

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

André Dias Freitas

**Investigação de Normalização de Corantes para
Classificação de Imagens da Cavidade Oral**

Uberlândia, Brasil

2021

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

André Dias Freitas

**Investigação de Normalização de Corantes para
Classificação de Imagens da Cavidade Oral**

Trabalho de conclusão de curso apresentado à Faculdade de Computação da Universidade Federal de Uberlândia, Minas Gerais, como requisito exigido parcial à obtenção do grau de Bacharel em Ciência da Computação.

Orientador: Marcelo Zanchetta do Nascimento

Universidade Federal de Uberlândia – UFU

Faculdade de Computação

Bacharelado em Ciência da Computação

Uberlândia, Brasil

2021

André Dias Freitas

Investigação de Normalização de Corantes para Classificação de Imagens da Cavidade Oral

Trabalho de conclusão de curso apresentado à Faculdade de Computação da Universidade Federal de Uberlândia, Minas Gerais, como requisito exigido parcial à obtenção do grau de Bacharel em Ciência da Computação.

Marcelo Zanchetta do Nascimento
Orientador

**Prof. Dr. Bruno Augusto Nassif
Travençolo**

Msc. Dalí Freire Dias dos Santos

Uberlândia, Brasil
2021

Agradecimentos

Este trabalho foi realizado durante o período de quarentena da COVID-19, e portanto, primeiramente, agradeço à minha família, que suportou comigo esse período de dificuldade, e agradeço ao meu orientador, Marcelo Zanchetta do Nascimento, pelos esforços, tempo e paciência, empenhados a me orientar na realização deste trabalho.

Resumo

O câncer é uma das doenças com maior taxa de mortalidade do mundo. A displasia é uma lesão pré-cancerígena de difícil diagnóstico, realizado a partir de amostras histológicas, que muitas vezes não possuem uma boa relação dos corantes Hematoxilina e Eosina (H&E), o que dificulta ainda mais o diagnóstico realizado pelo especialista em histologia. Nesse trabalho foi investigado um método de normalização de corantes nas imagens histológicas. Esse método emprega uma rede neural generativa baseada em uma *U-net* para a geração das imagens e uma arquitetura *PatchGAN* para discriminação das informações. As imagens histológicas normalizadas foram avaliadas em relação a classificação das lesões. Nesses experimentos foram usados modelos CNNs baseados nas arquiteturas ResNet e AlexNet assim como modelos híbridos que empregam numa junção das CNNs com algoritmos de aprendizado de máquina. O emprego da arquitetura ResNet e o algoritmo *Random Forest* proporcionou resultados com taxa de acurácia em torno de 97% para as imagens normalizadas com o método investigado.

Palavras-chave: Redes Neurais Convolucionais, Cavidade Oral, Normalização de Corantes, Classificação, Histologia.

Lista de ilustrações

Figura 1 – Exemplos de tecidos histológicos orais nos diferentes níveis de displasia, onde as setas vermelhas indicam os núcleos celulares e as azuis indicam o epitélio (SILVA, 2019).	15
Figura 2 – Etapas de um sistema CAD comum. Fonte: (RAMADAN, 2020).	17
Figura 3 – Exemplos de normalização de corantes H&E: (a) e (c) imagens originais e (b) e (d) imagens normalizadas.	18
Figura 4 – As principais subáreas de pesquisa em IA (ROSEBROCK, 2017).	19
Figura 5 – Exemplo de uma árvore binária, onde o nó do topo é chamado de nó raiz (representado pelo elemento número 4), e os nós na base são chamados de nós folhas (representados pelos elementos números 1, 3, 5, 7).	20
Figura 6 – Exemplo de uma árvore de decisão que possui cinco nós (LOUPPE, 2015).	20
Figura 7 – Exemplo de distribuição de duas classes (vermelho e azul) em um hiperplano, e um vetor capaz de separá-los em dois grupos distintos (KOWALCZYK, 2017).	22
Figura 8 – Exemplo de classificações com kernels linear, polinomial e RBF, respectivamente (KOWALCZYK, 2017).	22
Figura 9 – Um exemplo simples de rede neural com os neurônios representados nas cores azul, verde e vermelho.	23
Figura 10 – Funcionamento de um <i>kernel</i> sobre uma região da imagem denominada <i>patch</i> (ROSEBROCK, 2017).	24
Figura 11 – Exemplos de transformações por diferentes tipos de <i>kernels</i> (ROSEBROCK, 2017).	24
Figura 12 – Exemplos de <i>max pooling</i> e <i>average pooling</i>	25
Figura 13 – Arquitetura GAN simplificada (BISSOTO, 2019).	26
Figura 14 – Exemplos de tecidos histológicos da cavidade oral: (a) saudável, (b) displasia leve, (c) displasia moderada e (d) displasia severa.	30
Figura 15 – Modelo desenvolvido neste trabalho para avaliação das imagens histológicas.	30
Figura 16 – Exemplo do <i>cropping</i> realizado em que dois <i>patches</i> (representados nas cores vermelho e azul) são produzidos para cada imagem.	31
Figura 17 – Modelo da GAN para a normalização Stain-to-Stain (SALEHI; CHALECHALE, 2020)	31
Figura 18 – Modelo original da arquitetura de uma rede AlexNet (YU LAURENCE T. YANG, 2021).	34

Figura 19 – Bloco residual de uma ResNet. $F(x) + x$ é a função residual em que $F(x)$ representa as camadas não-lineares empilhadas e x representa a identidade necessária pra tornar a entrada equivalente à saída (YU LAURENCE T. YANG, 2021).	35
Figura 20 – Exemplo de uma matriz de confusão.	36
Figura 21 – Exemplo de normalizações de cada classe usando o método STST.	38
Figura 22 – Resultados do treinamento do modelo AlexNet.	39
Figura 23 – Resultados do treinamento do modelo ResNet50.	39
Figura 24 – Função ECDF para as características extraídas do modelo AlexNet	42
Figura 25 – Função ECDF para as características extraídas do modelo ResNet50	42

Lista de tabelas

Tabela 1 – Resultados (%) da classificação realizada com as imagens originais, sem investigação da normalização.	40
Tabela 2 – Resultados (%) da classificação com as imagens normalizadas empregadas apenas na fase de teste.	40
Tabela 3 – Classificação (%) com as imagens normalizadas empregadas na fase de treinamento e as imagens originais avaliadas na fase de teste.	40
Tabela 4 – Resultados (%) obtidos com as imagens normalizadas para as fases de treinamento e teste.	41
Tabela 5 – Valores das métricas para classificação dos modelos híbridos utilizando as imagens originais sem aplicação da normalização.	43
Tabela 6 – Resultados da classificação dos modelos híbridos utilizando as imagens normalizadas.	43

Lista de abreviaturas e siglas

CAD	Computer Aided Diagnosis
H&E	Hematoxilina e Eosina
IA	Inteligência Artificial
ANN	Artificial Neural Network
CNN	Convolutional Neural Network
RF	Random Forest
SVM	Support Vector Machine
GAN	Generative Adversarial Network
RGB	Red, Green, Blue
TIFF	Tagged Image File Format
ML	Machine Learning
FC	Fully Connected
ReLU	Rectified Linear Unit
STST	Stain-to-Stain
SC	Skip Connection
DA	Data Augmentation
ECDF	Empirical Cumulative Distribution Function

Sumário

1	INTRODUÇÃO	11
1.1	Objetivos	13
1.1.1	Geral	13
1.1.2	Específicos	13
2	FUNDAMENTAÇÃO TEÓRICA	14
2.1	Lesões da Cavidade Bucal	14
2.2	Histologia	15
2.3	Sistema de Apoio ao Diagnóstico	16
2.4	Normalização de Corantes H&E	16
2.5	Inteligência Artificial	17
2.5.1	Aprendizado de Máquina	18
2.5.1.1	Florestas Aleatórias	19
2.5.1.2	Máquina de Vetores de Suporte	21
2.5.2	Redes Neurais Convolucionais	23
2.5.3	Redes Adversárias Generativas	25
2.6	Estado da Arte	27
3	METODOLOGIA	29
3.1	Banco de Imagens	29
3.2	Classificação de Lesões em Imagens Histológicas Normalizadas	29
3.2.1	Pré-Processamento das Imagens	30
3.2.2	Normalização com Rede Adversária Generativa	31
3.2.3	Redes Neurais Convolucionais	33
3.2.3.1	AlexNet	33
3.2.3.2	ResNet	34
3.2.4	Algoritmos de Aprendizado de Máquinas	35
3.3	Avaliação de Desempenho das Abordagens	36
4	RESULTADOS	37
4.1	Normalização dos Corantes H&E com o Método STST	37
4.2	Investigação das Arquiteturas de Redes Neurais Convolucionais	37
4.3	Investigação dos Algoritmos de Aprendizado de Máquina	41
5	CONCLUSÃO	44
5.1	Desafios Encontrados	44

5.2	Trabalhos Futuros	45
	REFERÊNCIAS	46

1 Introdução

O câncer é uma das doenças que mais matam em todo o mundo, sendo responsável por mais de 9 milhões de mortes por ano. É uma doença genética, isto é, que surge a partir de uma mutação no DNA. Essa mutação pode ser herdada dos pais que possuem DNA anômalo, ou pode ser desenvolvida ao longo da vida pela exposição à agentes carcinógenos, como a radiação e o cigarro (INCA, 2019).

O processo de carcinogênese, ou seja, o processo de formação do câncer, é um processo lento e pode ser dividido em três principais estágios: o estágio de iniciação, quando os genes sofrem ação de algum agente cancerígeno, modificando esses genes. Nesse primeiro estágio ainda não é possível identificar clinicamente o tumor. O segundo estágio é o de promoção, quando uma célula geneticamente modificada no estágio de iniciação é afetada por agentes carcinógenos classificados como oncopromotores, e de forma lenta e gradual, a célula se transforma em uma célula maligna. Por fim ocorre o estágio de progressão, onde acontece a multiplicação descontrolada e irreversível das células alteradas e, já instaladas no corpo, podem evoluir até o surgimento das primeiras manifestações clínicas da doença (INCA, 2019). Desse modo, torna-se importante não apenas desenvolver métodos para o tratamento da doença, mas também para o diagnóstico precoce dela (JOTHI; RAJAM, 2017).

A displasia, uma lesão pré-câncer, se caracteriza como a ocorrência de anomalias relacionadas ao desenvolvimento de um órgão ou tecido, intimamente relacionadas a alterações genéticas e agressão continuada, alterando o tamanho, cor ou o formato da célula de forma não natural (FONSECA-SILVA et al., 2016; HAM; CORMACK, 1983). É um tipo de anormalidade que pode evoluir rapidamente e se tornar um tumor maligno. A análise e identificação de células displásicas é importante para auxiliar o especialista no processo de um diagnóstico precoce. Isso permite que o paciente possa receber o tratamento correto antes que a lesão torne-se mais grave ou possa evoluir para um câncer, interrompendo o desenvolvimento da doença antes mesmo do estágio de iniciação (SILVA, 2019).

Para realizar a análise e diagnóstico de células displásicas, são retiradas amostras de tecido do paciente e são coradas com Hematoxilina e Eosina (H&E) para visualizar as células e suas estruturas. A hematoxilina é fixada a núcleos celulares e ribossomos, o que lhes atribui cores roxeadas e a eosina cora citoplasmas, fibras musculosas, estruturas extracelulares e células vermelhas do sangue em tons róseos. Após esse processo, as imagens são observadas por meio de microscópio e podem ser digitalizadas para futura análise. Antigamente, essa análise era feita somente por observações do próprio especialista, realizado de forma "manual". Dessa maneira a análise dessas imagens histológicas depende diretamente da qualidade dos corantes utilizados, sendo que podem sofrer di-

ferentes reações químicas dependendo do ambiente em que são utilizados, alterando os resultados da qualidade das imagens (MICHAEL et al., 2014; TOSTA et al., 2019). Isso é uma problemática, pois desse modo a coloração de cada amostra poderá ser diferente, e a avaliação dos especialistas na maioria das vezes não é unânime. Além disso, o nível de experiência e carga de trabalho do especialista podem influenciar a qualidade da análise das regiões da imagem (SILVA, 2019).

O diagnóstico dos diversos estágios das células displásicas geralmente é uma tarefa muito difícil. Nas últimas décadas, sistemas de apoio ao diagnóstico (do inglês, *computer-aided diagnosis* - CAD) vêm contribuindo fornecendo uma segunda opinião ao especialista durante a análise do estágio dessa lesão. Um sistema CAD possui diversas etapas para análise de uma imagem histológica. As principais etapas são pré-processamento, segmentação, extração de características e classificação (ROBERTO et al., 2021; LI et al., 2018). A etapa de pré-processamento busca melhorar as informações presentes na imagem minimizando ruídos que porventura existam no processo de aquisição ou digitalização da imagem. A segmentação tem como objetivo identificar as diferentes estruturas celulares das amostras, o que é um grande obstáculo em análises computacionais, visto que essa tarefa depende diretamente da qualidade da imagem original e da remoção de ruídos realizada na etapa de pré-processamento. A extração de características permite a comparação quantitativa de imagens histológicas, e realizar uma boa extração de características implica diretamente na otimização da classificação. Por fim, a etapa de classificação utiliza os resultados das etapas anteriores para identificar corretamente as classes do problema especificado, como por exemplo, a severidade das lesões da displasia (GONZALEZ; WOODS, 2018).

Das diversas etapas de um CAD, o estágio inicial é extremamente importante para o sucesso das etapas subsequentes do sistema. Dentre os métodos utilizados, as técnicas de normalização apresentam-se como uma estratégia importante para o ajuste de cores das imagens histológicas. Ela deve ser feita de forma a manter um bom contraste com a preservação de todas as informações de origem na imagem processada. Diferentes estratégias de normalização de coloração têm sido propostas para reduzir a inconsistência dos tecidos corados em sistemas automatizados (SALEHI; CHALECHALE, 2020). Na literatura, esses métodos de normalização são classificados em correspondência de histogramas, transferência de cores, correspondência espectral e aprendizado generativo (TOSTA et al., 2019; BENTAIEB; HAMARNEH, 2017). As diversas técnicas de normalização restringem a correspondência a ser realizada com base em uma imagem de referência, que, se não forem representativas podem resultar em erros de normalização. A abordagem baseada no aprendizado trata o problema da normalização como um problema de transferência de estilo, usando redes neurais adversárias generativas (do inglês, *generative adversarial networks* - GANs) com a tarefa auxiliar de um classificador. Investigar a influência que esses métodos de normalização exercem em etapas de classificação de sistemas CAD para

lesões de displasia ainda é um desafio para a área.

Com base nessas argumentações, esse estudo apresenta a investigação de métodos que realizam a normalização dos corantes H&E para imagens histológicas de displasia. Pretende-se ainda analisar o comportamento de métodos de normalização na classificação dos estágios das lesões provocadas pela displasia.

1.1 Objetivos

1.1.1 Geral

Este projeto tem como objetivo investigar o método de normalização de corantes H&E denominado *Stain-to-Stain Translation* (STST) para a classificação de imagens histológicas da cavidade oral.

1.1.2 Específicos

- Implementação do modelo denominado *Stain-to-Stain Translation* (STST) baseado em uma arquitetura de rede GAN para a normalização de corantes H&E em imagens histológicas;
- Investigação das arquiteturas de redes convolucionais AlexNet e ResNet para classificação de imagens histológicas da cavidade oral;
- Desenvolvimento de novas abordagens híbridas baseadas nos modelos CNNs e algoritmos de aprendizagem de máquina, *Random Forest* (RF) e *Support Vector Machine* (SVM), para classificação das lesões da cavidade oral;
- Estudo do desempenho da técnica de normalização na classificação dos diversos níveis das lesões das imagens da cavidade oral;

2 Fundamentação Teórica

Neste capítulo serão expostos conceitos fundamentais para o entendimento deste trabalho. Abordando-se temas de diferentes áreas como Computação e Biologia, estes temas seguem apresentados neste capítulo de forma introdutória.

2.1 Lesões da Cavidade Bucal

As anormalidades pré-cancerosas são lesões que podem levar ao desenvolvimento de um câncer. Nesse estágio, as células ainda não sofreram mutações, o que torna difícil a diferenciação entre as células saudáveis das células com essas anormalidades. O desenvolvimento ainda pode ser bastante lento, demorando algum tempo até que possam ser notadas transformações mais severas (SILVA, 2019).

Como principais lesões e tipos de anormalidades têm-se as metaplasias, hiperplasias e displasias. As metaplasias são lesões que causam a substituição de células sensíveis a algum tipo de estresse por células mais resistentes. Nesse processo, características importantes das células originais podem ser perdidas e o estresse causado pela troca das células pode resultar em uma transformação maligna da célula. Um outro estágio é a hiperplasia, que é um processo natural do aumento do número de células em determinado lugar do corpo. Um exemplo é a proliferação epitelial mamária que ocorre na puberdade. Apesar de ser natural, as hiperplasias podem não ser reguladas corretamente pelo corpo, causando o surgimento de massas celulares anormais com potencial maligno que podem evoluir para o câncer. Já o terceiro processo, a displasia, é um tipo de lesão encontrado geralmente em tecidos epiteliais, caracterizada pela transformação morfológica das células, como alterações no tamanho, no formato e nas características internas das células. Ela pode ser classificada como lesões leves, moderadas ou severas, de acordo com a intensidade das alterações nas células do epitélio (ver Figura 1).

As displasias apresentam chances de 6% a 36% de evoluírem para um câncer, e agentes externos como o álcool e o tabaco podem contribuir para a evolução das displasias orais. O câncer de células escamosas desenvolvido a partir de displasias orais têm taxas de mortalidade consideráveis, portanto se faz importante a identificação precoce e a caracterização da severidade das lesões (SILVA, 2019).

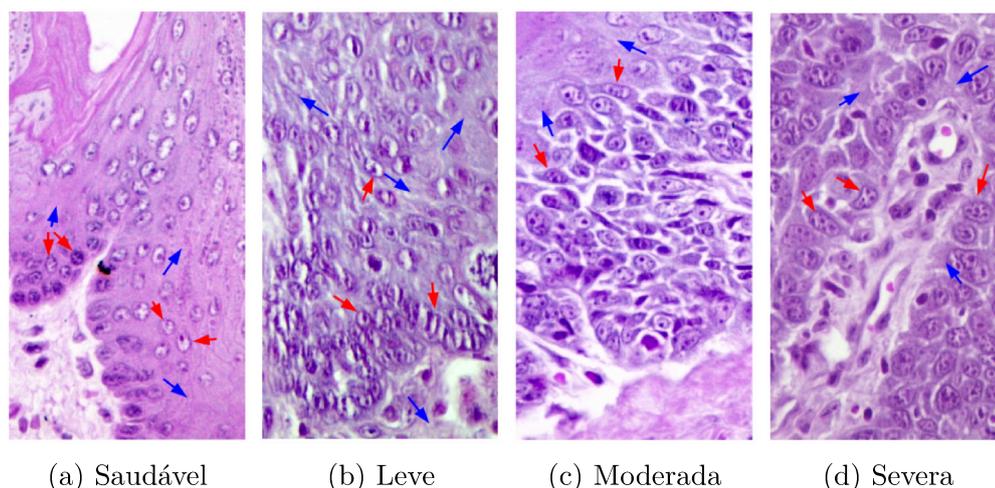


Figura 1 – Exemplos de tecidos histológicos orais nos diferentes níveis de displasia, onde as setas vermelhas indicam os núcleos celulares e as azuis indicam o epitélio (SILVA, 2019).

2.2 Histologia

A histologia pode ser definida como o estudo de tecidos biológicos, isto é, tecidos humanos, de animais, assim como tecidos vegetais. Essa área busca investigar como esses tecidos se organizam para constituírem os órgãos. Dessa maneira, são estudados as células que formam os tecidos, os órgãos e os sistemas de órgãos. Devido a pequena dimensão das células, faz-se necessário a utilização de microscópios, possibilitando analisar seu estado natural. Para realizar a análise das informações presentes no tecido e as observações com uso de um microscópio, os tecidos devem ser preparados. Desta maneira, os tecidos devem ser submetidos a diversas etapas para coloração do tecidos (SILVA, 2019).

Entre essas etapas, na fase de fixação, os tecidos são imersos em soluções químicas com o objetivo de retardar a decomposição e preservar sua estrutura e composição molecular. No processo de desidratação, aplica-se o material em composto de álcool e xilol para se tornar transparente. A inclusão tem objetivo de distinguir os componentes que constituem o tecido através de parafina, de modo que seja enrijecido para ser seccionado em cortes finos. Já etapa de microtomia realiza o corte do bloco em seções uniformes com o auxílio de um micrótomo. Após isso, os cortes são colocados sobre lâminas de vidro para, posteriormente, serem corados. Finalmente, tem-se a montagem e coloração, em que os tecidos são tratados com corantes que diferenciam os elementos básicos e ácidos das células. Geralmente, os corantes baseados em H&E são os mais utilizados. A eosina cora o citoplasma e o colágeno em cor de rosa, enquanto a hematoxilina cora os núcleos celulares e outros componentes ácidos de roxo ou azul. Ao fim da coloração, uma lâmina de vidro é posicionada acima do corte de tecido com a finalidade de protegê-la.

Os tecidos são digitalizados com o uso de microscópios equipados com câmeras para aquisição de imagens desses tecidos. Com as câmeras de alta resolução torna-se possível

a digitalização de imagens, a fim de serem usadas em computadores para análise por meio de sistemas CAD. Os objetos que não são visíveis à olho nu podem ser visualizados por uma câmera, dada sua alta resolução. Esses sistemas também são úteis para estudar células por períodos longos, porque usam luz de baixa intensidade e evitam o dano celular que pode resultar de uma iluminação intensa. Geralmente as imagens são armazenadas em arquivos de imagens com quantização de 8 bits no modelo de cores vermelho, verde e azul (do inglês: red, green and blue - RGB).

2.3 Sistema de Apoio ao Diagnóstico

Um sistema de apoio ao diagnóstico (do inglês, *computer aided diagnosis* - CAD) pode ser considerado como uma ferramenta especializada, que tem a função de auxiliar o médico especialista a fazer o diagnóstico do paciente. O sistema auxilia o médico obtendo informações adicionais para dar uma "segunda opinião", a qual contribui para a decisão final do diagnóstico dada pelo especialista (JUNIOR, 2018). A resposta do computador é útil, dado que a decisão do especialista está sujeita à variações provocadas por fatores intra e interpessoais.

Os sistemas CAD podem ter duas aplicações, a primeira é o auxílio à detecção de lesões, a partir da detecção de padrões anormais nas imagens de amostra. A segunda aplicação é o auxílio ao diagnóstico, realizado a partir da classificação das características identificadas nas imagens em padrões normais ou anormais (JUNIOR, 2018). Em sua maioria, os sistemas CAD utilizam técnicas de duas áreas de pesquisa computacional: visão computacional, que estuda a segmentação e extração de características de imagens, e inteligência artificial, que estuda métodos capazes de reconhecer padrões e realizar classificações.

Apesar das principais áreas de atuação dos sistemas CAD atualmente serem as áreas de mamografia e diagnóstico de outros tecidos epiteliais, eles também podem ser aplicados a diversas outras modalidades de obtenção de imagens, como radiografia, tomografia, ressonância magnética, ultra-sonografia e medicina nuclear (JUNIOR, 2018). Na Figura 2 é apresentado um diagrama das etapas empregadas em um sistema CAD para classificação de lesões em imagens médicas.

2.4 Normalização de Corantes H&E

Os sistemas de apoio ao diagnóstico (ou sistemas CAD) são capazes de realizar quatro etapas da análise e processamento de imagens histológicas: pré-processamento, segmentação, extração de características e classificação. A normalização de corantes ocorre na etapa de pré-processamento, onde também são feitos outros aperfeiçoamentos nas imagens, como a eliminação de ruídos.

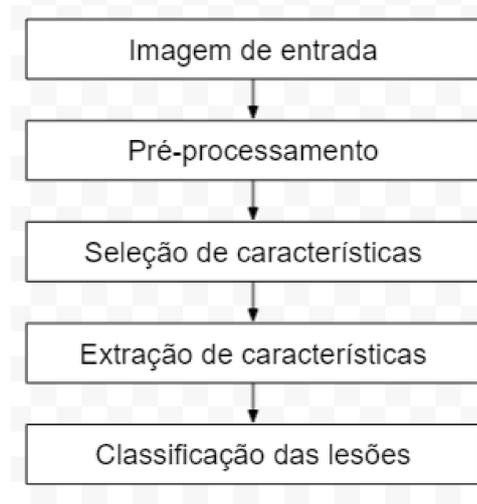


Figura 2 – Etapas de um sistema CAD comum. Fonte: (RAMADAN, 2020).

O processo de normalização, no contexto da classificação de imagens histológicas, é responsável por nivelar, isto é, normalizar, os canais de cores da imagem de forma a manter um bom contraste com a preservação de todas as características e informações da imagem original (SALEHI; CHALECHALE, 2020). Na Figura 3 é representado um conjunto (pares) de imagens originais, sem normalização, e suas respectivas versões normalizadas. O resultado da normalização são imagens com coloração estatisticamente semelhantes. Dessa forma, os sistemas classificadores conseguem compreender mais claramente as mudanças de cor nas imagens que representam as diferentes estruturas das células e tecidos, para o contexto de imagens histológicas.

Existem diferentes métodos para realizar a normalização dos corantes de imagens histológicas. A maioria das técnicas clássicas de normalização de imagens usam apenas uma imagem de referência selecionada por um especialista, que a classifica como um bom exemplo para classificação e tem um efeito significativo no resultado dos métodos. Esses métodos de normalização são classificados em correspondência de histogramas, transferência de cores, correspondência espectral e aprendizado generativo (TOSTA et al., 2019; BENTAIEB; HAMARNEH, 2017). Há também métodos de normalização como o *stain – to – stain*, um método baseado em *pix2pix*, que desfaz a necessidade da imagem de referência e também atinge grande semelhança visual com o resultado esperado. Essa abordagem é baseada no aprendizado, e trata o problema da normalização como um problema de transferência de estilo, usando as GANs.

2.5 Inteligência Artificial

A definição do termo Inteligência Artificial (IA) é motivo de discussão entre os pesquisadores da área, mas em geral, pode-se definir como o estudo de grupos de algoritmos capazes de resolver problemas que envolvam aprendizado (ROSA, 2011). Em que

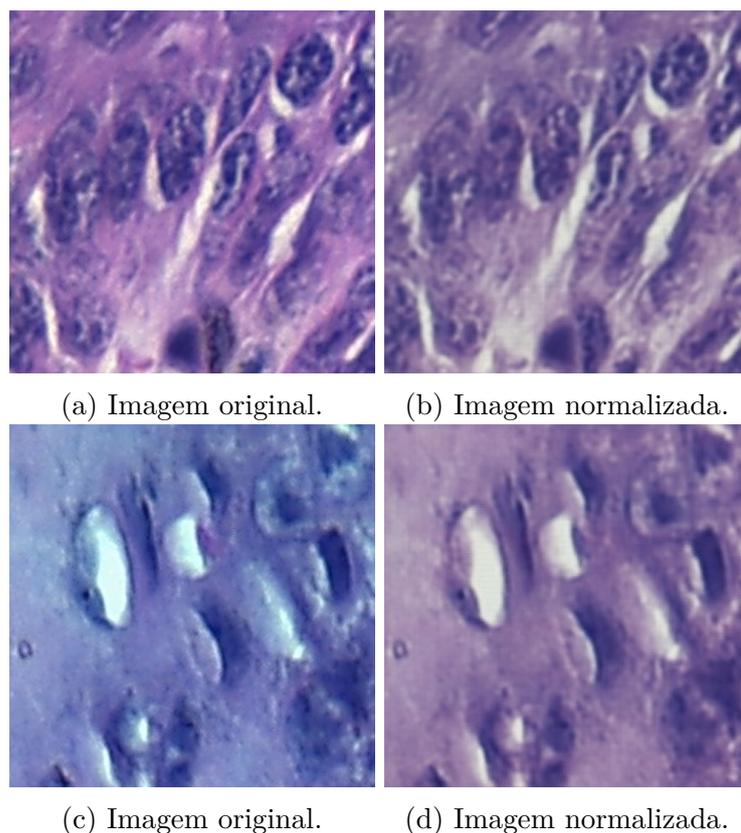


Figura 3 – Exemplos de normalização de corantes H&E: (a) e (c) imagens originais e (b) e (d) imagens normalizadas.

suas aplicações vão de tarefas corriqueiras, como um chatbot capaz de simular uma conversa, até tarefas especializadas como provas complexas de teoremas. As áreas de atuação dessa tecnologia podem englobar várias outras, como jogos, processamento de linguagem natural, geometria, cálculo, visão computacional e diagnósticos médicos. Na Figura 4 é possível observar as principais subáreas e as relações no caso do estudo de IA.

2.5.1 Aprendizado de Máquina

O aprendizado de máquina (do inglês, *machine learning* - ML) se refere especificamente ao processo de aprendizado realizado pelos algoritmos de IA. Existem inúmeros algoritmos, de modo que podem ser categorizados de acordo com os métodos e tecnologia dos quais fazem uso. Nesse caso, podemos dividir os algoritmos de ML no tipo de aprendizado: aprendizado supervisionado, não-supervisionado e semi-supervisionado (ROSEBROCK, 2017).

De forma sucinta, a diferença entre esses três tipos de aprendizado se baseia na relação do algoritmo com os dados de entrada e os dados resultantes de seu processo de execução. Os algoritmos de aprendizado supervisionado realizam um processo de treinamento, previsões são feitas sobre os dados de entrada e então são corrigidas quando essas previsões estão erradas. O aprendizado não-supervisionado se baseia na ideia de que o



Figura 4 – As principais subáreas de pesquisa em IA (ROSEBROCK, 2017).

algoritmo não é capaz de saber se acertou sua predição, isto é, nesse tipo de aprendizado o algoritmo é capaz de aprender com seus próprios resultados mesmo sem saber se estão corretos, e por isso é o tipo de aprendizado mais complicado de se obter bons resultados. Já o aprendizado semi-supervisionado é basicamente a combinação dos outros dois tipos de aprendizado, isto é, tem-se as respostas para apenas algumas predições, mas não todas, conforme afirma os autores em (ROSEBROCK, 2017).

Os algoritmos dessas categorias devem ser usados para criação de um modelo, o qual é usado numa fase de treinamento com um conjunto de dados e então pode ser aplicado em dados de teste para fazer a predição. Vários modelos têm sido usados em pesquisas na área de ML: árvore de decisão, máquina de vetores de suporte, regressão logística, *Naïve Bayes*, métodos do *ensemble*, etc. A seguir, são descritos os algoritmos selecionados e investigados nesse trabalho.

2.5.1.1 Florestas Aleatórias

O algoritmo florestas aleatórias (do inglês, *random forests* -RFs) são estruturas de dados baseadas em árvores de decisão, que por sua vez são baseadas na clássica estrutura de dados árvore, as quais geralmente são empregados as árvores binárias. Uma árvore lógica é um grafo em que quaisquer dois nós (ou vértices) estão conectados por exatamente um único caminho, havendo então nós raízes no topo da árvore, e nós folhas na base da árvore (ver Figura 5).

Em uma árvore de decisão tem-se que cada etapa de ramificação da árvore repre-

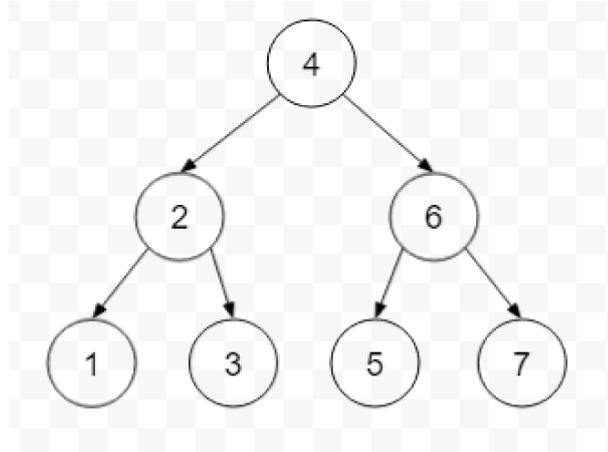


Figura 5 – Exemplo de uma árvore binária, onde o nó do topo é chamado de nó raiz (representado pelo elemento número 4), e os nós na base são chamados de nós folhas (representados pelos elementos números 1, 3, 5, 7).

senta uma decisão. Isto é, dado que o fluxo de execução do algoritmo se inicia na raiz da árvore e segue até sua base, podem existir diferentes caminhos até que a execução alcance as folhas. A escolha de qual caminho seguir é feita na etapa de decisão, quando um teste estatístico é realizado entre os dados de entrada e as regras de cada decisão (LOUPPE, 2015; RAMOS; NASCIMENTO; PEREIRA, 2012). Na Figura 6 é exemplificado a execução de uma árvore de decisão. Nesse modelo ocorre o particionamento do dado de entrada $X = [X_1: 0, X_2: 1]$, as soluções possíveis (ou classes) são c_1 e c_2 , e as regras de decisão são $X_1 < 0.7$ e $X_2 \leq 0.5$. Como $X_1 = 0$ é menor ou igual a 0,7 e $X_2 = 1$ é maior do que 0,5, percorre-se o caminho $t_0 > t_1 > t_4$ e a classe c_1 é alcançada.

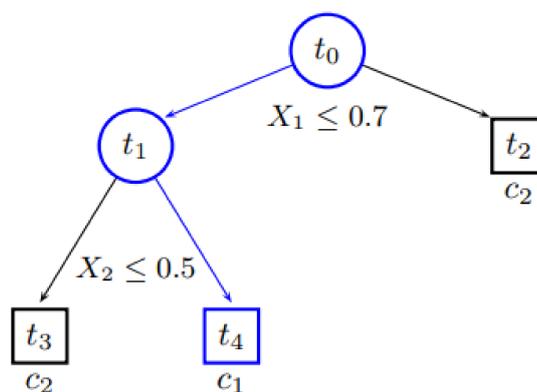


Figura 6 – Exemplo de uma árvore de decisão que possui cinco nós (LOUPPE, 2015).

A árvore de decisão é um algoritmo de aprendizado supervisionado. Isso significa que seu treinamento é feito a partir de dados de entrada que já possuem respostas. Dessa maneira, são geradas as regras de decisão e a árvore é construída a partir de funções indutivas baseadas em valores gerado aleatoriamente, garantindo baixa tendência (bias)

e alta variância.

Desse modo, as chamadas florestas aleatórias são estruturas que agrupam várias árvores de decisão. Como as árvores de decisão possuem alta variância em seus resultados, muitas vezes duas árvores diferentes resultam em decisões diferentes. Assim, ao se fazer uso de um conjunto de árvores, é possível calcular a decisão de todas as árvores da floresta e computar a porcentagem de respostas iguais, quase como uma votação, de modo que a decisão com maior porcentagem é aquela tomada como resposta pela floresta (LOUPPE, 2015). O algoritmo RF deve ser aplicado da seguinte forma:

- Seleciona-se o número de crescimento de árvores, através do parâmetro N , e um número m não maior que o número de descritores no classificador;
- Para $i = 1$ até N aplica-se uma função $Train_forest()$;
- Seleciona-se aleatoriamente, uma amostra *bootstrap* dos descritores. Os descritores que não são selecionados nessa amostra são nomeados como dados *out-of-bag*;
- Cria-se uma árvore aleatória, onde, em cada nó, o melhor conjunto é escolhido dentre m variáveis selecionadas aleatoriamente;
- Em seguida, a árvore é utilizada para prever os dados *out-of-bag*;
- No final, os resultados dos dados *out-of-bag*, são utilizados para indicar em qual classe deve classificar os descritores.

2.5.1.2 Máquina de Vetores de Suporte

O classificador máquina de vetores de suporte (do inglês, support vector machine - SVM) é um algoritmo de aprendizado supervisionado que possui o objetivo de fazer a separação dos dados de entrada utilizando o conceito estatístico da separabilidade linear e gradiente descendente (VAPNIK, 1999). Com os dados de entrada e suas características é possível distribuí-los de forma a separá-los com um vetor, também chamado de hiperplano, como ilustrado na Figura 7.

Para aplicar o SVM deve-se criar um conjunto de treinamento X com n descritores pertencentes a X com seus respectivos rótulos $y_i \in Y$, sendo que X constitui o espaço de entrada e $Y = \pm 1$ as possíveis classes para a classificação. Se X for linearmente separável pelas classes $+1$ e -1 , então, X pode ser classificado aplicando a equação 2.1:

$$H_x = w \cdot x + b \quad (2.1)$$

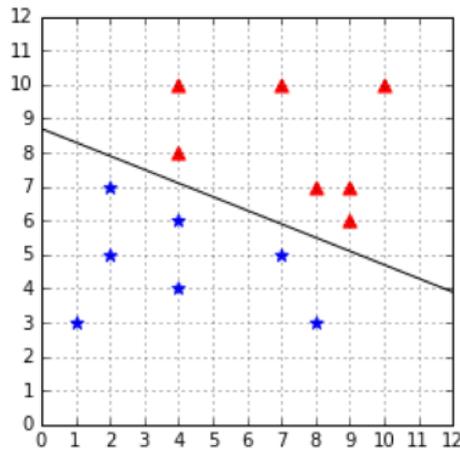
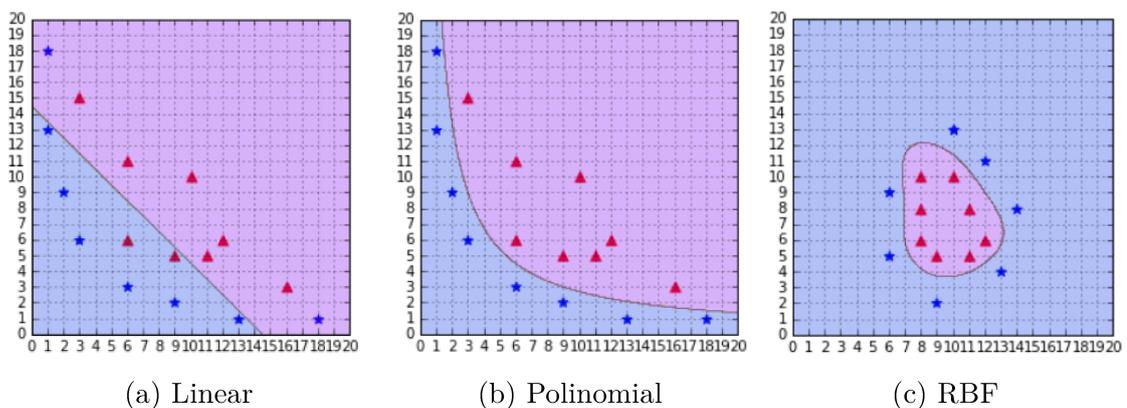


Figura 7 – Exemplo de distribuição de duas classes (vermelho e azul) em um hiperplano, e um vetor capaz de separá-los em dois grupos distintos (KOWALCZYK, 2017).

onde, $w \cdot x$ é o produto escalar entre os vetores w e x , e w pertence a X . A margem é um vetor normal ao hiperplano, descrito por $b/\|w\|$, sendo $b \in \mathbb{R}$, e w correspondente à distância do hiperplano em relação a origem.

No entanto, muitas vezes os dados não podem ser separados por um vetor linear, como na Figura 7, e para resolver esse problema podem ser utilizados outros núcleos (*kernels*). O *kernel* é a função capaz de realizar uma transformação indireta no vetor original. Desse modo o vetor original não é afetado, mas o vetor resultante do processo de execução sofre transformações. Existem alguns tipos de *kernels* e os mais utilizados são o linear, o polinomial, o RBF (*Radial Basis Function*) ou Gaussiano. O kernel linear é o mais simples deles, pois apenas faz a multiplicação de dois vetores. O kernel polinomial aplica um grau de potencialização polinomial na multiplicação dos vetores, resultando em possíveis curvas no vetor de separação. Já o kernel RBF é o mais complexo, pois faz uma projeção do vetor em um espaço dimensional infinito (KOWALCZYK, 2017). Na Figura 8 são exemplificados os diversos tipos de *kernel* para um SVM.



(a) Linear

(b) Polinomial

(c) RBF

Figura 8 – Exemplo de classificações com kernels linear, polinomial e RBF, respectivamente (KOWALCZYK, 2017).

2.5.2 Redes Neurais Convolucionais

Para compreender Redes Neurais Convolucionais (do inglês, *convolutional neural networks* - CNNs) é necessário conhecer primeiramente o conceito de redes neurais artificiais (do inglês, *artificial neural networks* - ANNs). De forma semelhante à rede de neurônios presente no cérebro de qualquer ser humano, as ANNs são capazes de receber informações na camada de entrada, processar esses dados e propagá-los adiante para as outras camadas de neurônios, até a última camada ou camada de saída. Dessa maneira o processamento realizado durante a comunicação entre as camadas é capaz de resolver diversos problemas do escopo da inteligência artificial. Na Figura 9 é exemplificado uma ANN simples. Esse modelo é um tipo de técnica de aprendizado supervisionado.

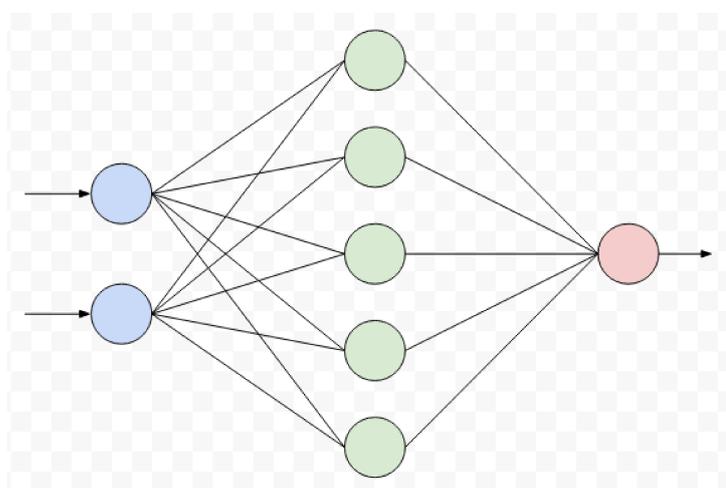


Figura 9 – Um exemplo simples de rede neural com os neurônios representados nas cores azul, verde e vermelho.

As CNNs são ANNs criadas especificamente para realizar o processamento de imagens e visão computacional. A principal modificação das CNNs quando comparadas às ANNs são as camadas convolucionais, as quais representam a principal característica desse tipo de rede. O núcleo das CNN é geralmente construído a partir das camadas: convolução, pooling, camadas totalmente conectadas (do inglês, *fully connected* - FC). Enquanto as duas primeiras são responsáveis pela extração de características, as camadas FC realizam a classificação dos atributos e geram como saída um rótulo para ser atribuído aos dados de entrada.

A camada convolucional é um procedimento linear que efetua a multiplicação dos elementos dos dados de entrada e matrizes numéricas denominadas *kernels*. Essas matrizes deslizam pela imagem, onde cada posição da matriz representa um pixel da imagem de entrada, aplicando um filtro através da convolução (ver Figura 10). Diferentes valores do *kernel* representam diferentes filtros, desse modo a rede é capaz de aprender a utilizar novos valores para essas matrizes (ROSEBROCK, 2017).

131	162	232	84	91	207
104	-1	109	+1	237	109
243	-2	202	+2	135	26
185	-15	200	+1	61	225
157	124	25	14	102	108
5	155	116	218	232	249

Figura 10 – Funcionamento de um *kernel* sobre uma região da imagem denominada *patch* (ROSEBROCK, 2017).

A utilização diferentes kernels resultam em imagens com diferentes transformações. Na Figura 11 é exemplificado o resultado sobre diferentes transformações em imagens. A partir dessas diferentes versões da imagem original, a rede é capaz de extrair características e processá-las a fim de realizar uma classificação da imagem, por exemplo.

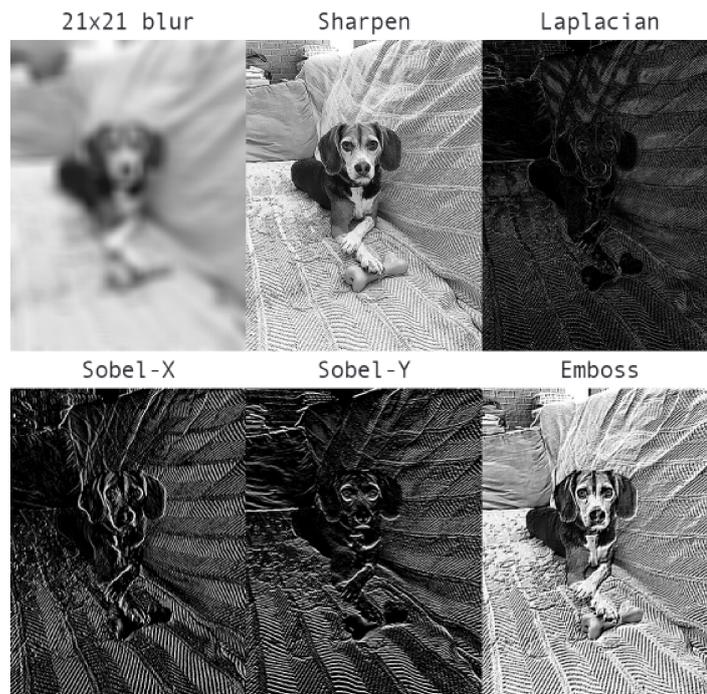


Figura 11 – Exemplos de transformações por diferentes tipos de *kernels* (ROSEBROCK, 2017).

A segunda operação consiste em aplicar uma função de ativação não-linear no mapa resultante da convolução. Entre as diversas abordagens, a unidade linear corrigida (do inglês, *rectified linear unit* - ReLU) é a mais empregada nessa etapa.

As camadas de *pooling* são responsáveis por diminuir o tamanho da matriz, ou seja, reduzir o tamanho da imagem processada. Para isso é utilizado um algoritmo semelhante à camada convolucional: uma janela pequena de tamanho variável, chamada de *pool*, desliza sobre a matriz de entrada, mas não se aplica uma convolução. Ao invés disso, aplica-se uma função de redução.

As principais funções de redução são a *max pooling* e *average pooling* (Figura 12). O *max pooling* seleciona o maior valor da *pool*, e o *average pooling* calcula a média entre os valores da *pool*.

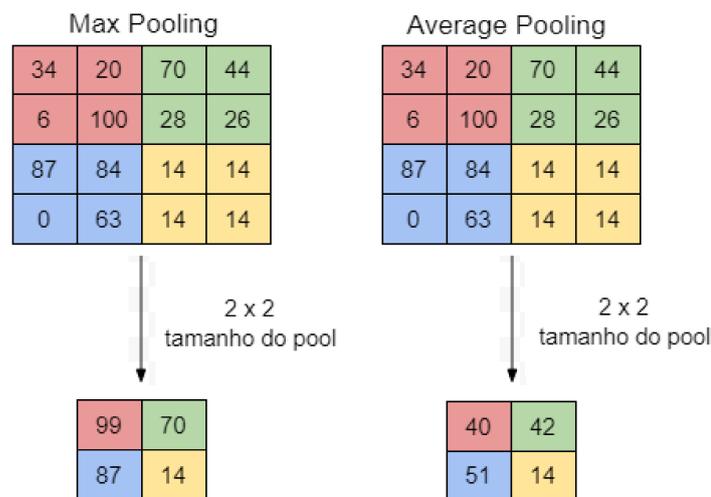


Figura 12 – Exemplos de *max pooling* e *average pooling*.

No geral, as primeiras camadas das CNNs são para extrair as características da imagem original, e então as últimas camadas são neurônios também utilizados em ANNs. As camadas FC consistem em uma série de camadas nas quais todas as saídas são conectadas as entradas da camada seguinte (semelhante à Figura 9) por meio de um peso adaptado por meio de treinamento. O ajuste dos pesos geralmente é feito pelo uso do algoritmo retropropagação. Em classificação, as saídas correspondem às probabilidades de classe, que são obtidas por meio de uma função de ativação. Normalmente, a função mais adequada para esta tarefa é a softmax.

Assim como há diversas abordagens para as ANNs, existem também diversos modelos de CNNs. Diversas formas de combinar tais camadas foram publicadas na literatura, disponibilizando os mais variados modelos de CNN para pesquisadores da área. Exemplos de redes CNNs são os modelos AlexNet, VGG-16 e ResNet50.

2.5.3 Redes Adversárias Generativas

Redes Adversariais Generativas (do inglês, *Generative Adversarial Networks* - GANs) são modelos generativos não-supervisionados que envolvem duas redes neurais artificiais, um gerador e um discriminador, que são treinados simultaneamente (SALEHI;

CHALECHALE, 2020). O primeiro modelo de GAN foi criado em 2014 por (GOODFELLOW et al., 2014). Até alguns anos depois, em 2017, as GANs ainda eram consideráveis instáveis e de difícil treinamento, não sendo utilizáveis para a maioria dos problemas. Após esse período, várias arquiteturas, funções de perda, técnicas condicionais e métodos de restrição foram explorados e desenvolvidos, aumentando a usabilidade desses modelos (BISSOTO, 2019).

Nessa abordagem, os modelos são denominados gerador e discriminador, os quais são redes neurais com arquitetura própria e funções de aprendizado independentes de seu adversário. Enquanto o discriminador tenta classificar as imagens em imagens reais ou imagens criadas sinteticamente, o gerador cria imagens sintéticas com o objetivo de enganar o discriminador e ter suas imagens classificadas como imagens reais. Na Figura 13 é mostrado um modelo simples de uma arquitetura de uma GAN.

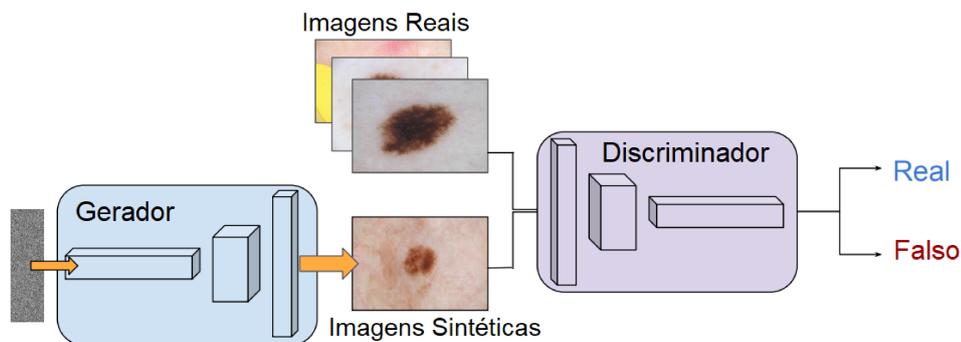


Figura 13 – Arquitetura GAN simplificada (BISSOTO, 2019).

No início do processo, o gerador produz imagens ruidosas e o discriminador prevê aleatoriamente. Uma das grandes características das GANs está em como alternar o treinamento das duas redes, de modo que o gerador se torna mais preciso para “enganar” a arquitetura discriminadora. O objetivo é adaptar para manter a capacidade de identificar corretamente quais observações são falsas. Esse processo faz o gerador definir novas maneiras de enganar o discriminador e assim o ciclo continua. Esse modelo usa camadas convolucionais para prever as informações no discriminador. Também é empregado camadas de normalização de lotes (batch) no discriminador para GANs. A entrada para o gerador são dados extraídos de uma distribuição padrão. A saída é uma imagem do mesmo tamanho que uma imagem nos dados originais do treinamento. O gerador de um GAN tem um decodificador que converte os dados em uma imagem.

Neste trabalho será utilizado um modelo de GAN que se baseia no modelo Pix2Pix. O Pix2Pix requer pares de imagens que consistem da imagem original e a imagem transformada, que representa a transformação a ser alcançada pelo treinamento. A arquitetura GAN utiliza uma rede U-Net como geradora, e uma PatchGAN como rede discriminadora. (SALEHI; CHALECHALE, 2020). A U-Net é uma rede usada geralmente para

realizar a segmentação de imagens biomédicas, e a PatchGAN é uma rede que classifica se cada patch de tamanho $N \times N$ da imagem é real ou falso. Os modelos U-Net e a PatchGAN são redes adversárias e durante a fase de treinamento elas estão sempre competindo entre si para gradualmente alcançar os objetivos. O gerador nessa abordagem é responsável por aprender a mapear um vetor de ruído e produzir imagens e o discriminador tem a tarefa de classificar se as imagens são verdadeiras, ou se são imagens falsas (SALEHI; CHALECHALE, 2020). Outros detalhes da implementação dessa abordagem são tratados na seção da Metodologia.

2.6 Estado da Arte

Trabalhos correlatos ao tema de normalização de imagens histológicas e seus respectivos métodos podem ser categorizados em alguns grupos de acordo com sua metodologia, entre eles estão os métodos de separação de mancha e os métodos de correspondência de cores do modelo.

Como as diferentes manchas das células possuem várias características das imagens, ser capaz de separar as informações sobre as manchas é de fundamental importância, assim como extrair cores da imagem original para as imagens normalizadas, de forma automática. Para isso, o método Macenko (MACENKO M. NIETHAMMER; THOMAS, 2009) propõe uma solução que assume que as manchas de hematoxilina e eosina são linearmente separáveis no espaço de cor de densidade óptica. Então, são encontradas as duas maiores direções de valor singular usando decomposição de valor singular e são projetados os valores de pixel de densidade óptica neste plano. No entanto, este tipo de método nem sempre pode estimar os vetores de coloração corretos se uma forte variação de coloração estiver presente nas lâminas histopatológicas.

Algoritmos de correspondência de cores do modelo fazem uso do espectro de cores RGB e tentam alterar os níveis dos canais de cores para combinar com os níveis da imagem de referência. Para isso, o método Reinhard (REINHARD M. ASHIKHMIN; SHIRLEY, 2001) propõe combinar as estatísticas de histogramas de cores de uma imagem de referência e as imagens de origem após transformar as imagens do modelo RGB em imagens do modelo LAB. No entanto, quando várias cores estão presentes da imagem, não existe a suposição da distribuição unimodal de pixels em cada canal do espaço de cores LAB. Portanto, isso pode resultar em áreas de fundo sendo mapeadas como regiões coloridas (SALEHI; CHALECHALE, 2020).

O método proposto por Vahadane (VAHADANE T. PENG; NAVAB, 2016) desenvolve uma abordagem de normalização de mancha com base na técnica de fatoração de matriz não negativa esparsa para preservar as informações estruturais da imagem de origem. A complexidade computacional desse método é consideravelmente maior se com-

parado ao método anterior, e também não preserva todas as informações de cores da imagem de origem. Embora essas soluções levem a uma melhor estimativa de coloração, elas são limitadas às informações de cor da imagem empregada como referência (SALEHI; CHALECHALE, 2020).

Esses métodos produzem bons resultados de normalização, porém todos eles exigem uma imagem de referência que seja classificada como uma boa representação da doença estudada por um especialista da área. Como existe a dependência de um especialista, e diferentes especialistas podem ter diferentes opiniões sobre as classificações de boas imagens representativas, faz-se necessário o desenvolvimento de um método que não necessite de uma imagem de referência. Para tanto, o método *Stain-to-Stain* (STST) (SALEHI; CHALECHALE, 2020) foi desenvolvido e faz uso de GANs baseadas no framework Pix2Pix para realizar a normalização das imagens histológicas. Essas soluções podem ser promissoras para normalização de corantes H&E para alguns tipos de imagens histológicas. No contexto de imagens da cavidade oral esse tipo de técnica ainda não foi investigado, assim como sua influência em relação às características de classificação dos diversos níveis de uma lesão em imagens desse padrão.

3 Metodologia

Neste capítulo são apresentadas as informações a respeito do banco de imagens, as etapas do desenvolvimento que representam aspectos relevantes da estrutura das soluções propostas, além das tecnologias utilizadas para sua construção dos algoritmos.

3.1 Banco de Imagens

O banco de imagens utilizado nesse trabalho foi construído pelo pesquisador Adriano B. Silva ([SILVA, 2019](#)) no estudo proposto em seu trabalho de Mestrado. As imagens são amostras de tecidos de língua de camundongos contendo lesões induzidas experimentalmente utilizando o carcinógeno 4NQO e aprovadas pelo Comitê de Ética na Utilização de Animais.

As imagens histológicas digitais foram obtidas com o uso do microscópio óptico Leica DM500, em magnificação de 400x, e armazenadas em arquivos em formato TIFF, utilizando o modelo de cores RGB e resolução original de 2048×1536 pixels. Essas imagens foram recortadas em regiões de interesse com resolução de 450×250 pixels para áreas de presença da displasia. O banco de imagens possui 296 imagens, divididas igualmente nas classes severa, moderada, leve e saudável. Na Figura 14 é mostrado exemplos das diversas classes das imagens histológicas da cavidade oral.

3.2 Classificação de Lesões em Imagens Histológicas Normalizadas

O sistema proposto é composto de três fases: pré-processamento, normalização de corantes H&E e classificação. Para a etapa de normalização foi implementado um método de normalização de imagens baseado em um modelo de rede GAN, e para a classificação foram desenvolvidos sistemas híbridos que combinam CNN e algoritmos de aprendizado de máquina para a classificação das lesões nas imagens histológicas. Nessa etapa, duas arquiteturas de redes neurais artificiais foram estudadas para avaliação da normalização na etapa de classificação: os modelos AlexNet e ResNet50 e os algoritmos de aprendizado de máquina RF e SVM. Na Figura 15 é exemplificado as etapas do sistema proposto.

O sistema foi desenvolvido utilizando a linguagem Python e as bibliotecas do Keras e scikit-learn. Os experimentos foram conduzidos em um computador com processador Intel i3-6006U, 4GB de memória RAM e sistema operacional Windows 10.

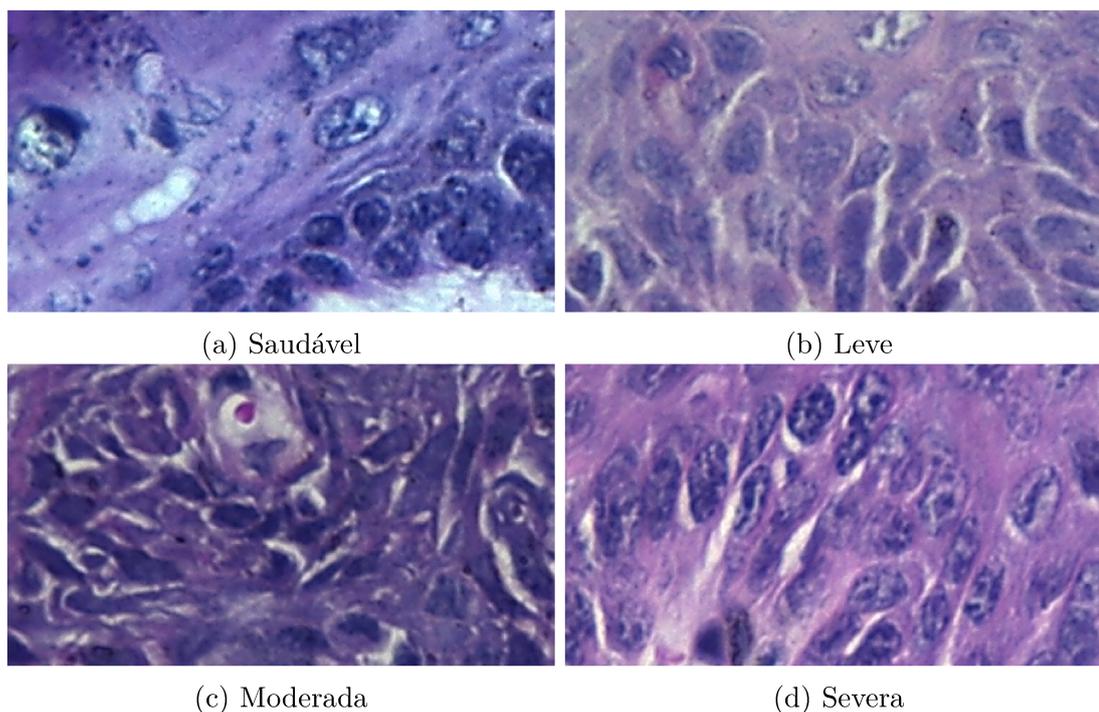


Figura 14 – Exemplos de tecidos histológicos da cavidade oral: (a) saudável, (b) displasia leve, (c) displasia moderada e (d) displasia severa.

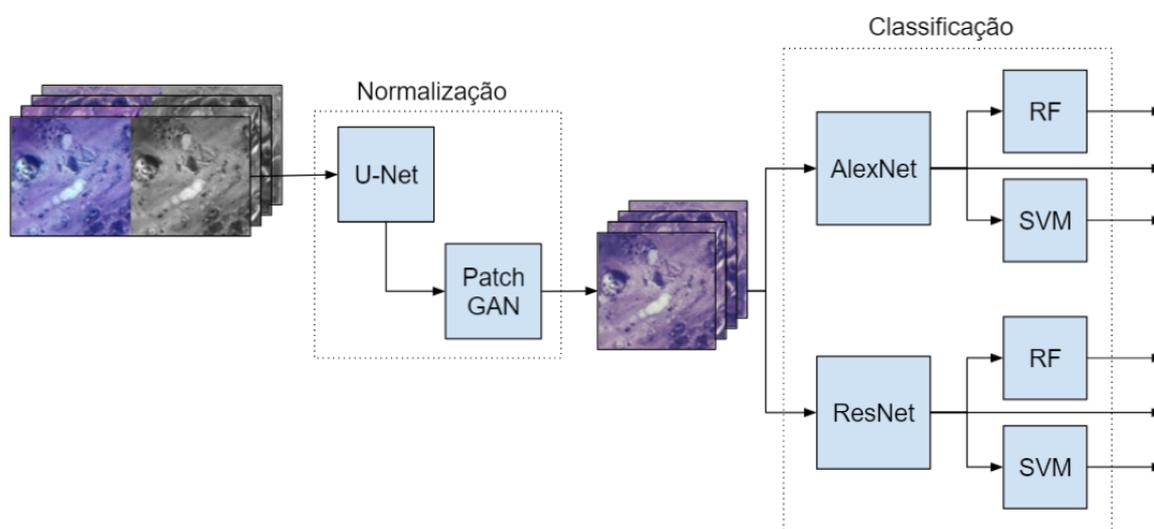


Figura 15 – Modelo desenvolvido neste trabalho para avaliação das imagens histológicas.

3.2.1 Pré-Processamento das Imagens

Na etapa de pré-processamento foi realizado o recorte, ou *cropping*, das imagens das regiões de interesse. Os *crops*, ou *patches*, gerados a partir desse processo devem possuir as dimensões 256x256 pixels para que sejam utilizáveis pela implementação utilizada do modelo de normalização *Stain-to-Stain*. Para isso, as imagens de 460x256 foram divididas em dois *patches* (ver na Figura 16), produzindo 148 imagens de 256x256 pixels para

cada classe, totalizando 592 imagens.

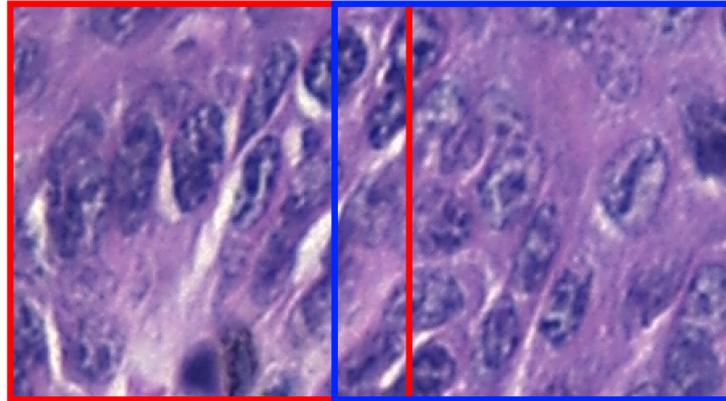


Figura 16 – Exemplo do *cropping* realizado em que dois *patches* (representados nas cores vermelho e azul) são produzidos para cada imagem.

3.2.2 Normalização com Rede Adversária Generativa

Para realizar a normalização das imagens foi utilizado o modelo denominado *Stain-to-Stain* (STST) *normalization* proposto por (SALEHI; CHALECHALE, 2020). Na Figura 17 é ilustrado a arquitetura do modelo STST em que são representados os dados como entrada, as imagens reais e suas respectivas versões em níveis de cinza. Essa abor-

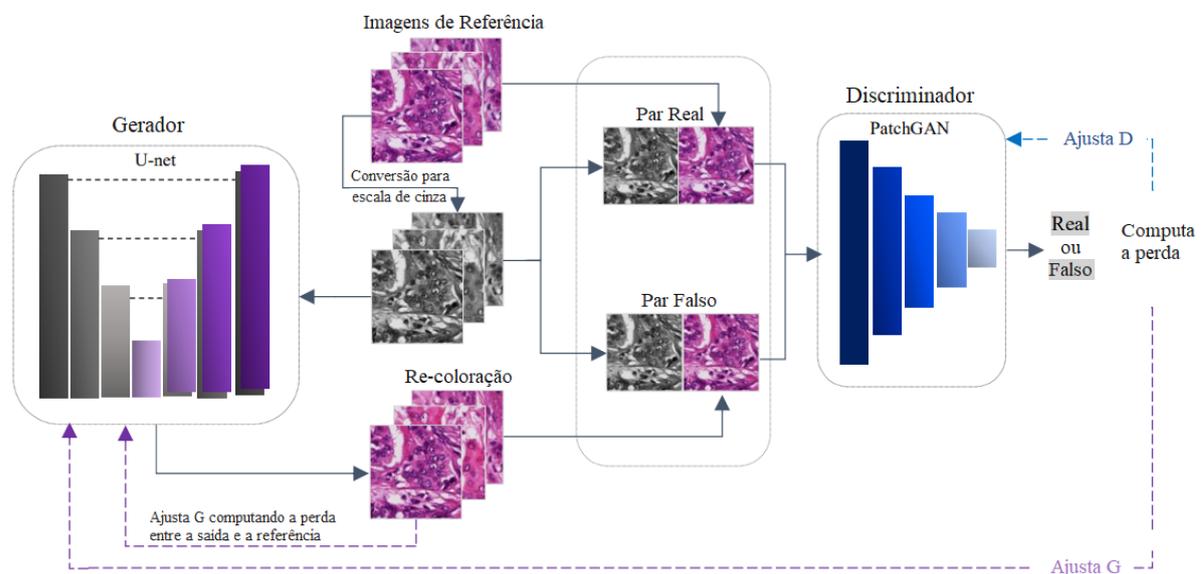


Figura 17 – Modelo da GAN para a normalização Stain-to-Stain (SALEHI; CHALECHALE, 2020)

dagem emprega um processo de representação imagem-a-imagem para a normalização de corantes em imagens histológicas. Para essa tarefa nos modelos tradicionais de normalização, uma imagem de referência deve ser usada para a normalização dos corantes.

No modelo estudado nesse trabalho esse processo ocorre por meio de GANs condicionais (cGANs), definida pelos autores em (MIRZA; OSINDER, 2014).

A cGAN aprende a mapear por meio de observação em uma imagem x e um vetor de ruídos randômicos z para uma saída y , $G:z \rightarrow y$. O gerador G é treinado para produzir saídas que não podem distinguir de imagens reais por um discriminador D , o qual é treinado para fazer a detecção das gerações falsas. Essa abordagem emprega uma rede U – *net* para a geração das imagens e uma arquitetura *PatchGAN* para discriminação das informações. No modelo gerador, as camadas de codificadores e decodificadores são diretamente conectados por *Skip Connection (SC)* ("pular conexão"). Uma vez que a conexão por SC pode transportar as informações de baixo nível ao longo das camadas da rede codificadora e decodificadora. Isso pode contribuir no desempenho da representação dos corantes. No modelo *PatchGAN*, cada imagem é dividida em segmentos $n \times n$, para fazer a predição de cada parte em verdadeiro ou falso. Finalmente, a média de todas as respostas baseadas nos segmentos é calculada e a classificação é realizada.

O objetivo de uma cGAN é dada pela Equação 3.1:

$$L_{cGAN}(G.D) = E_{x \sim p_{data}(x)}[\log(D(x.y))] + E_{z \sim p_z(z)}[\log(1 - D(G(z.y).y))], \quad (3.1)$$

onde G tenta minimizar esse objetivo em comparação com uma rede adversária D , que tenta maximizá-lo. O estudo aqui considerou o modelo proposto por Isola et al. (ISOLA et al., 2017), o qual emprega um abordagem *Pix2Pix*, em que elimina a necessidade de uma imagem de referência. Essa abordagem permite realizar o mapeamento da imagem de entrada para uma saída junto com uma função de perda do treinamento para o mapeamento.

No modelo *Pix2Pix*, a perda $L1$ (Equação 3.2) faz com que o gerador produza uma amostra que se assemelha a variável x (imagem de entrada).

$$L_{L1}(G) = E_{x,y,z}[\|x - G(z.y)\|_1]. \quad (3.2)$$

Então, o termo de normalização $L1$ é adicionado a Equação 3.2, o qual é usado para analisar a função de perda adversária. A função de perda é dada por:

$$L(G.D) = L_{cGAN}(G.D) + \lambda L_{L1}(G), \quad (3.3)$$

onde λ denota um hiper-parâmetro que controla os pesos. Nessa etapa, o valor 100 foi empregado conforme sugerido pelos autores (SALEHI; CHALECHALE, 2020). Durante o treinamento, o objetivo é encontrar o gerador G baseado em resolução de problemas por meio de otimização:

$$G^* = \operatorname{argmin}(G) \operatorname{argmax}(D) L(G.D) \quad (3.4)$$

Nesse modelo, pares de imagens devem ser usadas no treinamento, as quais representam a imagem original e a correspondente imagens transformadas. Como não há

esse conjunto disponível, as imagens em escala de cinza e a correspondente RGB são re-passadas como entrada. Essa abordagem segue as recomendações sugeridas por (SALEHI; CHALECHALE, 2020). Após várias iterações e gerações das redes, a acurácia é melhorada e a imagem resultante proporciona um padrão normalizado de cores para os corantes H&E. Nessa etapa foi realizado o treinamento a partir de 20 imagens selecionadas por um especialista em histologia que as classificou com uma boa distribuição de cores. Dada essas 20 imagens foram gerados 481 patches de 256 x 256 pixels para serem utilizados no treinamento e os 592 patches de 256 x 256 pixels para o teste, e conseqüentemente produzindo as imagens normalizadas. Foram usadas 30 épocas para treinamento desse modelo.

3.2.3 Redes Neurais Convolucionais

Após essa etapa, as imagens histológicas devem ser classificadas para uma avaliação em relação ao processo de normalização. Inicialmente, um novo recorte foi realizado nas imagens, resultando em 11.840 imagens de 32 x 32 pixels, sem sobreposição de patches, e sem seleção de características, de modo que existiram imagens sem núcleos celulares. Esse recorte ocorreu devido a capacidade de processamento dos dados e necessidade de memória para processamento. Após isso, variações das imagens foram criadas em uma etapa denominada aumento de dados (*do inglês, data augmentation - DA*), produzindo cópias baseadas em imagens com modificações em rotação e espelhamento. Esses dados foram produzidos para melhorar a capacidade da rede. As imagens foram geradas em diferentes direções, totalizaram-se 94.720 imagens de tamanho de dimensão 32 x 32 pixels divididos para as quatro classes de imagens histológicas.

Nessa etapa, a divisão das imagens foi realizada da seguinte maneira: 20% foi retirada para ser utilizada na fase de teste. Dos 80% restantes, 20% foram retirados para validação, e o restante foi utilizado para treinamento. Dessa maneira, temos que 64% do total das imagens foi utilizado para treinamento, 20% foi utilizada para teste e 16% foi utilizado para validação. Isto é, do total de 94.720 imagens, foram utilizadas 60.620 imagens para treinamento, 15.156 imagens para validação e 18.944 imagens para teste. As redes foram treinadas por 50 épocas.

3.2.3.1 AlexNet

A rede AlexNet é o primeiro modelo de CNN estudado nesse trabalho. Sua arquitetura é definida na Figura 18, foi proposta em 2012 por Alex Krizhevsky, Geoffrey Hinton e Ilya Sutskever. Basicamente, a rede AlexNet tem uma arquitetura um pouco mais avançada do que modelos clássicos, em que emprega cinco camadas convolucionais com função de ativação ReLU, camadas de *pooling* após a primeira, segunda e quinta

camadas de convolução, e por fim duas camadas densas.

A rede AlexNet foi criada com o objetivo de vencer o desafio “*ImageNet Large Scale Visual Recognition Challenge*” e alcançou a acurácia em reconhecimento de imagens maiores do que modelos anteriores a ela, resultando em um momento histórico no aumento de interesse pela tecnologia de redes neurais pela comunidade científica (ALOM TAREK M. TAHA; ASARI, 2018; YU LAURENCE T. YANG, 2021). Apesar de possuir uma arquitetura relativamente já difundida, a AlexNet ainda gera bons resultados para problemas de classificação de imagem, e por isso ainda é bastante estudada e utilizada.

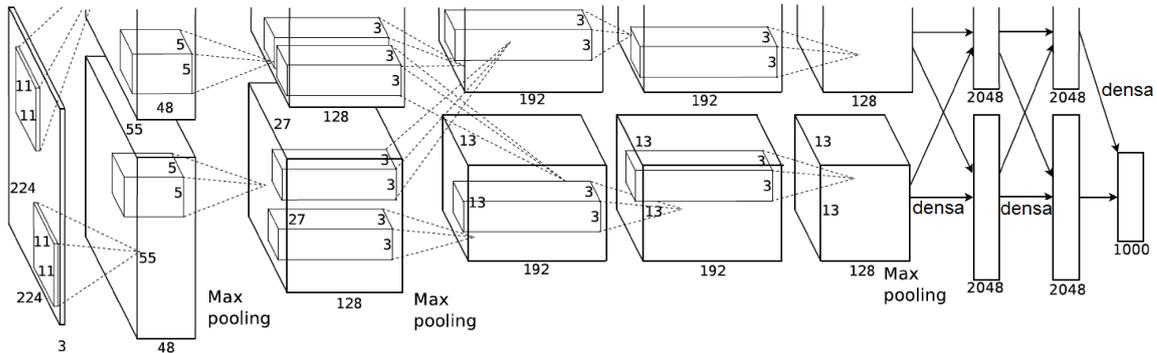


Figura 18 – Modelo original da arquitetura de uma rede AlexNet (YU LAURENCE T. YANG, 2021).

Na implementação da rede AlexNet foram feitas algumas modificações na arquitetura com objetivo de melhorar os resultados. Essas modificações foram a inserção de duas camadas de *dropout*, sendo uma no início e uma no final da arquitetura, e a utilização de regularizadores de *kernel* e *bias*. Essas adaptações são ofertadas pela biblioteca do Keras. A camada de *dropout* é responsável por desativar aleatoriamente alguns neurônios da camada anterior, e as regularizações utilizadas são capazes de aplicar penalidades nas funções de ativação do modelo, aumentando a otimização. O principal motivo de realizar essas alterações foi evitar *overfitting* do modelo original. Isso ocorreu porque o modelo apresentava um bom desempenho na fase de treinamento, mas que não representava bons resultados na fase de teste, indicando que o modelo “decorou” as classificações das imagens.

3.2.3.2 ResNet

A principal característica da ResNet (*Residual Network*) é as conexões de salto ou atalho, também chamadas de blocos residuais, que são capazes de pular uma ou mais camadas para obter o valor de entrada adicionado à camada de saída (ver Figura 19). Dessa maneira, são mantidas intactas as informações de entrada tanto quanto possível, produzindo um erro de treinamento menor do que os modelos que não utilizam blocos

residuais (YU LAURENCE T. YANG, 2021).

As camadas da rede ResNet podem variar, e arquiteturas com 50, 101 ou 152 camadas estão entre as mais utilizadas. Neste trabalho foi utilizado o modelo ResNet50, com 50 camadas, um dos modelos mais simples da arquitetura capaz de produzir ótimos resultados.

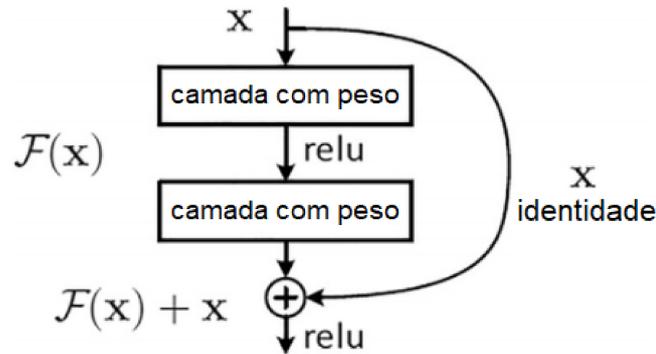


Figura 19 – Bloco residual de uma ResNet. $F(x) + x$ é a função residual em que $F(x)$ representa as camadas não-lineares empilhadas e x representa a identidade necessária pra tornar a entrada equivalente à saída (YU LAURENCE T. YANG, 2021).

3.2.4 Algoritmos de Aprendizado de Máquinas

A partir do treinamento dos modelos CNNs, foi realizada a extração de características da última camada antes da classificação com o *softmax*. Essas informações foram extraídas do “*flattening*” e foram empregadas para construir os vetores de características. Para os algoritmos de aprendizado de máquina foram extraídos 1024 características por imagem. Os vetores de características foram avaliados com os algoritmos RF e SVM.

Os algoritmos RF e SVM foram utilizados a partir da biblioteca scikit-learn, do Python. Para o RF foi utilizada uma floresta com 20 árvores e os parâmetros padrões da biblioteca, como por exemplo, máximo de *features* igual a raiz quadrada do número de características, número mínimo de amostras em um nó folha igual a 1, e número mínimo de amostras necessárias para dividir um nó interno igual a 2. Para o SVM foi utilizado o *kernel* polinomial de grau 3 e os parâmetros padrões da biblioteca, como gama escalar, tolerância para critério de parada igual a 1e-3 e tamanho do cache para o processamento igual a 200 MB. Ambos os algoritmos foram treinados sob as características de treino e avaliadas sob as características de teste.

3.3 Avaliação de Desempenho das Abordagens

Algumas métricas e testes estatísticos foram implementados para validar os resultados gerados a partir das execuções dos algoritmos. Para avaliar o resultados da CNN, RF e SVM foram utilizadas as métricas sensibilidade, especificidade, precisão e acurácia, retiradas a partir da matriz de confusão. A matriz de confusão combina as classes reais e as classificações realizadas, como exemplificado na Figura 20. O número de verdadeiros positivos representa a quantidade de classificações corretas da classe, a quantidade de verdadeiros negativos representa quantas vezes o classificador acertou em não classificar o dado como essa classe, os falsos positivos representam os erros na classificação da classe, e os falsos negativos representam os erros na classificação contrária, quando se erra em não classificar o dado como esta classe.

		Classificação	
		Sim	Não
Real	Sim	Verdadeiro Positivo (VP)	Falso Negativo (FN)
	Não	Falso Positivo (FP)	Verdadeiro Negativo (VN)

Figura 20 – Exemplo de uma matriz de confusão.

A sensibilidade (SE) também pode ser chamada de taxa de verdadeiros positivos, pois é calculada sob a soma dos verdadeiros positivos e falsos negativos:

$$\text{Sensibilidade} = \frac{vp}{vp + fn}$$

A especificidade (ES) é a taxa de verdadeiros negativos, pois é calculada sob a soma dos verdadeiros negativos e falsos positivos:

$$\text{Especificidade} = \frac{vn}{vn + fp}$$

A precisão (PR) é a taxa de verdadeiros positivos dado todas as classificações positivas:

$$\text{Precisão} = \frac{vp}{vp + fp}$$

A acurácia (AC), que foi a métrica mais considerada nesse trabalho, é a taxa de acertos dados todas as classificações:

$$\text{Acurácia} = \frac{vp + vn}{vp + vn + fp + fn}$$

4 Resultados

Neste capítulo são apresentados os resultados obtidos para etapa de normalização dos corantes H&E para as imagens histológicas de displasia. São tratados os valores das métricas obtidas para etapa de classificação das diversas classes das lesões.

4.1 Normalização dos Corantes H&E com o Método STST

Na Figura 21 são mostrados os resultados da normalização com o método que emprega a GAN denominado STST para as imagens histológicas. Os resultados mostram que o método é capaz de fazer uma distribuição dos corantes para normalizar as imagens que estão mais degradadas em relação a distribuição dos corantes (ver Figuras 21(a), 21(b), 21(g) e 21 (h)). Nota-se nessas Figuras que a intensidade de brilho foi melhorada para uma uniformização da distribuição dos corantes. No entanto, para a imagem da Figura 21(f), essa modificação não é perceptível visualmente em relação a imagem original.

4.2 Investigação das Arquiteturas de Redes Neurais Convolucionais

Nas Figuras 22 e 23 são apresentadas as informações a respeito do desempenho das redes AlexNet e ResNet50 na fase de treinamento e validação, respectivamente. O eixo X representa as informações sobre as épocas empregadas no treinamento, das quais todas foram treinadas por 50 épocas, e o eixo Y representa a porcentagem de acurácia e perda das redes. O treinamento da AlexNet gerou maiores variações no processo de treinamento com uma acurácia próxima de 90%. Na ResNet50 ocorreu um pequeno distanciamento entre as curvas de treinamento e validação, porém a rede atingiu uma acurácia maior que 95% para a etapa de treinamento.

Neste experimento foram considerados avaliação das imagens: as imagens originais e as imagens normalizadas. Para essa investigação foram propostos quatro cenários: i) treinamento com imagens originais e teste com imagens originais; ii) treinamento com imagens originais e teste com imagens normalizadas; iii) treinamento com imagens normalizadas e teste com imagens originais e iv) treinamento com imagens normalizadas e teste com imagens normalizadas. Os resultados da classificação das redes sem uso da normalização é apresentado na Tabela 1. Nesta investigação observou-se que a ResNet proporcionou os melhores resultados em relação a métrica de acurácia para grande parte dos grupos investigados. Apenas para o grupo de lesões moderadas que a AlexNet foi

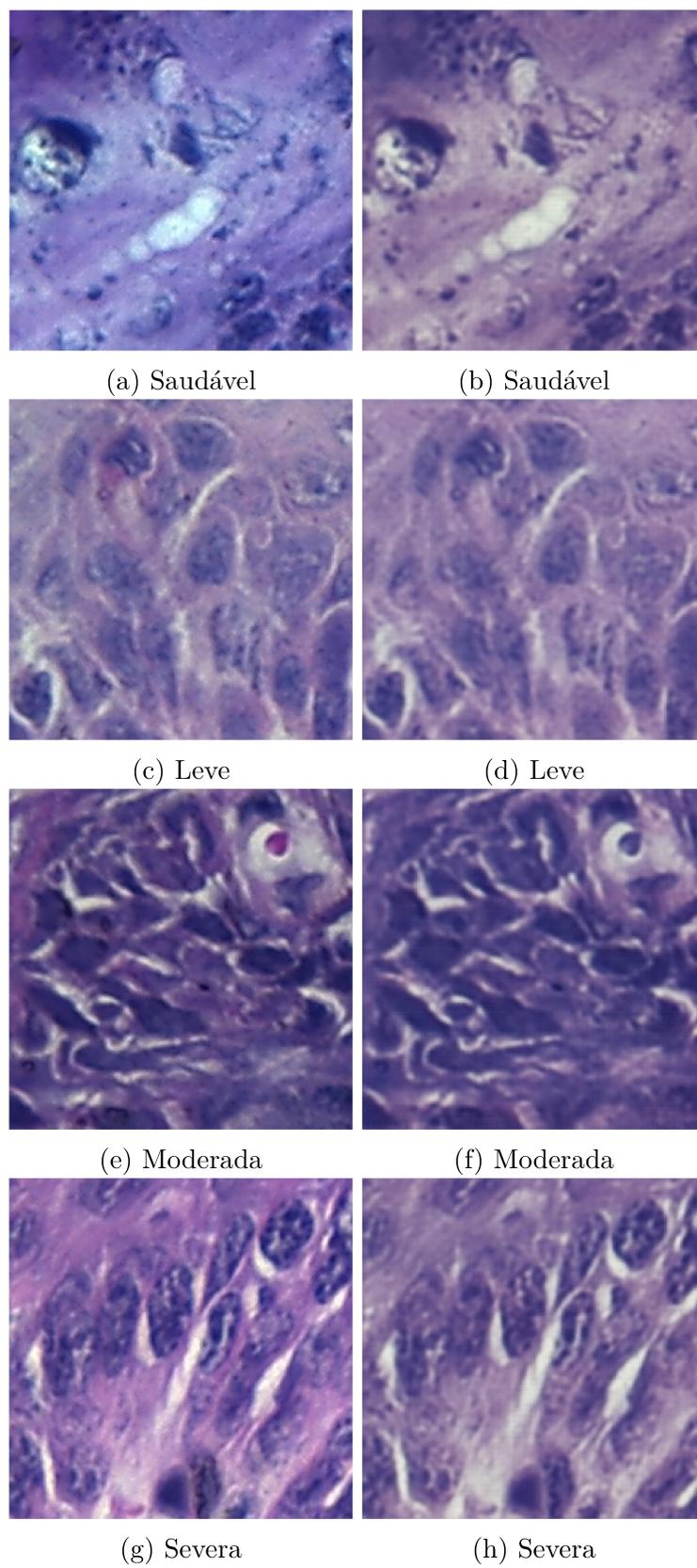


Figura 21 – Exemplo de normalizações de cada classe usando o método STST.

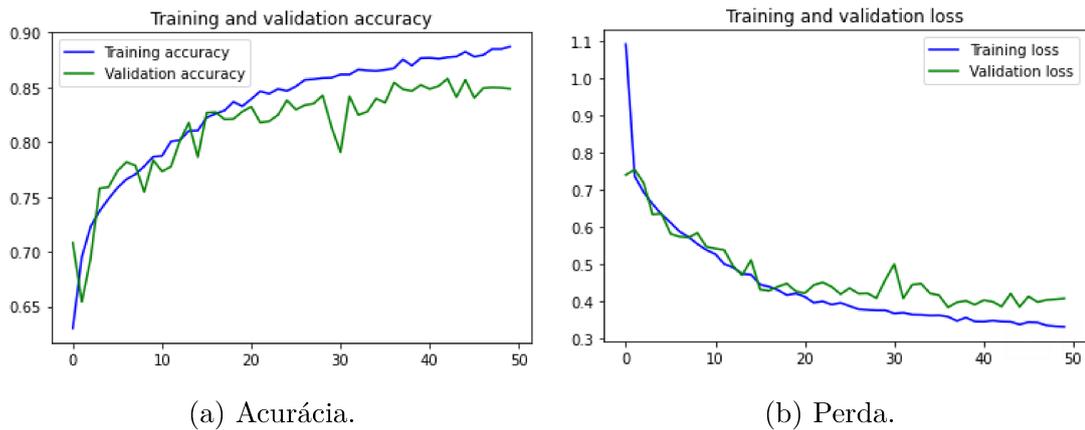


Figura 22 – Resultados do treinamento do modelo AlexNet.

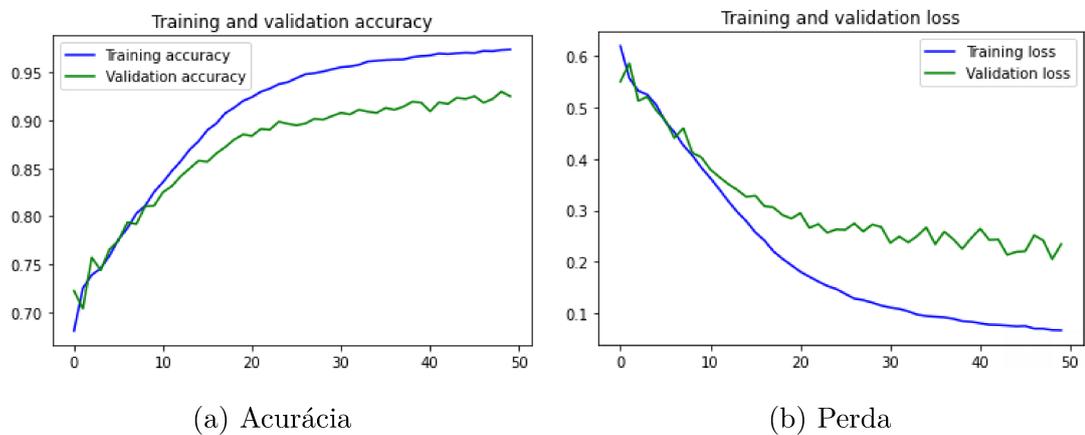


Figura 23 – Resultados do treinamento do modelo ResNet50.

superior na métrica acurácia, mas na média geral do desempenho da ResNet foi superior com uma acurácia de 90,60%.

Os resultados do segundo cenário são mostrados na Tabela 2. Nesta avaliação observa-se que os valores das medidas foram inferiores aos resultados das imagens originais, mostrando que os modelos são sensíveis as características de brilho das imagens, ou seja, as redes aprenderam as informações de intensidade de brilho sem o melhoramento inseridos pelo método de normalização. Um outro procedimento foi treinar com as imagens normalizadas e avaliar com as imagens originais (cenário 3). Os resultados são apresentados na Tabela 3. Essa investigação também mostra que a normalização provoca influência sobre os resultados das redes. As alterações nas características do contraste das imagens, treinamento e teste, influenciaram o desempenho dos modelos. O cenário 4 é apresentado na Tabela 4. Nesse experimento é possível observar que o modelo ResNet proporcionou os melhores resultados para todos os cenários (acurácia média de 97,01). No entanto, no modelo AlexNet esse comportamento não foi observado, ou seja, não usar a normalização é mais relevante para esse modelo.

Tabela 1 – Resultados (%) da classificação realizada com as imagens originais, sem investigação da normalização.

Modelo	Métrica	Saudável	Leve	Moderada	Severa	Média
AlexNet	Sensib.	91,29	78,12	80,08	68,97	79,60
	Especif.	90,82	91,19	91,95	98,88	93,20
	Precisão	76,86	74,21	77,28	95,35	80,90
	Acurácia	90,94	87,99	88,93	91,40	89,80
ResNet	Sensib.	65,65	82,78	79,17	97,63	81,30
	Especif.	99,65	93,20	91,64	90,54	93,80
	Precisão	98,45	79,80	76,40	77,47	83,00
	Acurácia	91,14	90,65	88,46	92,31	90,60

Tabela 2 – Resultados (%) da classificação com as imagens normalizadas empregadas apenas na fase de teste.

Modelo	Métrica	Saudável	Leve	Moderado	Severo	Média
AlexNet	Sensib.	7,15	61,28	29,78	39,09	34,33
	Especif.	93,32	58,86	80,39	79,79	78,09
	Precisão	26,34	32,58	34,17	39,19	33,07
	Acurácia	71,75	59,45	67,50	69,62	67,08
ResNet	Sensib.	0,11	81,90	27,65	9,48	29,78
	Especif.	99,77	33,70	73,45	99,35	76,57
	Precisão	13,51	28,61	26,25	82,99	37,84
	Acurácia	74,82	45,51	61,79	76,89	64,75

Tabela 3 – Classificação (%) com as imagens normalizadas empregadas na fase de treinamento e as imagens originais avaliadas na fase de teste.

Modelo	Métrica	Saudável	Leve	Moderado	Severo	Média
AlexNet	Sensib.	1,03	61,06	15,27	30,50	26,97
	Especif.	92,56	31,66	97,34	81,18	75,69
	Precisão	4,43	22,48	66,28	35,07	32,06
	Acurácia	69,65	38,86	76,44	68,51	63,37
ResNet	Sensib.	0,00	11,10	0,46	98,71	27,57
	Especif.	99,40	92,64	99,18	12,17	75,85
	Precisão	0,00	32,87	15,94	27,25	19,01
	Acurácia	74,51	72,67	74,03	33,80	63,75

Tabela 4 – Resultados (%) obtidos com as imagens normalizadas para as fases de treinamento e teste.

Modelo	Métrica	Saudável	Leve	Moderado	Severo	Média
AlexNet	Sensib.	93,27	76,90	31,59	35,54	59,33
	Especif.	64,73	83,93	97,74	99,17	86,39
	Precisão	46,90	60,83	82,69	93,45	70,97
	Acurácia	71,88	82,21	80,89	83,27	79,56
ResNet	Sensib.	96,61	93,00	90,74	95,80	94,03
	Especif.	98,32	98,74	97,57	97,39	98,01
	Precisão	95,06	96,00	92,73	92,44	94,06
	Acurácia	97,89	97,33	95,83	96,99	97,01

4.3 Investigação dos Algoritmos de Aprendizado de Máquina

Com objetivo de investigar as características aprendidas pela rede, modelos híbridos com extração de características baseados nas arquiteturas ResNet50 e Alexnet e algoritmos de aprendizado de máquina foram investigados.

Inicialmente, para avaliar as características extraídas dos modelos, foi realizado o teste estatístico de Mann-Whitney com uma função ECDF (*Empirical Cumulative Distribution Function*, ou Função de Distribuição Empírica Acumulativa). O teste de Mann-Whitney tem o objetivo de definir se os diferentes grupos de dados estudados pertencem ou não à mesma população. A partir do teste é calculado um p -valor, que representa a probabilidade da hipótese ser nula. Nesse caso, a hipótese de que características extraídas são separáveis é investigada, isto é, pertencem a grupos diferentes. Quanto menor for o p -valor melhor o resultado, respeitando sempre o intervalo de confiança de 95%.

A função ECDF é uma função acumulativa que tem como objetivo visualizar as distribuições dos resultados empíricos. Analisando as Figuras 24 e 25 é possível notar que grande parte das características obtidas da última camada dos modelos (*flattting*) possuem p -value próximo a zero. Isso significa que os resultados da extração de características são dados separáveis, isto é, realmente representam as classes.

Os resultados das classificações híbridas com os algoritmos RF e SVM são apresentadas nas Tabelas 5 e 6. Nesse experimentos foram considerados apenas os cenários i) e iv) devido aos resultados mais relevantes observados. Os dados das imagens sem uso de normalização mostram que essa abordagem híbrida foi promissora para os modelos das CNNs, pois a AlexNet tinha acurácia média de 89,80% foi para 92,73% com o algoritmo RF e 93,91% com o SVM. Já para a ResNet, o valor de acurácia era de 90,60% e passou para 97,34% com o RF e 97,05% com o SVM.

Com as imagens normalizadas, os valores médios das acurácias eram de 79,56% (AlexNet) e 97,01% (ResNet) passaram para 85,05% (AlexNet) e 97,05% (ResNet) com o RF. Com o SVM, esses valores foram modificados passando para 87,78% (AlexNet) e

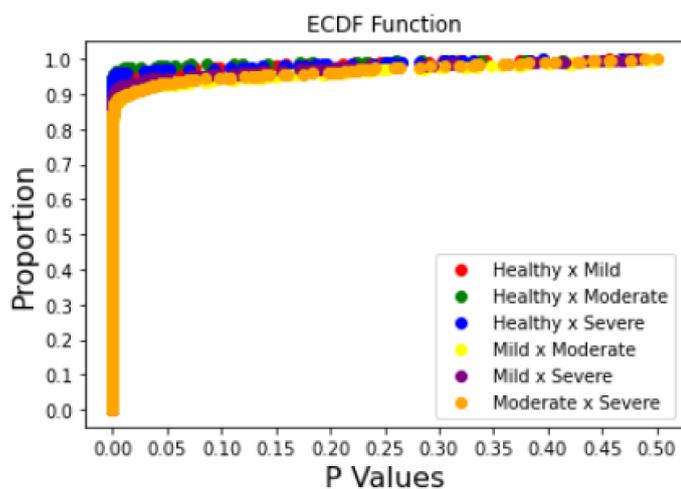


Figura 24 – Função ECDF para as características extraídas do modelo AlexNet

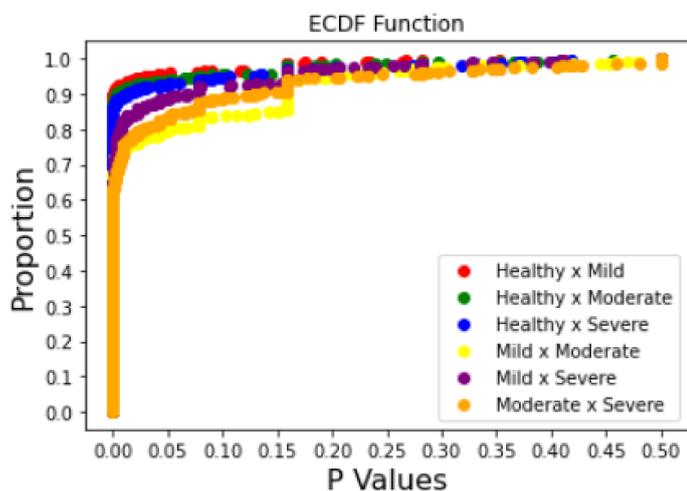


Figura 25 – Função ECDF para as características extraídas do modelo ResNet50

95,56% (ResNet). Nota-se que os algoritmos de aprendizado de máquina contribuíram de forma positiva com a arquitetura AlexNet na classificação das lesões. No entanto, os valores de acurácia ainda foram inferiores aos obtidos com a ResNet.

Tabela 5 – Valores das métricas para classificação dos modelos híbridos utilizando as imagens originais sem aplicação da normalização.

Modelo		Métrica	Saudável	Leve	Moderado	Severo	Média
RF	AlexNet	Sensib.	92,62	84,18	80,81	84,31	85,48
		Especif.	97,02	91,36	95,42	96,85	95,16
		Precisão	91,22	75,97	85,79	89,91	85,72
		Acurácia	95,92	89,60	91,70	93,71	92,73
	ResNet	Sensib.	97,53	93,90	91,32	96,03	94,70
		Especif.	99,13	97,01	97,95	98,82	98,23
		Precisão	97,41	91,06	93,84	96,46	94,69
		Acurácia	98,73	96,25	96,26	98,13	97,34
SVM	AlexNet	Sensib.	92,96	91,51	80,81	86,19	87,87
		Especif.	98,08	89,89	97,63	98,20	95,95
		Precisão	94,19	74,60	92,11	94,10	88,75
		Acurácia	96,80	90,29	93,35	95,20	93,91
	ResNet	Sensib.	98,08	93,36	90,28	94,70	94,11
		Especif.	98,39	96,76	97,94	99,04	98,03
		Precisão	95,33	90,35	93,74	97,04	94,11
		Acurácia	98,32	95,93	95,99	97,95	97,05

Tabela 6 – Resultados da classificação dos modelos híbridos utilizando as imagens normalizadas.

Modelo		Métrica	Saudável	Leve	Moderado	Severo	Média
RF	AlexNet	Sensib.	76,77	70,11	61,95	71,74	70,14
		Especif.	89,93	90,45	87,85	91,90	90,03
		Precisão	71,80	70,43	63,53	74,69	70,11
		Acurácia	86,63	85,47	81,25	86,86	85,05
	ResNet	Sensib.	96,69	94,23	91,36	94,19	94,12
		Especif.	98,87	98,36	97,00	97,89	98,03
		Precisão	96,63	94,92	91,24	93,70	94,12
		Acurácia	98,33	97,35	95,57	96,96	97,05
SVM	AlexNet	Sensib.	77,38	75,03	74,38	75,44	75,56
		Especif.	93,44	92,44	87,29	94,20	91,84
		Precisão	79,77	76,31	66,66	81,26	76,00
		Acurácia	89,42	88,18	84,00	89,51	87,78
	ResNet	Sensib.	93,30	89,53	92,70	88,91	91,11
		Especif.	98,80	98,43	92,82	98,09	97,03
		Precisão	96,30	94,86	81,52	93,93	91,65
		Acurácia	97,42	96,25	92,79	95,79	95,56

5 Conclusão

Neste trabalho foi realizado um estudo com parte de um sistema CAD que auxilie os profissionais especialistas em histologia a diagnosticar com maior exatidão a lesão pré-cancerígena chamada displasia. Visto que atualmente não existe uma opção definitiva de sistema de auxílio ao diagnóstico, é necessário o estudo e desenvolvimento de tecnologias capazes de exercerem esse auxílio. Além disso, a displasia é um tipo de lesão de difícil detecção, e por isso se faz importante o desenvolvimento de novas abordagens que possam melhorar a detecção para esta doença.

Desse modo, o modelo desenvolvido neste trabalho obteve resultados satisfatórios, pois a normalização foi realizada de forma a manter as características das imagens, e por consequência teve um impacto positivo na classificação para grupos de lesões em alguns dos cenários investigados. Os classificadores RF e SVM obtiveram resultados relevantes para as características obtidas com a CNN. Uma acurácia superior à 95% para o caso do modelo da arquitetura ResNet50. O estudo mostrou a importância dos modelos híbridos que combina arquiteturas CNNs para extração de características e algoritmos de ML para a classificação. Os resultados foram melhores do que os modelos que usam apenas a camada *softmax*.

As principais metas do estudo foram alcançadas por meio de um estudo introdutório à histologia, necessário para compreender o problema abordado, além do estudo de disciplinas do curso de Ciência da Computação, como Multimídia, para realizar o pré-processamento das imagens, Inteligência Computacional e Inteligência Artificial, para compreender os modelos de redes neurais convolucionais e os algoritmos de aprendizado de máquina. Além disso, a disciplina de Estatística foi de grande importância para interpretação e análise dos resultados apresentados nesse trabalho.

5.1 Desafios Encontrados

Durante o desenvolvimento do projeto foram encontradas algumas dificuldades, as quais são descritas aqui: i) entendimento de tópicos relacionados a Histologia, em razão falta de experiência e conhecimento prévio da área; ii) investigação sobre processo de normalização de corantes H&E e possíveis representações com simulações computacionais; iii) adaptações em frameworks que trabalham com CNN devido a limitação dos recursos computacionais e tempo de processamento para execução dos experimentos.

Esses desafios serviram como grandes oportunidades para incremento dos conhecimentos já obtidos, bem como para reforçar a importância de várias práticas e conceitos vistos durante o decorrer do curso. Além disso, a superação dos desafios encontrados con-

tribui para a consolidação de uma aplicação cada vez mais robusta que possa contribuir para área interdisciplinar de biologia em conjunto com a computação.

5.2 Trabalhos Futuros

Futuramente, pretende-se investigar outros algoritmos de normalização de corantes para uma comparação em relação ao desempenho do método investigado. Também serão investigados outras abordagens que trabalham com modelos GANs. Ainda pretende-se realizar a implementação dessas ferramentas em uma plataforma Web para que os especialistas possam utilizar no contexto de melhorar a normalização de corantes como também realizar a classificação de imagens histológicas.

Referências

- ALOM TAREK M. TAHA, C. Y. S. W. P. S. M. S. N. B. C. V. E. A. A. S. A. M. Z.; ASARI, V. K. The history began from alexnet: A comprehensive survey on deep learning approaches. *arXiv.org*, 2018. Citado na página 34.
- BENTAIEB, A.; HAMARNEH, G. Adversarial stain transfer for histopathology image analysis. *IEEE transactions on medical imaging*, IEEE, v. 37, n. 3, p. 792–802, 2017. Citado 2 vezes nas páginas 12 e 17.
- BISSOTO, A. E. *Improving Skin Lesion Analysis with Generative Adversarial Networks*. Dissertação (Mestrado) — Universidade Estadual de Campinas - Instituto de Computação, Campinas, 2019. Citado 2 vezes nas páginas 5 e 26.
- FONSECA-SILVA, T. et al. Association between histopathological features of dysplasia in oral leukoplakia and loss of heterozygosity. *Histopathology*, Wiley Online Library, v. 68, n. 3, p. 456–460, 2016. Citado na página 11.
- GONZALEZ, R. C.; WOODS, R. *Digital Image Processing*. [S.l.]: Pearson, 2018. Citado na página 12.
- GOODFELLOW, I. J. et al. Generative adversarial nets. *arXiv.org*, 2014. Citado na página 26.
- HAM, A.; CORMACK, D. H. Histologia. In: *Histologia*. [S.l.: s.n.], 1983. p. 906–906. Citado na página 11.
- INCA. *Instituto Nacional de Câncer*. 2019. Disponível em: <<https://www.inca.gov.br>>. Acesso em: 10 abr. 2021. Citado na página 11.
- ISOLA, P. et al. Image-to-image translation with conditional adversarial networks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2017. p. 1125–1134. Citado na página 32.
- JOTHI, J. A. A.; RAJAM, V. M. A. A survey on automated cancer diagnosis from histopathology images. *Artificial Intelligence Review*, Springer, v. 48, n. 1, p. 31–81, 2017. Citado na página 11.
- JUNIOR, W. L. M. *Uma Nova Abordagem de Descritor de Textura Baseada em Transformada Ripplet para Classificação de Lesões da Mama*. Dissertação (Mestrado) — Universidade Federal do ABC, 2018. Citado na página 16.
- KOWALCZYK, A. *Support Vector Machines Succinctly*. [S.l.]: Syncfusion, 2017. Citado 2 vezes nas páginas 5 e 22.
- LI, W. et al. Path r-cnn for prostate cancer diagnosis and gleason grading of histological images. *IEEE transactions on medical imaging*, IEEE, v. 38, n. 4, p. 945–954, 2018. Citado na página 12.

- LOUPPE, G. *Understanding Random Forests from theory to practice*. Tese (Doutorado) — University of Liège - Faculty of Applied Sciences Department of Electrical Engineering Computer Science, 2015. Citado 3 vezes nas páginas 5, 20 e 21.
- MACENKO M. NIETHAMMER, J. S. M. D. B. J. T. W. X. G. C. S. M.; THOMAS, N. E. *A method for normalizing histology slides for quantitative analysis*. [S.l.]: IEEE Int. Symp. Biomed. Imaging From Nano to Macro, 2009. Citado na página 27.
- MICHAIL, E. et al. Detection of centroblasts in h&e stained images of follicular lymphoma. In: IEEE. *2014 22nd Signal Processing and Communications Applications Conference (SIU)*. [S.l.], 2014. p. 2319–2322. Citado na página 12.
- MIRZA, M.; OSINDERO, S. Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*, 2014. Citado na página 32.
- RAMADAN, S. Z. Methods used in computer-aided diagnosis for breast cancer detection using mammograms: A review. *Hindawi - Journal of Healthcare Engineering*, 2020. Citado 2 vezes nas páginas 5 e 17.
- RAMOS, R.; NASCIMENTO, M.; PEREIRA, D. Texture extraction: An evaluation of ridgelet, wavelet and co-occurrence based methods applied to mammograms. *Expert Systems with Applications*, Elsevier, 2012. Citado na página 20.
- REINHARD M. ASHIKHMIN, B. G. E.; SHIRLEY, P. *Color transfer between images*. [S.l.]: IEEE Comput. Graph. Appl., vol. 21, no. 5, 2001. Citado na página 27.
- ROBERTO et al. Fractal neural network: A new ensemble of fractal geometry and convolutional neural networks for the classification of histology images. *Expert Systems with Applications*, Elsevier, v. 166, p. 114103, 2021. Citado na página 12.
- ROSA, J. L. G. *Fundamentos da Inteligência Artificial*. São Paulo: Grupo Editorial Nacional, 2011. Citado na página 17.
- ROSEBROCK, D. A. *Deep Learning for Computer Vision with Python*. [S.l.]: PyImageSearch, 2017. Citado 5 vezes nas páginas 5, 18, 19, 23 e 24.
- SALEHI, P.; CHALECHALE, A. Pix2pix-based stain-to-stain translation: A solution for robust stain normalization in histopathology images analysis. *2020 International Conference on Machine Vision and Image Processing (MVIP)*, 2020. Citado 9 vezes nas páginas 5, 12, 17, 26, 27, 28, 31, 32 e 33.
- SILVA, A. B. *Métodos Computacionais para Análise e Classificação de Displasias em Imagens da Cavidade Bucal*. Dissertação (Mestrado) — Universidade Federal de Uberlândia - Faculdade de Computação, 2019. Citado 6 vezes nas páginas 5, 11, 12, 14, 15 e 29.
- TOSTA, T. A. A. et al. Computational normalization of h&e-stained histological images: Progress, challenges and future potential. *Artificial intelligence in medicine*, Elsevier, v. 95, p. 118–132, 2019. Citado 2 vezes nas páginas 12 e 17.
- VAHADANE T. PENG, A. S. S. A. L. W. M. B. K. S. A. M. S. I. E. A.; NAVAB, N. *Structure-Preserving Color Normalization and Sparse Stain Separation for Histological Images*. [S.l.]: IEEE Trans. Med. Imaging, vol. 35, no. 8, 2016. Citado na página 27.

VAPNIK, V. An overview of statistical learning theory. *Neural Networks, IEEE Transactions on*, IEEE, v. 10, n. 5, p. 988–999, 1999. Citado na página 21.

YU LAURENCE T. YANG, Q. Z. D. A. M. J. D. H. Convolutional neural networks for medical image analysis: State-of-the-art, comparisons, improvement and perspectives. *Elsevier*, 2021. Citado 4 vezes nas páginas 5, 6, 34 e 35.