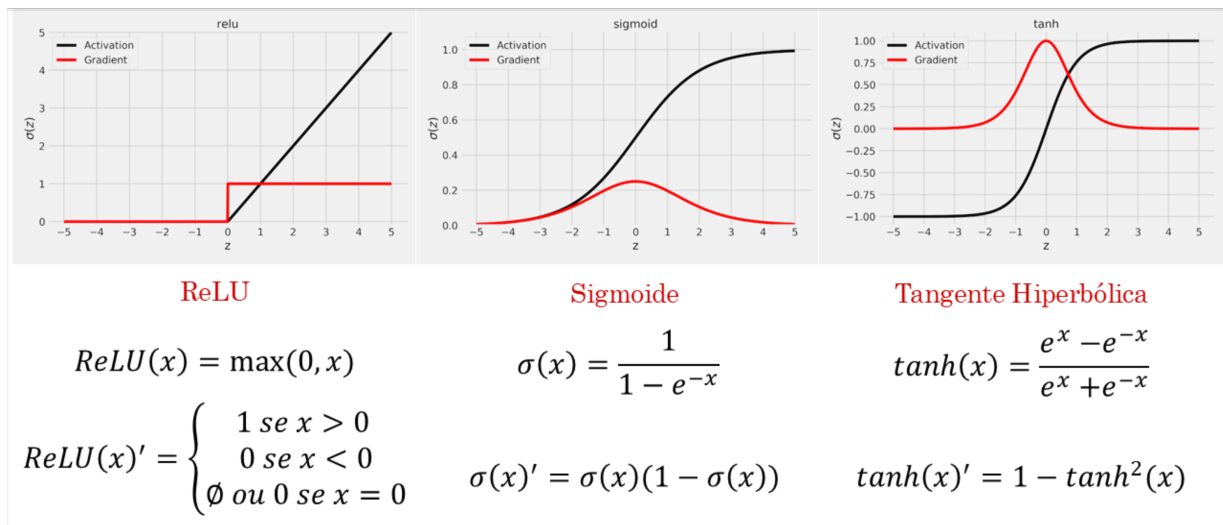


Figura 14: Diferentes funções de ativação e seus respectivos gradientes. Fonte: Deep Learning with PyTorch Step-by-Step, Daniel Voigt Godoy

da camada anterior e o vetor de pesos.

Determinação dos Pesos

Os pesos são fatores numéricos que determinam a intensidade de conexões entre camadas sucessivas, por meio do produto de seu valor com o valor do nó de origem, conforme apresentam as Figuras 10 e 11. Para que eles sejam determinados, utiliza-se o método da retropropagação, que é um algoritmo essencial para o ajuste de redes neurais e consiste de dois passos essenciais: a propagação da entrada e a retropropagação do erro.

Inicialmente, os dados de entrada alimentam a rede neural e cada neurônio aplica o produto interno da entrada com os pesos designados e repassa esta combinação linear à função de ativação. Este processo ocorre em camadas e assim segue até que se chegue na camada de saída. Na saída, utiliza-se a função perda para se determinar o tamanho do erro obtido com aquele conjunto de pesos. Desta maneira, a rede pode avaliar a sua própria performance.

Na fase da retropropagação, o erro é enviado de volta pela rede, passando pelas camadas interme-

diárias, até chegar à camada de entrada. Este processo é fundamentado na regra da cadeia (conceito de Cálculo Diferencial), que permite o cálculo do gradiente da função perda com respeito aos pesos da rede neural. Com o gradiente calculado, é possível ajustar os pesos de forma a minimizar o erro. O algoritmo *Gradient Descent* é habitualmente empregado nesta fase para determinar o tamanho do ajuste dos pesos e a direção deste ajuste.

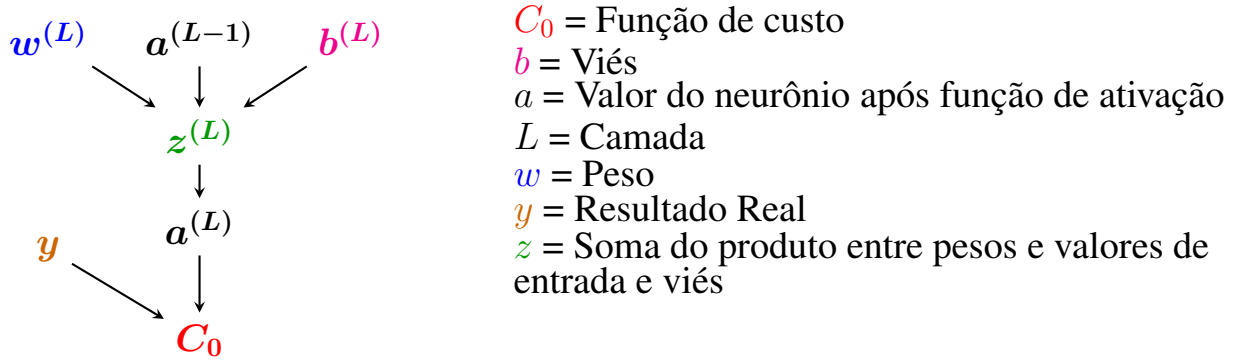


Figura 15: Fluxo de dados em uma rede neural: da camada anterior ao cálculo do custo. Adaptado do vídeo “Backpropagation calculus | Deep Learning Chapter 4”, 3Blue1Brown (Youtube, Acesso 22/06/2025).

Visto isso, é essencial abordar o conceito de derivada e como ela se relaciona com o esquema proposto. Esta ferramenta matemática, denotada por $f'(x)$ ou $\frac{df}{dx}$, calcula a taxa de variação de uma função em intervalos infinitesimais tendendo a zero. Assim, como mostra a Figura 15, na qual há a interligação de uma sequência de funções que variam infinitesimalmente à medida que se propagam no fluxo da rede neural, a regra da cadeia fornece a base para o cálculo de derivadas de funções compostas, representada em sua lei geral a seguir.

$$(f \circ g)'(x) = f'(g(x)) \cdot g'(x) \quad (5)$$

Desse modo, a regra da cadeia (5) se aplica à Figura 15 da seguinte forma.

$$\frac{\partial C_0}{\partial w^{(L)}} = \frac{\partial z^{(L)}}{\partial w^{(L)}} \cdot \frac{\partial a^{(L)}}{\partial z^{(L)}} \cdot \frac{\partial C_0}{\partial a^{(L)}} \quad (6)$$

O mesmo se dá em relação ao viés.

$$\frac{\partial C_0}{\partial b^{(L)}} = \frac{\partial a^{(L)}}{\partial z^{(L)}} \cdot \frac{\partial C_0}{\partial a^{(L)}} \quad (7)$$

É importante frisar que a variação da função de custo para a atualização de todos os pesos é igual

a um vetor que representa a variação do custo para cada um dos pesos na rede neural.

Os gradientes são ferramentas matemáticas que indicam a direção de um deslocamento vetorial para o aumento mais rápido de uma grandeza. No caso da retropropagação, temos um gradiente negativo com n dimensões (funções de custo obtidas), que aponta para que direção a função de custo desvanece mais rapidamente, assim chegando em um mínimo, que pode ser local ou global, visto que gradientes podem apresentar múltiplos picos e vales no espaço tridimensional.

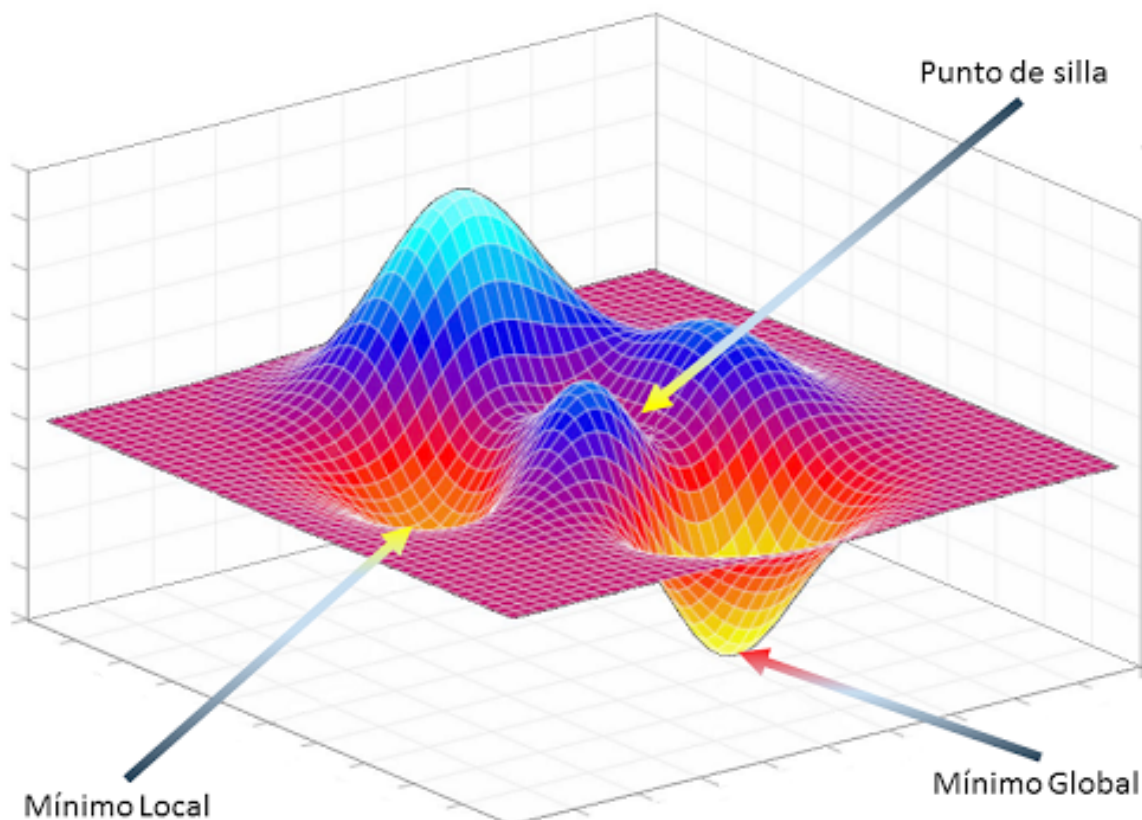


Figura 16: O processo de otimização deve encontrar um ponto ótimo na função a ser otimizada. Ao se tratar de funções de perda, este ponto é costumeiramente o mínimo global. Entretanto, o algoritmo utilizado nesta busca deve ser capaz de contornar situações em que o mínimo local possa ser erroneamente visto como ponto ótimo. Fonte: <https://numerentur.org> (Acesso 23/06/2025).

Assim, para minimizar a discrepância entre as previsões, é acrescentada ou descontada pequenos valores de mudança com base na variação da função de custo com respeito aos pesos, controlados pelo termo α , que regula o quão distantes são esses “passos”, permitindo uma gradual convergência em um ponto de erro mínimo.

$$\nabla C = \begin{bmatrix} \frac{\partial C_0}{\partial w_1} \\ \vdots \\ \frac{\partial C_0}{\partial w_n} \end{bmatrix} \rightarrow \Delta w \Leftrightarrow -\alpha \nabla C$$

$$\bar{w} \Leftarrow \bar{w} - \alpha \frac{\partial C_0}{\partial \bar{w}}$$

Por fim, para lidar com o desbalanceamento de classes, utilizamos a função *compute_class_weight*, da biblioteca *scikit-learn*, que calcula os pesos de cada classe com base na respectiva quantidade de observações disponíveis. Desse modo, erros em classes minoritárias são penalizados com intensidade inversamente proporcional à sua frequência no conjunto de dados, diminuindo vieses classificatórios decorrentes desse desequilíbrio.

3.4.3 Redes Convolucionais (CNNs)

No caso da aplicação de redes neurais para o reconhecimento de imagens, são usadas as redes neurais convolucionais que, inspiradas no mecanismo do córtex frontal de animais, se otimizam com a finalidade de abstrair características de imagens. A convolução — operação algébrica que denota a sobreposição de duas funções —, no contexto das redes neurais, trata-se do deslizamento iterativo de uma matriz bidimensional (*kernel*) ao longo de uma imagem tridimensional (largura, altura, canais), no qual é calculado o produto escalar da sobreposição de cada célula da matriz com o pixel da imagem no segmento em que é sobreposta. Assim, obtemos novas imagens, estas com características destacadas (e.g., bordas, texturas, formatos) exemplificado na Figura 17.

O aprendizado emerge no reajuste dos valores das matrizes de convolução, estes análogos aos pesos de uma rede neural simples, buscando estabelecer uma hierarquia de características relevantes para um certo fim.

Quando se trata do âmbito da medicina, as CNNs tornam-se ferramentas importantes para a classificação automatizada e mais robusta de lâminas histológicas. O processo de exame de uma amostra de tecido por um patologista requer tempo e pode desacelerar a triagem de pacientes, além de conter um certo grau de subjetividade dependendo do indivíduo que realiza o trabalho. As CNNs, após treinadas, são rápidas e mais robustas na classificação de tecidos, tornando-se ferramentas importantes para auxiliar no diagnóstico clínico com maior agilidade, padronização e reprodutibilidade,

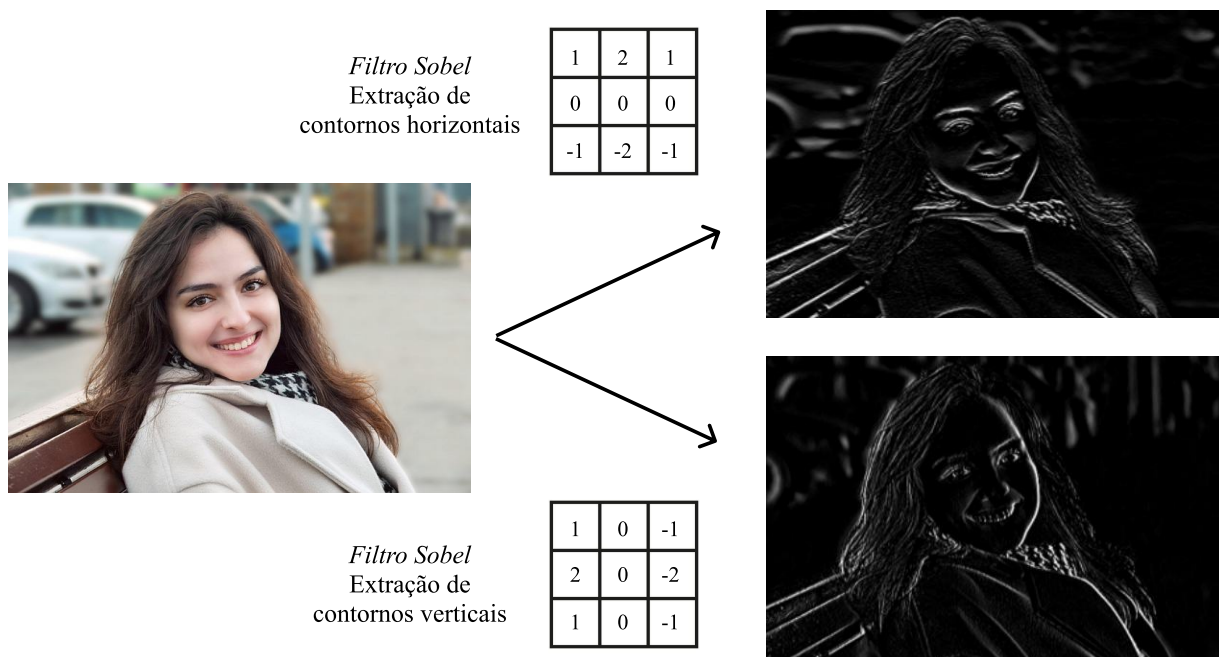


Figura 17: Exemplificação da convolução de uma imagem 640x427 com dois *kernels* diferentes. Fonte da imagem: <https://pixabay.com> (Acesso 05/07/2025).

também reduzindo vieses humanos e auxiliando decisões médicas.

Uma rede convolucional é comumente constituída de três elementos principais: camadas convolucionais, camadas *pooling* e *fully connected*.

As camadas convolucionais extraem características de imagens da camada anterior, sendo responsáveis pela criação de n mapas característicos, resultantes da convolução de n filtros — matrizes aqui denominadas por kernel — em cada uma dos canais da imagem (Vermelho, verde e azul na maioria das imagens coloridas).

As camadas de *pooling* servem para reduzir as dimensões espaciais dos mapas característicos obtidos, performando operações como o *pooling* de máximos, no qual uma matriz varre uma imagem e produz um novo mapa reduzido apenas com o maior valor de cada secção da imagem sob a iteração da matriz, e o *pooling* de médias, no qual um novo mapa reduzido é produzido através do cálculo da média dos valores para cada etapa da iteração da matriz.

As camadas *fully connected* (completamente conectadas) correspondem a camadas com uma quantia n de neurônios os quais completamente se conectam com os da camada seguinte e anterior. No contexto de redes convolucionais, exercem a importante função de achatar os filtros da camada anterior em um vetor unidimensional, para que assim se propague pelas camadas densas como uma rede neural simples.

Uma arquitetura simples de CNN é exemplificada na Figura 18.

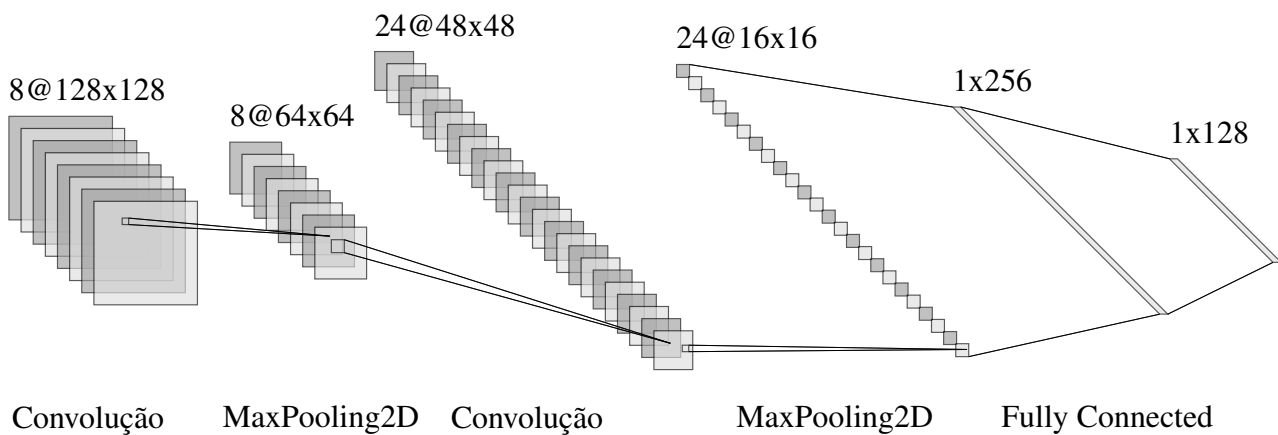


Figura 18: Arquitetura de Rede Convolutiva simples

Transfer Learning

O *Transfer Learning*, traduzido como Aprendizado por Transferência, consiste na implementação dos pesos e arquitetura de um modelo pré-treinado para desempenhar uma tarefa diferente da que foi originalmente treinado para realizar.

A vantagem evidente desse método é a capacidade de utilizar modelos treinados com elevada quantidade de observações e profundidade arquitetônica sem a necessidade de retrainar esse modelo, o que é inviável sob recursos computacionais limitados.

Comumente, as camadas do modelo pré-treinado são congeladas, ou seja, os seus pesos não serão reajustados e servirão apenas para extração de características da imagem, e no topo dessas camadas podem ser adicionadas as *head layers*, camadas que serão treinadas para aperfeiçoar o modelo pré-treinado para a execução da nova tarefa.

Uma subcategoria do Aprendizado de Transferência é o Ajuste Fino, em que parte do modelo pré-treinado é descongelado e este também passa a fazer parte do treino com sua arquitetura. Esse processo é executado gradualmente, com o descongelamento de camadas e monitoramento dos resultados.

Desse modo, utilizamos a ResNet50 como modelo de base para a aplicação do *transfer learning* nos conjuntos de dados utilizados e para subsequente comparação do seu desempenho com uma arquitetura de CNN Simples. A ResNet, descrita por He et al. [2015], é uma rede residual profunda

idealizada para visão computacional. Por fim, utilizamos os pesos da ResNet50 treinada no conjunto de dados ImageNet [Deng et al., 2009], que abriga 14.197.122 imagens anotadas.

A medida que são adicionadas mais camadas em uma rede neural comum, torna-se possível a ocorrência de dois problemas com os gradientes de função de perda em relação aos pesos da rede: o esvanecimento e a explosão. O esvanecimento ocorre quando esse gradiente passa a tender a zero a medida que se aprofunda na rede neural, impedindo a atualização dos pesos nas camadas iniciais. Já a explosão ocorre quando as matrizes dos gradientes (Eq. 6) têm valores maiores que 1 e seus produtos passam a crescer exponencialmente ao decorrer da cadeia. Desse modo, os pesos da rede crescem de forma brusca e impedem o aprendizado do modelo.

Dado esses problemas, a ResNet os mitiga através de *conexões residuais*, em que é estabelecida uma conexão entre a saída de uma camada (x) e uma saída posterior de modo que as camadas intermediárias são puladas. No ponto final de conexão, o valor dessa saída será equivalente a soma de x e $\mathcal{F}(x)$, que é a função aprendida pelas camadas intermediárias em um fluxo ordinário da rede neural, exemplificado na Figura 19. Assim, torna-se possível o treino de redes neurais profundas sem a perda de informação.

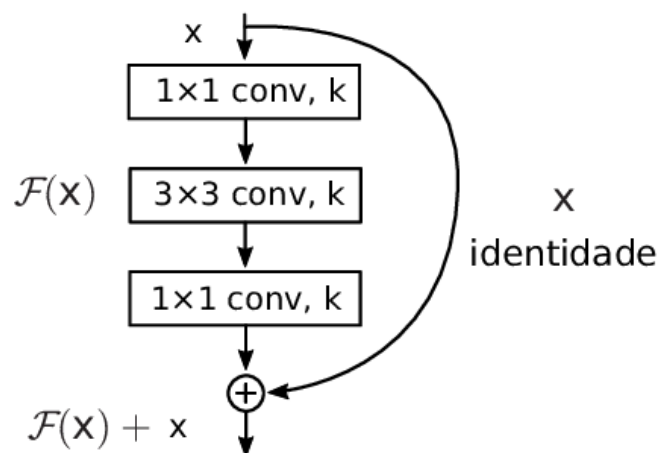


Figura 19: Bloco do Modelo Residual da ResNet [Bezerra, 2018]

Ajuste de Hiperparâmetros

Os hiperparâmetros são as variáveis que regem a estrutura, performance e função do modelo. Alguns exemplos de hiperparâmetros são a quantidade de neurônios por camada densa, quantidade de filtros por camada convolucional e taxa de aprendizado do modelo. Como estes são definidos antes do treinamento da rede, torna-se mais complexo decidir se os hiperparâmetros escolhidos vão retornar

o melhor desempenho no modelo para a tarefa desejada. Deste modo, os algoritmos de otimização de hiperparâmetros são responsáveis pela execução de uma busca iterativa do melhor conjunto de parâmetros para uma tarefa a partir da comparação dos resultados obtidos.

Existem variados algoritmos que performam uma busca de hiperparâmetros de forma automatizada. Apesar de diferenças, compartilham em comum a capacidade do usuário decidir qual espaço amostral pretende explorar. Ou seja, explicitar quais hiperparâmetros deseja variar e que espectro de valores cada um deve se encaixar. A Pesquisa Aleatória, conforme indicado pela nomenclatura, executa a rede neural com variações aleatórias, até que um valor ótimo seja alcançado. Esta é ideal para espaços amostrais grandes. Já a Pesquisa em Grade testa exaustivamente todas as permutações possíveis de acordo com o espectro de possibilidades fornecidas pelo usuário. Esta não é otimizada e funciona melhor em espaços amostrais menores.

Neste trabalho, utilizou-se o algoritmo de *Otimização Bayesiana* da biblioteca de *Python Keras Tuner* para otimizar a métrica de acurácia na validação. A busca Bayesiana, a cada iteração, escolhe novos hiperparâmetros para teste com base nos anteriores. Desse modo, sob influência de resultados prévios, essa busca explora com mais cautela microrregiões de combinação de hiperparâmetros com resultados promissores, dispensando aleatoriedade e necessitando de menos passos para convergir em um ponto ótimo.

3.4.4 Avaliação Estatística

Por fim, após o processo de treino de uma rede neural, é crucial a avaliação do desempenho do modelo por meio múltiplas métricas de quantificação que garantam uma consolidação homogênea da capacidade da rede neural de performar suas previsões para novos dados além dos que foram usados para treino. Para as explicações a seguir, serão utilizadas analogias com os conjuntos de dados usados nesse estudo.

- **Matriz de Confusão**

A Matriz de Confusão permite visualizar os resultados de um algoritmo classificador. As colunas representam as previsões de um modelo para uma determinada classe enquanto as linhas representam a classe real (verdade fundamental). Se em uma célula o seu par (x, y) corresponde a duas classes iguais, pode-se concluir que o modelo corretamente previu a classe da respectiva entrada. Um modelo que acerta todas as suas previsões apenas apresenta valores diferentes de

zero nas células da diagonal da matriz. Um esquema de Matriz de Confusão de duas classes é visualizado na Tabela 4.

		Classe Prevista	
		$A(-)$	$B(+)$
Classe Real	$A(-)$	Verdadeiro A (VA)	Falso A (FA)
	$B(+)$	Falso B (FB)	Verdadeiro B (VB)

Tabela 4: Matriz de Confusão Binária (2 classes)

- **Acurácia**

A Acurácia mede, de todas as previsões que o modelo fez acerca da classificação do tecido, a proporção das que foram corretamente previstas.

$$\text{Acurácia} = \frac{\text{Previsões Corretas}}{\text{Todas as Previsões}} = \frac{VA + VB}{VA + FA + FB + VB} \quad (8)$$

- **Sensibilidade e Especificidade**

A Sensibilidade é uma métrica que mede a habilidade do modelo de corretamente identificar os tecidos que se encaixam em um determinado tipo dentre todos os tecidos daquele tipo no conjunto utilizado para previsão. Ou seja, uma alta sensibilidade para uma classe de tecido significa que o modelo não deixou passar despercebida amostras do mesmo. A Equação 9 apresenta a definição matemática da métrica e sua representação de acordo com a Tabela 4.

$$\text{Sensibilidade} = \frac{\text{Verdadeiros Positivos}}{\text{Verdadeiros Positivos} + \text{Falsos Negativos}} = \frac{VB}{VB + FB} \quad (9)$$

A Especificidade é uma métrica que mede a habilidade do modelo de corretamente identificar os tecidos que não se encaixam em um determinado tipo dentre todos os tecidos que não são desse tipo no conjunto utilizado para previsão.

$$\text{Especificidade} = \frac{\text{Verdadeiros Negativos}}{\text{Verdadeiros Negativos} + \text{Falsos Positivos}} = \frac{VA}{VA + FA} \quad (10)$$

- **AUC** Uma outra estratégia para mensurar a qualidade de um modelo preditivo é utilizar a Área

Sob a Curva (AUC). Ela é determinada a partir da Curva de Resposta Característica (Curva ROC) e representa a capacidade de discriminação do modelo em questão.

Para melhor compreender o conceito da AUC, pode-se utilizar o exemplo de uma classificação binária, como a apresentada pela Tabela 4. Neste cenário, são apresentadas ao modelo duas observações simultaneamente: uma da classe positiva e outra da classe negativa. A AUC representa a probabilidade de o modelo identificar corretamente ambas as observações. Então, valores próximos de 1 indicam um modelo de altíssima capacidade de discriminação, ao passo que valores próximos a 0,5 indicam um modelo que faz alocação aleatória de classes.

- **Precisão**

A partir de todas as vezes que o modelo previu que um tecido é do tipo T , a Precisão mede a proporção de vezes em que essa previsão foi correta.

$$\text{Precisão} = \frac{\text{Previsões Positivas Corretas}}{\text{Todas as Previsões Positivas}} = \frac{VB}{FA + VB} \quad (11)$$

- **Score F1** O Score F1 é uma métrica que avalia o equilíbrio do modelo entre precisão e sensibilidade. Ou seja, no caso de um modelo que classifica se uma célula é cancerígena ou não, o Score F1 quantifica a relação entre não deixar nenhuma amostra cancerígena passar despercebida e, não classificar uma amostra como cancerígena quando não se trata de uma. Exposto na Equação 12, trata-se de uma média harmônica.

$$F_1 = 2 \cdot \frac{\text{Precisão} \cdot \text{Sensibilidade}}{\text{Precisão} + \text{Sensibilidade}} \quad (12)$$

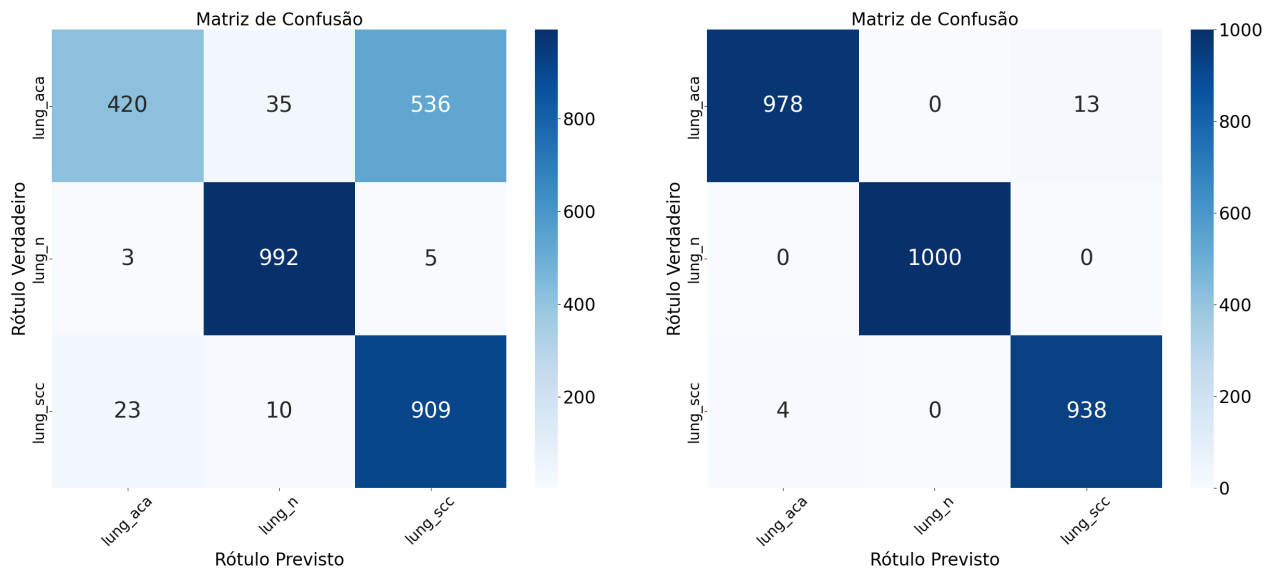
Essas métricas, em conjunto, permitem uma visualização consolidada do desempenho da rede neural em uma série de situações, permitindo ponderar quais são as prioridades para uma determinada tarefa e qual arquitetura utilizar.

4 Resultados

Nesta seção, serão apresentados os resultados obtidos pela aplicação do algoritmo de busca aleatória utilizando redes neurais convolucionais e, também, *transfer learning* via ResNet50. Serão apresentadas as matrizes de confusão, área sob a curva (AUC), especificidade, sensibilidade, precisão e Escore F1.

4.1 LC25000

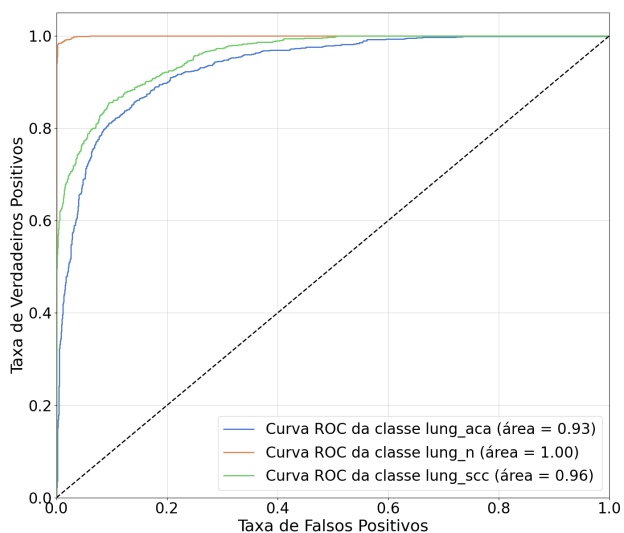
- **3 Classes** | Classificação do *patch* como **Adenocarcinoma Pulmonar** (lung_aca), **Carcinoma de Células Escamosas do Pulmão** (lung_ssc) ou **Tecido de Pulmão Benigno** (lung_n)



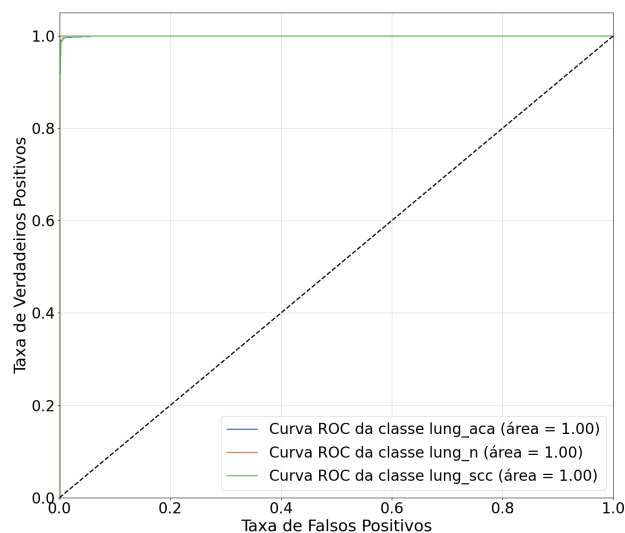
(a) CNN Simple: Matriz de Confusão de 3 classes

(b) ResNet50: Matriz de Confusão de 3 classes

Figura 20: Comparação entre duas matrizes de confusão de uma CNN Simple e ResNet50 para o conjunto LC25000



(a) CNN Simples: Curva ROC e AUC



(b) ResNet50: Curva ROC e AUC

Figura 21: Comparação entre curvas ROC e área sob a curva (AUC) de uma CNN Simples e ResNet50 para o conjunto LC25000

	Especificidade	Sensibilidade	Precisão	Escore F1
lung_aca	0.98	0.42	0.94	0.58
lung_n	0.97	0.99	0.95	0.97
lung_scc	0.72	0.96	0.62	0.76

Tabela 5: Métricas das classes com CNN Simples (LC25000)

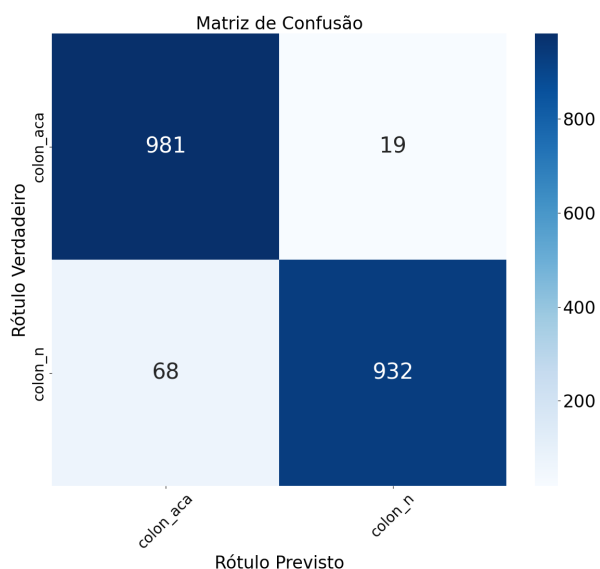
	Especificidade	Sensibilidade	Precisão	Escore F1
lung_aca	0.99	0.98	0.99	0.99
lung_n	1.00	1.00	1.00	1.00
lung_scc	0.99	0.99	0.98	0.99

Tabela 6: Métricas das classes com ResNet50 (LC25000)

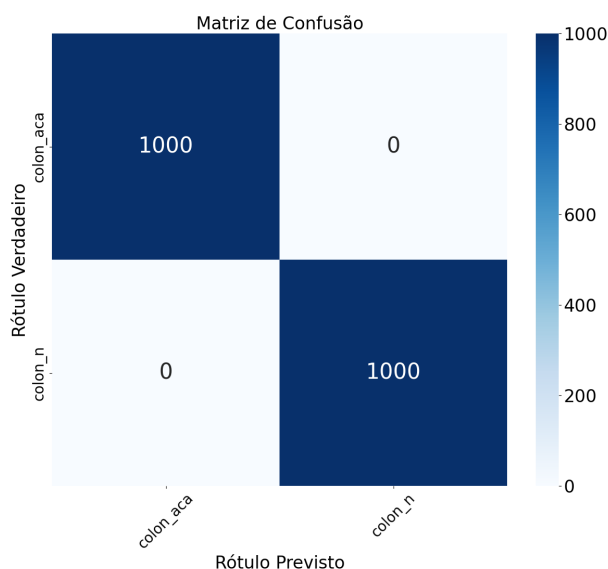
	CNN Simples	ResNet50
Acurácia de Validação	0.79	0.99

Tabela 7: Comparação da acurácia com uma CNN Simples e ResNet50 (LC25000)

- **2 Classes** | Classificação do *patch* como **Adenocarcinoma de Cólon** (colon_aca) ou **Tecido de Cólon Benigno** (colon_n)

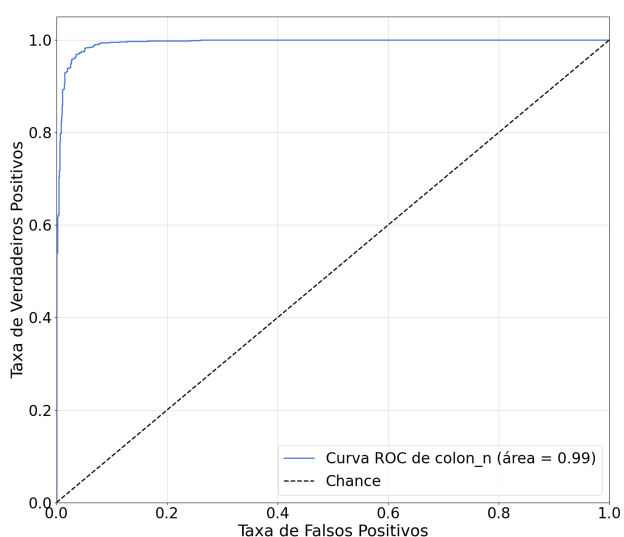


(a) CNN Simples: Matriz de Confusão de 2 classes

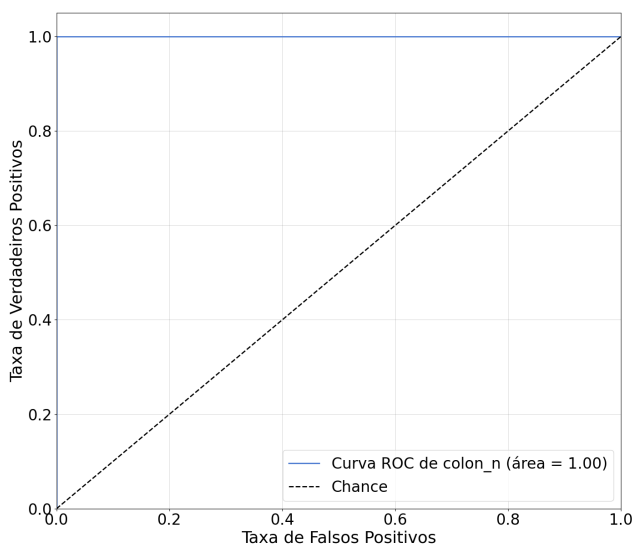


(b) ResNet50: Matriz de Confusão de 2 classes

Figura 22: Comparação entre duas matrizes de confusão de uma CNN Simples e ResNet50 para o conjunto LC25000



(a) CNN Simples: Curva ROC e AUC



(b) ResNet50: Curva ROC e AUC

Figura 23: Comparação entre curvas ROC e área sob a curva (AUC) de uma CNN Simples e ResNet50 para o conjunto LC25000

	Especificidade	Sensibilidade	Precisão	Escore F1
colon_aca	0.93	0.98	0.93	0.95
colon_n	0.98	0.93	0.98	0.95

Tabela 8: Métricas das classes com CNN Simples (LC25000 - 2 Classes)

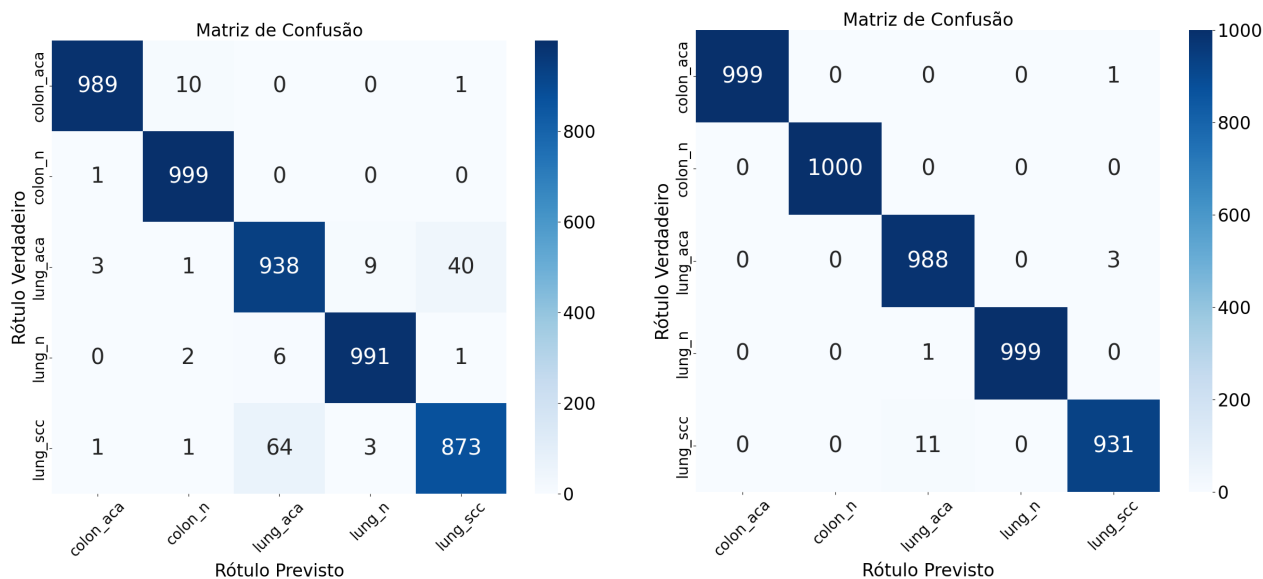
	Especificidade	Sensibilidade	Precisão	Escore F1
colon_aca	1.00	1.00	1.00	1.00
colon_n	1.00	1.00	1.00	1.00

Tabela 9: Métricas das classes com a ResNet50 (LC25000 - 2 Classes)

	CNN Simples	ResNet50
Acurácia de Validação	0.95	1.00

Tabela 10: Comparação da acurácia com uma CNN Simples e ResNet50 (LC25000 - 2 Classes)

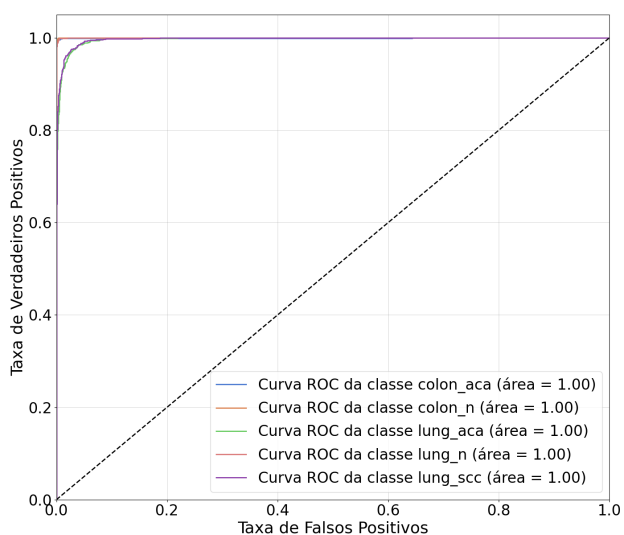
- **5 Classes** | Classificação do *patch* como **Adenocarcinoma Pulmonar** (lung_aca), **Carcinoma de Células Escamosas do Pulmão** (lung_scc), **Tecido de Pulmão Benigno** (lung_n), **Adenocarcinoma de Cólon** (colon_aca) ou **Tecido de Cólon Benigno** (colon_n)



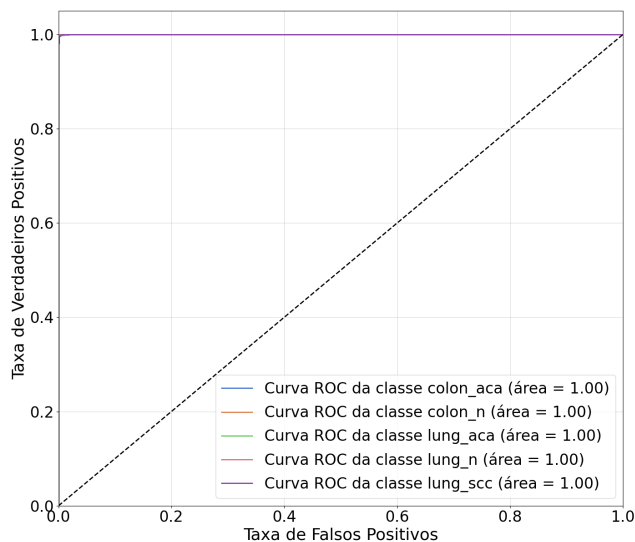
(a) CNN Simples: Matriz de Confusão de 5 classes

(b) ResNet50: Matriz de Confusão de 5 classes

Figura 24: Comparação entre duas matrizes de confusão de uma CNN Simples e ResNet50 para o conjunto LC25000



(a) CNN Simples: Curva ROC e AUC



(b) ResNet50: Curva ROC e AUC

Figura 25: Comparação entre curvas ROC e área sob a curva (AUC) de uma CNN Simples e ResNet50 para o conjunto LC25000

	Especificidade	Sensibilidade	Precisão	Escore F1
colon_aca	0.99	0.98	0.99	0.99
colon_n	0.99	0.99	0.98	0.99
lung_aca	0.98	0.94	0.93	0.93
lung_n	0.99	0.99	0.98	0.98
lung_scc	0.98	0.92	0.95	0.94

Tabela 11: Métricas das classes com CNN Simples (LC25000 - 5 Classes)

	Especificidade	Sensibilidade	Precisão	Escore F1
colon_aca	1.00	0.99	1.00	0.99
colon_n	1.00	1.00	1.00	1.00
lung_aca	0.99	0.99	0.98	0.99
lung_n	1.00	0.99	1.00	0.99
lung_scc	0.99	0.98	0.99	0.99

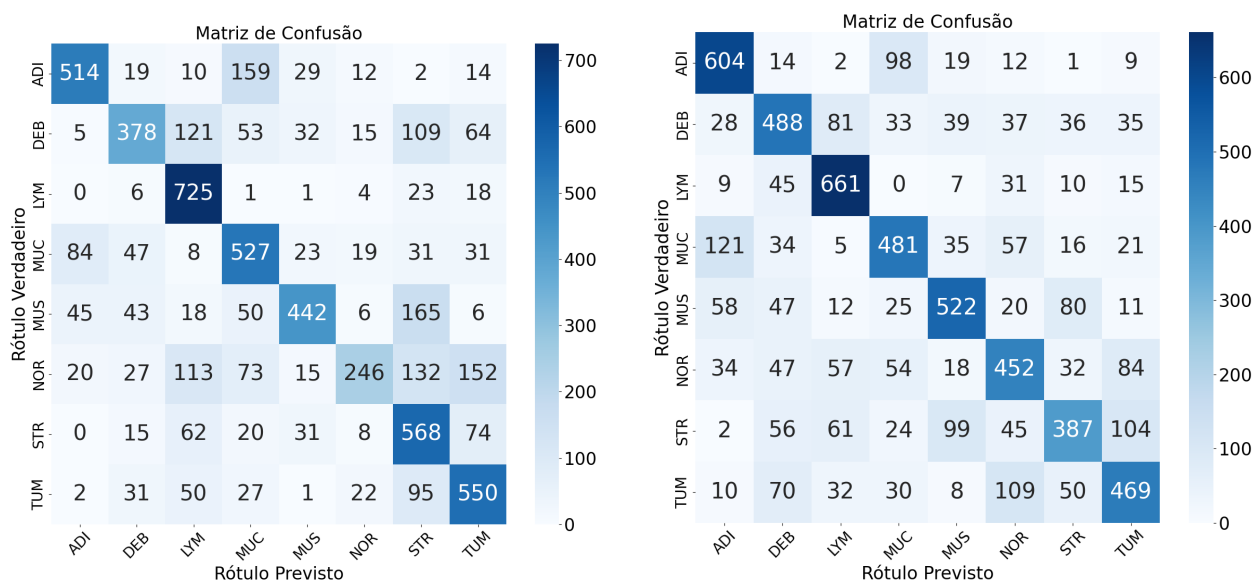
Tabela 12: Métricas das classes com ResNet50 (LC25000 - 5 Classes)

	CNN Simples	ResNet50
Acurácia de Validação	0.97	0.99

Tabela 13: Comparação da acurácia com uma CNN Simples e ResNet50 (LC25000 - 5 Classes)

4.2 Câncer Gástrico

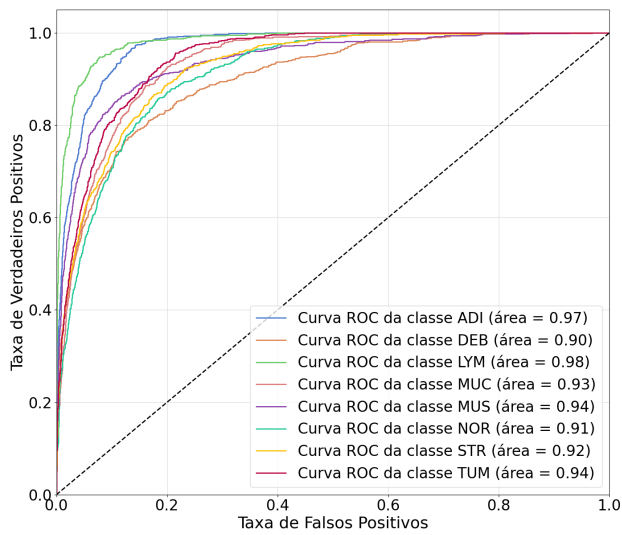
- **8 Classes** | Classificação do *patch* como **Tecido Adiposo** (ADI), **Detritos Celulares** (DEB), **Muco** (MUC), **Tecido Muscular** (MUS), **Tecido Linfóide** (LYM), **Tecido Estromal** (STR), **Mucosa normal** (NOR) ou **Tecido de Epitélio Tumoral** (TUM)



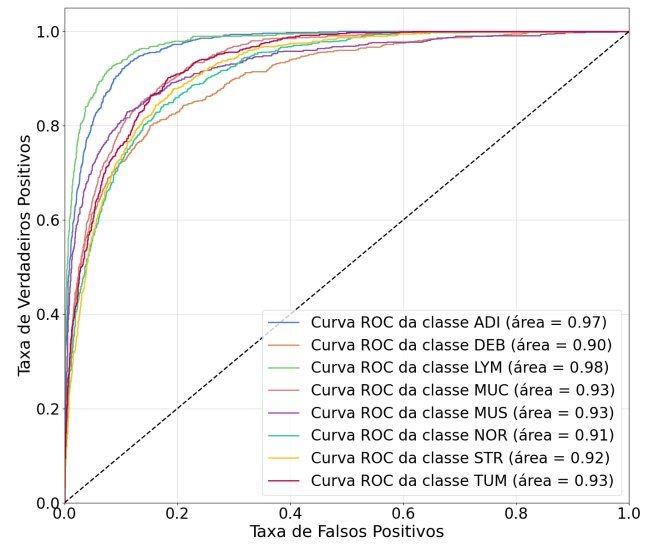
(a) CNN Simples: Matriz de Confusão de 8 classes

(b) ResNet50: Matriz de Confusão de 8 classes

Figura 26: Comparação entre duas matrizes de confusão de uma CNN Simples e ResNet50 para o conjunto de Câncer Gástrico



(a) CNN Simples: Curva ROC e AUC



(b) ResNet50: Curva ROC e AUC

Figura 27: Comparação entre curvas ROC e área sob a curva (AUC) de uma CNN Simples e ResNet50 para o conjunto de Câncer Gástrico

	Especificidade	Sensibilidade	Precisão	Escore F1
ADI	0.97	0.67	0.76	0.71
DEB	0.96	0.48	0.66	0.56
LYM	0.92	0.93	0.65	0.76
MUC	0.92	0.68	0.57	0.62
MUS	0.97	0.57	0.77	0.65
NOR	0.98	0.31	0.74	0.44
STR	0.89	0.73	0.50	0.59
TUM	0.93	0.70	0.60	0.65

Tabela 14: Métricas das classes com uma CNN Simples (Câncer Gástrico - 8 Classes)

	Especificidade	Sensibilidade	Precisão	Escore F1
ADI	0.95	0.79	0.69	0.74
DEB	0.94	0.62	0.60	0.61
LYM	0.95	0.84	0.72	0.78
MUC	0.95	0.62	0.64	0.63
MUS	0.95	0.67	0.69	0.68
NOR	0.94	0.58	0.59	0.58
STR	0.95	0.49	0.63	0.55
TUM	0.94	0.60	0.62	0.61

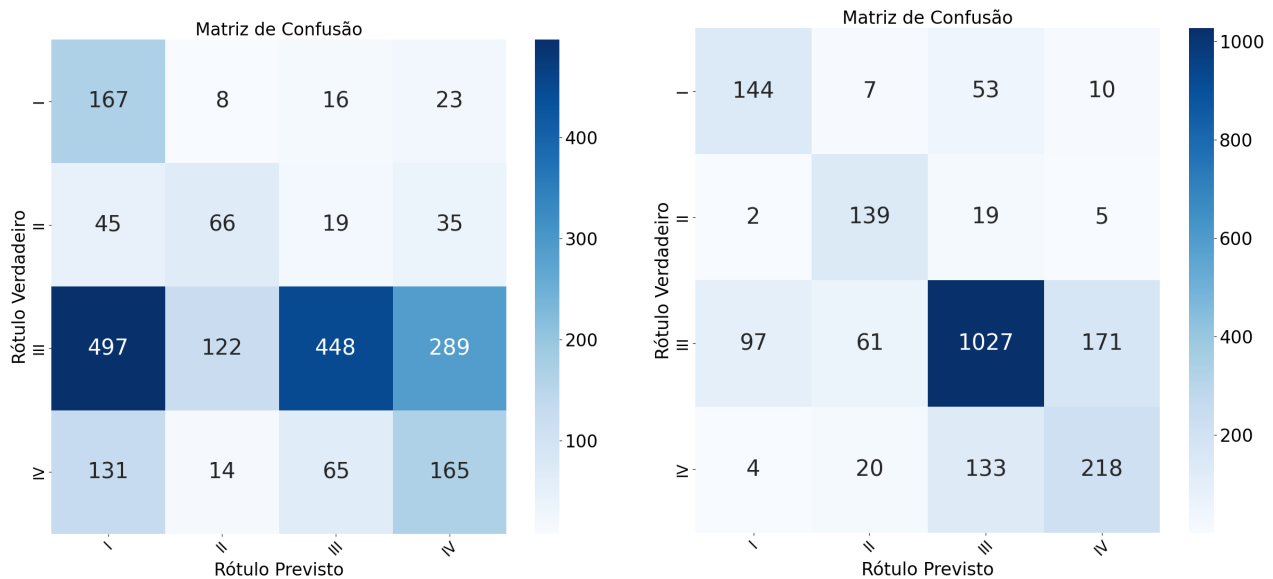
Tabela 15: Métricas das classes com a ResNet50 (Câncer Gástrico - 8 Classes)

	CNN Simples	ResNet50
Acurácia de Validação	0.63	0.65

Tabela 16: Comparação da acurácia com uma CNN Simples e ResNet50 (Câncer Gástrico - 8 Classes)

4.3 Carcinoma Seroso de Alto Grau - HGSOc

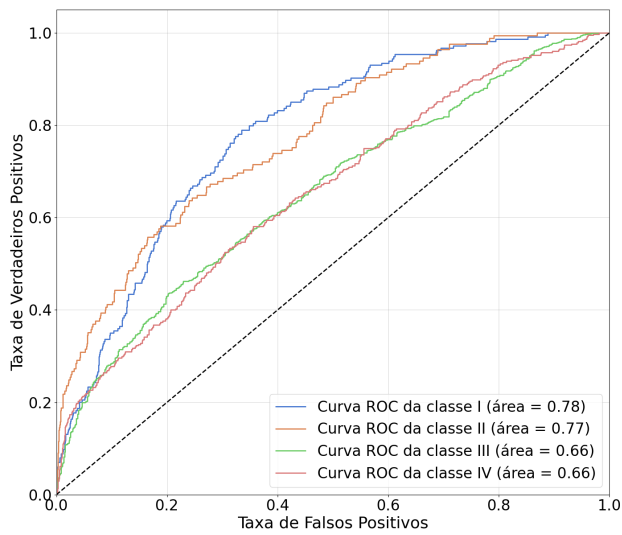
- **4 Classes** | Classificação do *patch* como **Estágio I (I)**, **Estágio II (II)**, **Estágio III (III)** e **Estágio IV (IV)**



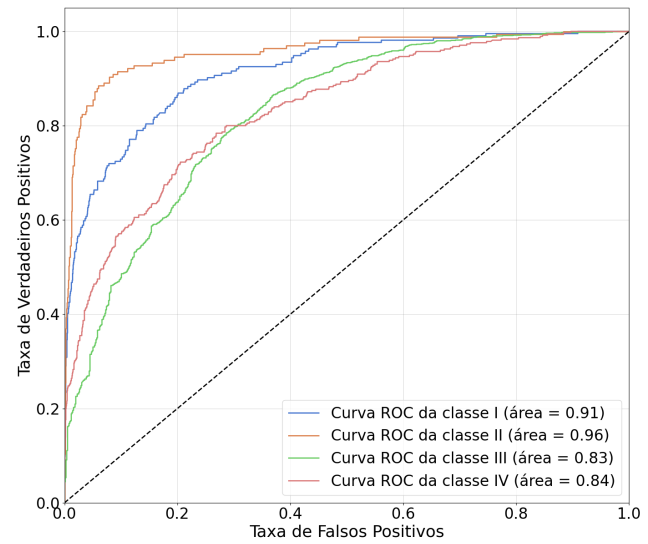
(a) CNN Simples: Matriz de Confusão de 4 classes

(b) ResNet50: Matriz de Confusão de 4 classes

Figura 28: Comparação entre duas matrizes de confusão de uma CNN Simples e ResNet50 para o conjunto de HGSOc



(a) CNN Simples: Curva ROC e AUC



(b) ResNet50: Curva ROC e AUC

Figura 29: Comparação entre curvas ROC e área sob a curva (AUC) de uma CNN Simples e ResNet50 para o conjunto de HGSOC

	Especificidade	Sensibilidade	Precisão	Escore F1
I	0.64	0.78	0.19	0.31
II	0.92	0.40	0.31	0.35
III	0.86	0.33	0.81	0.47
IV	0.80	0.44	0.32	0.37

Tabela 17: Métricas das classes com uma CNN Simples (HGSOC - 4 Classes)

	Especificidade	Sensibilidade	Precisão	Escore F1
I	0.94	0.67	0.58	0.62
II	0.95	0.84	0.61	0.70
III	0.72	0.75	0.83	0.79
IV	0.89	0.58	0.53	0.55

Tabela 18: Métricas das classes com a ResNet50 (HGSOC - 4 Classes)

	CNN Simples	ResNet50
Acurácia de Validação	0.40	0.72

Tabela 19: Comparação da acurácia com uma CNN Simples e ResNet50 (HGSOC - 4 Classes)

5 Discussão

No geral, é possível inferir o potencial de redes neurais convolucionais na classificação de lâminas histológicas junto da utilidade que o aprendizado por transferência fornece nesse propósito.

Especialmente no conjunto de amostras de tecidos pulmonares e colônicos, as arquiteturas utilizadas atingiram acurácias de 79 a 100% na determinação do tipo tecidual quando o espaço amostral apenas contia classes do mesmo domínio histológico [Figuras 20 e 22]. Na Tabela 5 é perceptível a baixa capacidade de distinção entre o Adenocarcinoma de Pulmão e Tecido Escamoso de Pulmão na CNN simples. Todavia, a ResNet50 mostrou a sua superioridade na distinção desses dois tecidos, dada a sua profundidade de 50 camadas, juntamente das *head layers*, que corretamente abstraíram as divergências morfológicas de ambos atingindo uma acurácia global de 99%. Com a junção das duas entidades histológicas — pulmão e cólon — o desempenho do modelo, apesar do aumento de classes e acentuada dissonância das entidades histológicas, permaneceu com capacidade quase completa de classificar cada tipo de tecido. Nesses três casos, a ResNet50 atingiu, em todas as métricas utilizadas, maiores valores do que as variadas arquiteturas de CNN Simples utilizadas, reforçando a capacidade da busca bayesiana de localizar o melhor modelo

As imagens do conjunto LC25000, em sua totalidade, demonstraram homogeneidade na qualidade e utilidade de cada *patch*, auxiliando no aprendizado eficiente do modelo. Ademais, o tamanho razoável de cada *patch* (768x768 pixels) pode ter contribuído para a compreensão de estruturas morfológicas de maior escala no ambiente celular, estas úteis para classificação do tipo do tecido.

O conjunto do Câncer Gástrico apresentou resultados similares para ambas arquiteturas, ainda com a ResNet50 a frente. A matrizes de confusão da Figura 26, ambas ResNet50 e CNN Simples demonstram esparsidade nas previsões. Da mesma forma, a média aritmética das sensibilidades da ResNet50, 65%, se aproxima da média aritmética das sensibilidades da CNN, 63%, evidenciando que os dois modelos se assemelham na capacidade de corretamente classificar classes diante da frequência dessa classe no conjunto. Não obstante, as classes que menos passaram despercebidas variaram entre as duas. A CNN Simples demonstrou maior capacidade de detectar os tipos **LYM**, **STR** e **TUM** (Tabela 14), enquanto a ResNet50 demonstrou maior capacidade em detectar os tipos **LYM**, **ADI** e **MUS** (Tabela 15).

O conjunto do HGSOC teve, na ResNet50 e CNN Simples, respectivamente, precisões de 83% e 81% no estágio III, contrastando com precisões de 19% na classe I (CNN Simples, 17) e 53% na

classe IV (ResNet50, 18). No contexto de severo desbalanceamento do conjunto de dados, que favorece a classe III, apesar da correção de pesos das classes realizado antes do treino do modelo, este não se mostrou suficiente para prevenir a inclinação da ResNet50 a focalizar a maioria das suas previsões para o estágio dominante. Por outro lado, a CNN Simples, apesar de ter tido uma distribuição percentual homogênea de previsões por classe, se mostrou errônea na maioria e sua maior precisão foi, também, na classe III, com precisão de 81%. Torna-se claro, portanto, que para um treinamento eficiente no conjunto de HGSOc, serão necessárias medidas adicionais para lidar com o desequilíbrio de dados. É provável que o aumento de dados (*oversampling*) das classes minoritárias através da criação de cópias com rotações e simetrias variadas elimine o favorecimento de uma classe em detrimento das outras por parte do modelo e resulte na maior compreensão de padrões morfológicos. Outra técnica promissora é o *undersampling*, em que todas as classes têm suas observações reduzidas ao mesmo número de observações da classe minoritária. Apesar de igualar as classes, esse método resulta na perda de informação, então é ideal a repetição do treino com múltiplas combinações de seleção de dados dos conjuntos majoritários. Adicionalmente, o HGSOc tem como uma de suas principais características a elevada heterogeneidade intratumoral e consequente variada morfologia [Azzalini et al., 2023], que pode dificultar a habilidade de generalização do modelo.

Sob essa óptica, a criação de um modelo baseado em *machine learning* com o intuito de estadiamento do HGSOc pode se beneficiar da junção de informações quantitativas como dados clínicos e da utilização de modelos capazes de abstrair contextos gerais da imagem como os *vision transformers*. Além disso, a utilização de um conjunto de dados com mais observações disponíveis e com maior homogeneidade na qualidade das lâminas histológicas tende a favorecer resultados mais robustos e satisfatórios, ainda no contexto de CNNs.

A normalização de *Macenko* demonstrou capacidade de uniformização de imagens de diferentes domínios em um mesmo espectro de cores. Apesar disso, resultou em histogramas RGB com padrões irregulares e que carecem de uma explicação lógica. Possivelmente a normalização de *Macenko* se destacaria em *slides* inteiros ao invés da abordagem de cada *patch* como imagem individual.

Ademais, a busca Bayesiana se mostrou eficiente em achar arquiteturas eficientes para uma dada tarefa, evidenciado nos resultados do conjunto LC25000 em suas três classificações. Desse modo, é possível priorizar outros fatores como agentes causadores dos resultados do conjunto HGSOc e de Câncer Gástrico ao invés da própria CNN. Contudo, ainda é sugerido que sejam realizadas buscas

Bayesianas com espaços amostrais maiores com maiores variações em hiperparâmetros, visto que trabalhamos com limites reduzidos na experimentação arquitetônica da busca e não variamos hiperparâmetros como otimizador e funções de ativação.

Outras potenciais melhorias além das citadas são a experimentação de outros modelos pré-treinados para o aprendizado por transferência, dado que apenas exploramos a ResNet50. Também, um maior volume de dados potencialmente otimizaria os resultados do conjunto de HGSOc e Câncer Gástrico.

Por fim, esses modelos potencialmente podem ser aplicados em sistemas computadorizados no âmbito da saúde. Após o protocolo de extração de uma lâmina histológica por meio cirúrgico, torna-se mais eficaz e consistente a classificação desses tecidos e subsequente apoio ao diagnóstico médico.

6 Conclusão

Este trabalho evidenciou o potencial das redes neurais convolucionais e do aprendizado por transferência na tarefa de classificação histológica de tecidos cancerígenos de domínios diversos que, no contexto do estudo realizado, trataram-se de pulmão, cólon, estômago e ovários. Foi consolidada uma compreensão teórica e aplicação prática das redes neurais, juntamente com métricas de avaliação de seu desempenho, refletindo a capacidade da estatística de esclarecer o contexto de cada sub tarefa realizada.

Os resultados obtidos permitiram visualizar de forma clara o potencial da aplicação de modelos computacionais baseados em aprendizado profundo na análise anatomopatológica de tecidos humanos e no suporte ao diagnóstico médico. A convergência entre redes neurais e métricas estatísticas na compreensão histológica das morfologias tumorais evidencia a maturidade crescente da ciência de dados como uma ferramenta relevante, capaz de promover novas descobertas e otimizar significativamente o processo humano envolvido na análise de lâminas histológicas e no diagnóstico clínico. Sob esse viés, limitações técnicas foram identificadas por meio da análise de dados, sugerindo medidas a serem consideradas em futuras investigações.

O estudo realizado evidencia a variedade de possibilidades futuras na interseção entre *machine learning* e o campo médico, sendo promissoras a expansão para novos experimentos com diferentes conjuntos de dados, arquiteturas computacionais, abordagens estatísticas e eventuais implementações em interfaces clínicas, como ferramentas assistivas para patologistas.

7 Agradecimentos

Agradeço ao professor Dr. Benilton Carvalho, que me orientou durante a síntese do meu primeiro trabalho científico. Tive o privilégio de trabalhar ao lado de um mestre do ensino, que me guiou nas dificuldades e, sem dúvidas, me trouxe muitos ensinamentos para futuros trabalhos — e até para a vida (tenho até uma lista com anotações). Obrigado por sua disponibilidade e por toda a paciência durante esse processo, que foi ímpar em minha vida. Por fim, vale ressaltar os novos interesses que pude descobrir através do senhor, visto que, quando iniciei a pesquisa pensando em biotecnologia em um contexto amplo, não imaginava que mergulharia no universo da inteligência artificial e da estatística — e que gostaria tanto de ambos. És inspirador.

Agradeço ao Instituto Principia e a todas as pessoas que fazem — e já fizeram — a Escola de Talentos acontecer: Marcelo Guzzo, Welington Barbosa, Ricardo Matheus, Leonardo de Lima, Rickson Mesquita, André Paschoal e todos os outros envolvidos! Não exagero ao dizer que meu ensino médio não teria sido nada do que foi sem essa oportunidade que me foi dada. Uma sequência de possibilidades se revelou para mim após o meu ingresso nessa instituição, que preza pela ciência e reconhece o nosso potencial. O impacto da Escola de Talentos em minha vida é imensurável. Uma vez E.T., para sempre E.T. Obrigado.

Agradeço à minha família, por me dar apoio e amor incondicional, e por me acompanharem de perto nessa jornada. Pai e mãe, eu não seria quem sou hoje sem vocês. O meu amor por vocês não cabe em palavras!

Agradeço à minha amiga Clara, que é a pessoa com quem posso contar para compartilhar minhas ideias mais loucas, falar de ciência, livros, sonhos, universidade e, também, das dificuldades e alegrias da vida. Nós ainda vamos criar muitas coisas magníficas!

Agradeço a todos os meus amigos, que são uma rede de apoio que me traz imensa gratidão. É incrível ter vocês para criar memórias inesquecíveis.

Por fim, agradecemos o uso do cluster Coaraci (Fapesp Proc. 2019/17874-0) e ao Center for Computing in Engineering and Sciences at Unicamp (Fapesp Proc.2013/08293-7).

Referências

- Eros Azzalini, Giorgio Stanta, Vincenzo Canzonieri, and Serena Bonin. Overview of Tumor Heterogeneity in High-Grade Serous Ovarian Cancers. *International Journal of Molecular Sciences*, 24(20):15077, October 2023. ISSN 1422-0067. doi: 10.3390/ijms242015077.
- Cides Bezerra. *UMA ABORDAGEM DE SEGMENTAÇÃO SEMÂNTICA DE ÍRIS PARA FINS BIOMÉTRICOS USANDO APRENDIZAGEM PROFUNDA*. PhD thesis, 09 2018.
- Andrew A. Borkowski, Marilyn M. Bui, L. Brannon Thomas, Catherine P. Wilson, Lauren A. DeLand, and Stephen M. Mastorides. Lung and Colon Cancer Histopathological Image Dataset (LC25000), December 2019. URL <http://arxiv.org/abs/1912.12142>. arXiv:1912.12142 [eess] version: 1.
- Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255, 2009. doi: 10.1109/CVPR.2009.5206848.
- Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition, 2015. URL <https://arxiv.org/abs/1512.03385>.
- International Agency for Research on Cancer. Cancer today. <https://gco.iarc.who.int/today/en/dataviz/tables?mode=population>, 2022. URL <https://gco.iarc.who.int/today/en/dataviz/tables?mode=population>. Data visualization tool.
- Shenghan Lou, Jianxin Ji, Huiying Li, Xuan Zhang, Yang Jiang, Menglei Hua, Kexin Chen, Kaiyuan Ge, Qi Zhang, Liuying Wang, Peng Han, and Lei Cao. A large histological images dataset of gastric cancer with tumour microenvironment annotation for AI. *Scientific Data*, 12(1):138, January 2025. ISSN 2052-4463. doi: 10.1038/s41597-025-04489-9. URL <https://www.nature.com/articles/s41597-025-04489-9>. Publisher: Nature Publishing Group.
- Marc Macenko, Marc Niethammer, J. S. Marron, David Borland, John T. Woosley, Xiaojun Guan, Charles Schmitt, and Nancy E. Thomas. A method for normalizing histology slides for quantitative analysis. In *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 1107–1110, Boston, MA, USA, June 2009. IEEE. ISBN 978-1-4244-3931-7. doi: 10.1109/ISBI.2009.5193250. URL <http://ieeexplore.ieee.org/document/5193250/>.

Steven Narod. Can advanced-stage ovarian cancer be cured? *Nature Reviews Clinical Oncology*, 13(4):255–261, April 2016. ISSN 1759-4782. doi: 10.1038/nrclinonc.2015.224. URL <https://www.nature.com/articles/nrclinonc.2015.224>. Publisher: Nature Publishing Group.