



CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
PROGRAMA DE PÓS-GRADUAÇÃO

CLASSIFICAÇÃO AUTOMÁTICA DE LEUCÓCITOS EM AMOSTRAS SANGUÍNEAS DE CÃES BASEADA EM TÉCNICAS DE VISÃO COMPUTACIONAL E REDES NEURAIS CONVOLUCIONAIS

RÓGENES REIS FIRMINO

Orientador: Flávio Luis Cardeal Pádua
Centro Federal de Educação Tecnológica de Minas Gerais

Coorientador: Marco Túlio Alves Nolasco Rodrigues
Universidade de Itaúna

Coorientadora: Natália Cosse Batista
Centro Federal de Educação Tecnológica de Minas Gerais

BELO HORIZONTE
AGOSTO DE 2023



SERVIÇO PÚBLICO FEDERAL
MINISTÉRIO DA EDUCAÇÃO
CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
COORDENAÇÃO DO CURSO DE MESTRADO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL

***“CLASSIFICAÇÃO AUTOMÁTICA DE LEUCÓCITOS EM
AMOSTRAS SANGUÍNEAS DE CÃES BASEADA EM TÉCNICAS DE
VISÃO COMPUTACIONAL E REDES NEURAIS
CONVOLUCIONAIS”***

Dissertação de Mestrado apresentada por **Rógenes Reis Firmino**, em 07 de agosto de 2023, ao Programa de Pós-Graduação em Modelagem Matemática e Computacional do CEFET-MG, e aprovada pela banca examinadora constituída pelos professores:

Prof. Dr. Flávio Luis Cardeal Pádua
Centro Federal de Educação Tecnológica de Minas Gerais

Prof. Dr. Marco Túlio Alves Nolasco Rodrigues
Centro Federal de Educação Tecnológica de Minas Gerais

Prof^ª. Dr^ª. Natália Cosse Batista
Centro Federal de Educação Tecnológica de Minas Gerais

Prof^ª. Dr^ª. Fabíola de Oliveira Paes Leme
Universidade Federal de Minas Gerais

Prof. Dr. Gustavo Campos Menezes
Centro Federal de Educação Tecnológica de Minas Gerais

Visto e permitida à impressão,

Prof. Dr. Thiago de Souza Rodrigues
Coordenador do Programa de Pós-Graduação em
Modelagem Matemática e Computacional

Firmino, Rógenes Reis
F525c Classificação automática de leucócitos em amostras sanguíneas de cães baseada em técnicas de visão computacional e redes neurais convolucionais / Rógenes Reis Firmino. – 2023.
58 f.

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Modelagem Matemática e Computacional.
Orientador: Flávio Luis Cardeal Pádua.
Coorientador: Marco Túlio Alves Nolasco Rodrigues.
Dissertação (mestrado) – Centro Federal de Educação Tecnológica de Minas Gerais.

1. Leucócitos – Detecção – Teses. 2. Leucócitos – Classificação – Teses. 3. Visão por computador – Teses. 4. Redes neurais (Computação) – Teses. I. Pádua, Flávio Luís Cardeal. II. Rodrigues, Marco Túlio Alves Nolasco. III. Centro Federal de Educação Tecnológica de Minas Gerais. IV. Título.

CDD 006.31

Agradecimentos

Primeiramente agradeço a Deus por conceder a mim sabedoria, força e determinação para a finalização deste trabalho.

Agradeço aos meus pais, Geraldo e Rose, por todo o amor, suporte e encorajamento, mesmo frente às grandes dificuldades.

Agradeço à minha amada Isabela, por todo o amor, compreensão e companheirismo ao longo de toda essa jornada.

Agradeço ao meu orientador, Flávio Luis Cardeal Pádua, por aceitar conduzir este trabalho, além de todo o apoio disponibilizado.

Agradeço aos meus coorientadores, Marco Túlio Alves Nolasco Rodrigues e Natália Cosse Batista por todo o acompanhamento, auxílio e incentivo durante este período.

Agradeço à equipe parceira da veterinária da UFMG, em especial professora Fabíola e Lucas, por auxiliarem e darem suporte necessário até as etapas finais deste trabalho.

Por fim, agradeço a todos os que contribuíram, de alguma forma, para a realização deste trabalho.

Resumo

Na rotina clínica veterinária, o hemograma é o exame complementar mais solicitado, cujo objetivo é avaliar as diferentes células sanguíneas do animal para auxiliar no diagnóstico e acompanhamento da evolução das doenças. O hemograma examina as diferentes células do sangue animal de modo a realizar a contagem de hemácias, leucócitos e plaquetas em uma amostra sanguínea. Os métodos atuais para contagem e análise das células sanguíneas podem ser divididos em técnicas manuais, impedância, análise quantitativa da camada flogística, citometria de fluxo e por visão computacional. A técnica manual é morosa e apresentam um fluxo de trabalho repetitivo, o que pode gerar influências negativas nos resultados apresentados, pois além do prazo extenso para se obter essa resposta, ela ainda pode conter os efeitos subjetivos do profissional. Já as técnicas automáticas, apresentam outros tipos de problemas, como a baixa exatidão na contagem diferencial de leucócitos na técnica de impedância e o alto custo de aquisição e manutenção dos equipamentos relacionados à citometria de fluxo. Este trabalho propõe o desenvolvimento de um método para localizar e classificar automaticamente leucócitos caninos por meio de técnicas de visão computacional e aprendizado de máquina, levando em consideração a variação morfológica celular do animal, a partir de imagens de esfregaços sanguíneos. Dessa forma, é possível fornecer resultados mais precisos do que as metodologias existentes, reduzindo o efeito da subjetividade dos resultados, além de diminuir o tempo gasto com a leitura de lâmina do esfregaço, liberando o médico veterinário para o desempenho de outras atividades. Para a construção do método, uma base de imagens com 433 exemplos de amostras sanguíneas de cães foi criada. Em seguida, a partir da análise de diversos espaços de cores, a componente S do espaço HSV proporcionou o melhor realce dos núcleos dos leucócitos. Com o uso de um limiar, as células sanguíneas de interesse foram localizadas na cena, identificando corretamente 1390 elementos celulares de um total de 1436, o que corresponde a uma taxa de acertos de 96.76%. Por fim, com o auxílio da VGG16, uma rede neural convolucional com transferência de aprendizado, uma acurácia média de 98,5% foi obtida para as células das seguintes classes: linfócitos, neutrófilos, eosinófilos, monócitos e metarrubríctos.

Palavras-chave: Leucócitos. Visão computacional. Detecção e classificação automática de leucócitos. Redes Neurais Convolucionais.

Abstract

In veterinary clinical practice, the blood count is the most commonly requested complementary examination. Its objective is to evaluate different blood cells in animals, assisting in the diagnosis and monitoring of disease progression. The blood count analyzes and counts various blood cells, including red blood cells, leukocytes, and platelets in a blood sample. Current methods for blood cell counts can be classified into manual techniques, impedance, quantitative analysis of the phlogistic layer, flow cytometry, and computer vision. The manual technique is time-consuming and involves a repetitive workflow, which can negatively impact the accuracy of the results. Besides the prolonged waiting period for obtaining results, there is also a risk of subjective interpretation by the professional. On the other hand, automatic techniques present their own set of issues. The impedance technique, for instance, has low accuracy in the differential leukocyte count. Additionally, flow cytometry equipment is costly to acquire and maintain. This study proposes the development of a computer vision and machine learning-based method to automatically locate and classify canine leukocytes, considering the morphological variations of the cells. The method utilizes images of blood smears to achieve this. By employing this approach, more accurate results can be obtained compared to existing methodologies. It minimizes the subjectivity of the results and reduces the time required for reading smear slides, enabling veterinarians to allocate their time to other activities. To construct the method, a database of 433 examples of blood samples from dogs was created. After analyzing various color spaces, it was determined that the S component of the HSV color space provided the best enhancement of the leukocyte nuclei. By applying a threshold, the blood cells of interest were successfully located in the scene, correctly identifying 1390 cell elements out of a total of 1436, resulting in a hit rate of 96.76%. Finally, utilizing VGG16, a convolutional neural network with transfer learning, an average accuracy of 98.5% was achieved for classifying lymphocytes, neutrophils, eosinophils, monocytes, and metarubricites.

Keywords: Leukocytes. Computer vision. Leukocytes automatic detection and classification. Convolutional Neural Network.

Lista de Figuras

Figura 1 – Diagrama de um sistema de visão computacional	16
Figura 2 – Orientação de uma imagem digital	17
Figura 3 – Imagem RGB e suas componentes	18
Figura 4 – Cores primárias aditivas	19
Figura 5 – Cubo RGB	20
Figura 6 – Cilindro HSV	21
Figura 7 – Imagens original e segmentada	23
Figura 8 – Estágios do sistema nervoso humano	25
Figura 9 – Neurônio artificial	26
Figura 10 – Perceptron Multi Camadas	27
Figura 11 – Arquitetura de uma CNN	29
Figura 12 – Transferência de aprendizado	30
Figura 13 – Imagem de amostra sanguínea	35
Figura 14 – Componentes dos espaços de cor utilizados para estudo de segmentação	37
Figura 15 – Segmentação sem refinamento	38
Figura 16 – Segmentação com refinamento	39
Figura 17 – Recorte em cada célula	39
Figura 18 – Exemplos das células capturadas	41
Figura 19 – Scanner 3D	43
Figura 20 – Segmentação incorreta - Eliminação	45
Figura 21 – Segmentação incorreta - Junção e divisão	45

Lista de Tabelas

Tabela 1 – Resumo dos trabalhos relacionados	34
Tabela 2 – Modelo 1 sem <i>data augmentation</i> - Matriz de confusão e métricas	48
Tabela 3 – Modelo 2 sem <i>data augmentation</i> - Matriz de confusão e métricas	48
Tabela 4 – Modelo 3 sem <i>data augmentation</i> - Matriz de confusão e métricas	49
Tabela 5 – Modelo 4 sem <i>data augmentation</i> - Matriz de confusão e métricas	49
Tabela 6 – Modelo 5 sem <i>data augmentation</i> - Matriz de confusão e métricas	49
Tabela 7 – Modelo 1 com <i>data augmentation</i> - Matriz de confusão e métricas	50
Tabela 8 – Modelo 2 com <i>data augmentation</i> - Matriz de confusão e métricas	50
Tabela 9 – Modelo 3 com <i>data augmentation</i> - Matriz de confusão e métricas	50
Tabela 10 – Modelo 4 com <i>data augmentation</i> - Matriz de confusão e métricas	51
Tabela 11 – Modelo 5 com <i>data augmentation</i> - Matriz de confusão e métricas	51
Tabela 12 – Média das métricas dos modelos sem <i>data augmentation</i>	51
Tabela 13 – Média das métricas dos modelos com <i>data augmentation</i>	51
Tabela 14 – Média das métricas - 5 treinamentos com augmentation - Base LISC	52
Tabela 15 – Taxa de acerto dos profissionais	52

Lista de Abreviaturas e Siglas

ABNT	Associação Brasileira de Normas Técnicas
DECOM	Departamento de Computação
HSI	Hue-Saturation-Intensity
HSV	Hue-Saturation-Value
SVM	Support Vector Machine
CNN	Convolutional Neural Network
HSG	Hue-Saturation-Green
RGB	Red-Green-Blue
TLU	Threshold logic unit

Lista de Símbolos

μ	Intensidade do pixel
Ω	Representação de uma imagem digital

Sumário

1 – Introdução	11
1.1 Motivação	12
1.2 Problema de pesquisa	12
1.3 Objetivos	13
1.4 Contribuições	14
1.5 Organização do trabalho	14
2 – Fundamentação Teórica	15
2.1 Sistema de visão computacional	15
2.2 Imagens Digitais	17
2.2.1 Espaço de cor	19
2.2.1.1 Espaço RGB	20
2.2.1.2 Espaço HSV	21
2.3 Segmentação	22
2.4 Aprendizado de máquina	23
2.4.1 Redes neurais - O cérebro humano	25
2.4.2 Redes neurais artificiais	26
2.4.2.1 Perceptron	27
2.4.2.2 Perceptron Multi Camadas	27
2.4.2.3 <i>Backpropagation</i>	28
2.4.3 Redes Neurais Profundas	28
2.4.3.1 Redes Neurais Convolucionais	28
2.4.3.2 Transferência de aprendizado	29
2.5 Conclusão	31
3 – Trabalhos Relacionados	32
3.1 Conclusão	34
4 – Metodologia	35
4.1 Segmentação	36
4.1.1 Espaços de cor	36
4.1.2 Contornos dos núcleos	36
4.1.3 Refinamento	37
4.1.4 Localização	38
4.2 Classificador	39
5 – Resultados	42

5.1	Configuração de experimentos	42
5.1.1	Montagem da base de dados	42
5.1.2	Métricas de avaliação	43
5.1.3	Segmentação	44
5.1.4	Classificação	46
5.1.5	Testes em outra bases de dados	46
5.1.6	Testes com patologistas	47
5.2	Resultados	47
5.3	Análise e discussão dos resultados	53
6	– Conclusão	54
6.1	Próximos passos	55
	Referências	56

1 Introdução

O sangue é um importante tipo de tecido conjuntivo presente na grande maioria dos animais. Ele é formado basicamente por dois grupos, o plasma e os elementos celulares. O plasma representa a parte líquida do sangue e, representando aproximadamente 55% do volume desse tecido, é composto basicamente por água e alguns componentes dissolvidos, como proteínas, vitaminas, entre outros. Já os elementos celulares são compostos basicamente por células vermelhas, células brancas e plaquetas (KUAN et al., 2018).

Os elementos celulares possuem funcionalidades variadas. De acordo com Meijden e Heemskerk (2019), um dos propósitos das plaquetas (trombócitos) no corpo é realizar a manutenção das paredes dos vasos sanguíneos. Já Alzubaidi et al. (2020) informam que células vermelhas (eritrócitos) são responsáveis por carregar oxigênio e assim realizar o transporte deste para todo o corpo. Em relação às células brancas (leucócitos), Reena e Ameer (2020) alegam que possuem como principal função combater infecções e doenças.

Existem diversas doenças que afetam as células sanguíneas, como a anemia falciforme, condição em que inexistem células vermelhas saudáveis o suficiente para que o oxigênio seja distribuído por todo o sistema do indivíduo, o que pode causar fadiga, dores, infecções e até outras complicações (STUART; NAGEL, 2004). Por outro lado, o tecido sanguíneo pode, de certo modo, ser utilizado para a análise do estado da saúde do indivíduo em um âmbito mais geral, não somente focalizando em doenças presentes no sangue. Por serem agentes defensivos do organismo, os leucócitos se multiplicam rapidamente quando percebem a entrada de corpos estranhos, germes e agentes infecciosos no corpo, de modo a combatê-los. Devido a esse comportamento, a quantidade de leucócitos presente em determinada amostra de sangue pode auxiliar no diagnóstico de doenças como leucemia e câncer (YILDIRIM; CINAR, 2019).

Segundo (SANTOS; LEME, 2021), além de atuar como uma excelente ferramenta de triagem no campo veterinário, o hemograma pode determinar o diagnóstico definitivo de certas enfermidades, ao empregar diversas técnicas, manuais e automáticas, para a análise celular dos animais em questão. Estes fatos evidenciam a importância, não só da realização do exame em si, mas também da obtenção de resultados devidamente acurados obtidos por meio dele.

Ainda segundo (SANTOS; LEME, 2021), o leucograma, cujo objetivo é realizar a contagem total e por classe de leucócitos (linfócitos, neutrófilos, monócitos, eosinófilos e basófilos) é parte fundamental de um hemograma veterinário. Junto do eritrograma e do plaquetograma, é capaz de expor quantidade significativa de informações clínicas para o médico veterinário e fará parte do foco principal deste trabalho.

1.1 Motivação

De acordo com [Lin e Wang \(2018\)](#), atualmente existem métodos manuais e automáticos para a contagem e reconhecimento de leucócitos presentes em uma amostra sanguínea. Dentre os métodos manuais, é possível citar a inspeção manual das células, realizada por patologistas profissionais e entre as automáticas, existem técnicas de citometria de fluxo e contagem automática.

Ambos os casos apresentam potencial de evolução e pontos que, se refinados, trariam melhorias para o processo de análise do sangue de uma maneira geral. Segundo [Lin e Wang \(2018\)](#), os métodos automáticos citados não conseguem validar qualitativamente as células brancas, apenas contá-las. No caso da citometria de fluxo, existe o agravante do alto custo para aquisição e manutenção dos equipamentos, enquanto a contagem diferencial dos leucócitos no método de impedância apresenta baixa exatidão ([SANTOS; LEME, 2021](#)). Os métodos manuais por sua vez, possuem outros tipos de problemas: segundo [Saraswat e Arya \(2014\)](#), o processo manual despende muito tempo, é moroso e sempre depende de um profissional treinado para a realização da leitura do esfregaço sanguíneo. Esta técnica consiste na análise dos elementos celulares por meio da disposição de uma gota de sangue em uma lâmina de vidro. O sangue então passa por um processo de fixação e coloração, a fim de que possa possibilitar a observação das células no microscópio. Mesmo contando com a qualificação do profissional, este está sujeito a vieses naturais, que podem ser causados pela própria grande variabilidade biológica presente nos tecidos sanguíneos.

Ao analisar a relação entre as demandas do hemograma, principalmente quanto a análise de leucócitos e os pontos de melhoria existentes nos métodos atuais, é possível observar indicativos de que há espaço para inovação e melhoria nos processos envolvidos, inclusive com o auxílio de tecnologia na esfera computacional. Atualmente existem métodos compostos por técnicas de visão computacional e inteligência artificial que se dispõem a tentar resolver o problema da contagem e classificação de leucócitos em amostras sanguíneas, porém pouquíssimos trabalhos abordam essa tema com foco em animais domésticos, como cães. A expressiva maioria é focada em células de origem humana, portanto um método que trate especificamente de células caninas tem muito a agregar nesse âmbito, pois além de propor uma resolução do problema da classificação dos leucócitos para cães, abre precedentes para a pesquisa dessa mesma matéria em outros animais, evidenciando e revisitando este assunto no meio científico.

1.2 Problema de pesquisa

Apesar de os métodos manuais de classificação de leucócitos geralmente alcançarem uma boa acurácia, como afirma [Zhao et al. \(2017\)](#), seus pontos negativos podem impactar muito no processo de análise sanguínea.

Ao se analisar manualmente amostras sanguíneas de animais em busca de células de interesse, é necessário que um profissional capacitado e treinado, com um esfregaço em mãos,

encontre e classifique os elementos manualmente. Este processo é naturalmente demorado, além de ser moroso para o profissional, que uma vez afetado pelo cansaço, pode ter diminuída sua percepção e análise, fazendo com que decaia o resultado da classificação.

Além da acurácia, o fator tempo é de suma importância para que demandas cada vez mais altas possam ser atendidas com mais agilidade, evitando prazos de espera muito longos por parte dos interessados.

Este trabalho se propõe a detectar e classificar automaticamente alguns tipos de leucócitos presentes em amostras de sangue de cães, a fim de auxiliar profissionais, estudantes e educadores durante o processo de análise sanguínea desses animais, aumentando o desempenho e qualidade dos resultados produzidos por essa atividade.

Dessa maneira o problema de pesquisa pode ser definido da seguinte maneira:

Dada uma imagem de amostra sanguínea de determinado cão, adquirida por microscópio, é possível dizer, por meio de um algoritmo composto por um método automático para detecção, segmentação e classificação de leucócitos, quantas dessas células estão presentes, assim como seus respectivos tipos ou classes. A saída do método é uma imagem contendo os rótulos dos leucócitos presentes na amostra.

Para que o método possa funcionar de maneira correta, existem algumas restrições que devem ser consideradas quando esse for efetivamente utilizado:

- As novas imagens utilizadas devem ter dimensões próximas às da base de dados utilizada para construção do método (1145 píxeis de largura e 707 de altura) e se possível, adquiridas por equipamento igual ou semelhante
- O corante utilizado para gerar o realce nos leucócitos nas novas imagens deve ser do mesmo tipo (ou o mais próximo possível) do utilizado para a construção da base de imagens (panótico rápido)

1.3 Objetivos

O objetivo principal deste trabalho é desenvolver um método que, dada uma imagem de determinada amostra sanguínea de cão, seja capaz de localizar e classificar os leucócitos presentes. O método proposto é o cerne de uma possível aplicação maior, seja *Web*, *Mobile* ou *Desktop*, que visa auxiliar no processo de classificação dos leucócitos, podendo ser utilizada em laboratórios, por exemplo. Para que o método seja justificável, ele deve atender alguns requisitos mínimos de acurácia e tempo, próximos ou superiores aos métodos manuais que apresentam resultados semelhantes.

Para que o objetivo principal seja concluído, os seguintes objetivos específicos devem ser concretizados:

- Criar base de dados com imagens de amostras sanguíneas de cães a serem estudadas, contendo os diversos tipos de leucócitos;
- Estudar os padrões e as características que diferenciam os tipos de leucócitos;
- Definir as melhores técnicas de tratamento de imagens, visão computacional e de aprendizado de máquina que mais se adéquem ao problema em questão;
- Desenvolver um método para detecção e segmentação automática dos leucócitos;
- Desenvolver um método a classificação automática dos leucócitos encontrados pelo método de detecção e segmentação;

1.4 Contribuições

O desenvolvimento de um método específico para a detecção e classificação de leucócitos em células animais pode ser de grande valia para diversos interesses. O método pode ser incorporado em um sistema maior de classificação e pode ser utilizado no processo de educação de alunos da área veterinária, acelerando análises, por exemplo, e ainda conta com potencial mercadológico.

Um produto contendo uma solução de classificação automática de leucócitos em cães pode trazer diversos benefícios para a população, principalmente para aquela envolvida diretamente com essa espécie. O aumento da agilidade e acurácia dos resultados de processos laboratoriais envolvidos em diagnósticos de doenças caninas, pode trazer maior bem-estar aos animais, por auxiliar na aceleração da inicialização de um possível tratamento.

Além das contribuições citadas anteriormente, este trabalho pode servir de base para trabalhos futuros que possuam objetivos similares, dado que todo um arcabouço será gerado em relação a este tema, como a construção de base de dados e análise de técnicas específicas para resolver o problema.

1.5 Organização do trabalho

Este trabalho está dividido em 6 capítulos, incluindo o presente. O [Capítulo 2](#) apresenta e descreve os principais conceitos e técnicas relacionadas ao tema deste trabalho, compondo a fundamentação teórica. O [Capítulo 3](#) é um apanhado de técnicas, ferramentas e métodos utilizados atualmente para solucionar problemas similares ao encontrado neste trabalho. É a apresentação do estado da arte por meio de trabalhos relacionados. O [Capítulo 4](#) trata da metodologia, descrevendo os passos, ferramentas e métodos utilizados para solucionar o problema descrito inicialmente. O [Capítulo 5](#) apresenta os experimentos e resultados encontrados ao longo do desenvolvimento deste trabalho e por fim, o [Capítulo 6](#) traz as conclusões sobre o trabalho, assim como perspectivas sobre possíveis continuações por meio de trabalhos futuros.

2 Fundamentação Teórica

Este capítulo explica as principais técnicas, métodos e ferramentas necessárias para a compreensão deste trabalho. Serão abordados conceitos relacionados à visão computacional, como processamento de imagens digitais, espaços de cor e processos de segmentação. Além disso, noções de aprendizado de máquina, métodos de classificação e redes neurais artificiais. Essas duas áreas são a base técnica para o desenvolvimento deste trabalho, pois para se construir um método de classificação automática de leucócitos, é necessário que haja a localização das células de maneira automática (por meio de técnicas de visão computacional) e em seguida o reconhecimento ou classificação desses elementos (utilizando aprendizado de máquina).

Além das informações a respeito da base para a construção de um método automático de classificação de leucócitos, serão apresentadas informações sobre os métodos não automáticos utilizados atualmente, importantes para informar sobre as diretrizes do que seria um bom resultado esperado ao se utilizar um processo automático guiado por métodos computacionais.

2.1 Sistema de visão computacional

Um sistema de visão computacional é inicialmente criado para resolver um problema utilizando como base, imagens. No caso deste trabalho, o problema a ser resolvido é de realizar a classificação automática de leucócitos de cães utilizando imagens digitais. Geralmente esse tipo de sistema é composto por um ou mais dispositivos de aquisição de imagens (câmeras, sensores, entre outros) e algoritmos que, após análise baseada nas informações presentes nas imagens adquiridas, produzem determinada saída ou resultado.

A respeito do entendimento sobre o que efetivamente é um sistema de visão computacional [Filho e Neto \(1999\)](#) afirmam que sistemas de visão computacional são aqueles capazes de adquirir, processar e interpretar imagens correspondentes a cenas reais. Já no trabalho de [Klette \(2014\)](#) é relatado que a visão computacional está associada à utilização de câmeras para entender e analisar o mundo real mediante imagens digitais, enquanto em [Gonzalez, Woods et al. \(2002\)](#) é explicado que a visão computacional é uma ciência que dá permissão às máquinas enxergarem, tornando atividades como detectar objetos em imagens de maneira automática, algo real.

A [Figura 1](#) exemplifica um sistema de visão computacional, contando com etapas que vão desde a aquisição das imagens, até o resultado. A etapa de aquisição geralmente utiliza equipamentos físicos como câmeras e sensores para transformar a cena do mundo real que se deseja estudar em uma imagem digital. A etapa de pré-processamento consiste em preparar a imagem para a análise, aplicando diversas técnicas, como redução de ruído, transformação do espaço de cor, binarização, entre outras. Após esse processo, são geralmente utilizadas técnicas para separar o objeto de interesse do fundo, a chamada segmentação. No caso deste trabalho, por exemplo, os objetos de interesse são os leucócitos, enquanto o fundo é representado pelo restante

dos elementos sanguíneos, como plasma e eritrócitos. Após a etapa de segmentação, a imagem (que já está devidamente modificada pelos processos anteriores) chega aos processos finais: extração de características e reconhecimento. De maneira geral essas últimas etapas utilizam o auxílio de técnicas de aprendizado de máquina, pois normalmente há padrões a serem aprendidos e reconhecidos, como o que é e o que não é um leucócito, ou a qual classe de glóbulos brancos essa célula pertence. Por fim, após a imagem passar por todas as etapas do processo, é obtido um resultado, que responde à pergunta inicial do problema em questão. Por fim ainda é possível citar o bloco da base de conhecimento, que se comunica e é paralelo a todas as etapas do processo. Pode representar, por exemplo, o tamanho em píxeis de um leucócito considerado aceitável durante o processo de segmentação, sendo que esse limite só pode ser conhecido por estudo da própria base de imagens, realizando diversos testes empíricos. Outra componente dessa base, pode ser o limite do valor da saturação de determinado elemento, para caracterizá-lo como leucócito ou como elemento de não interesse.

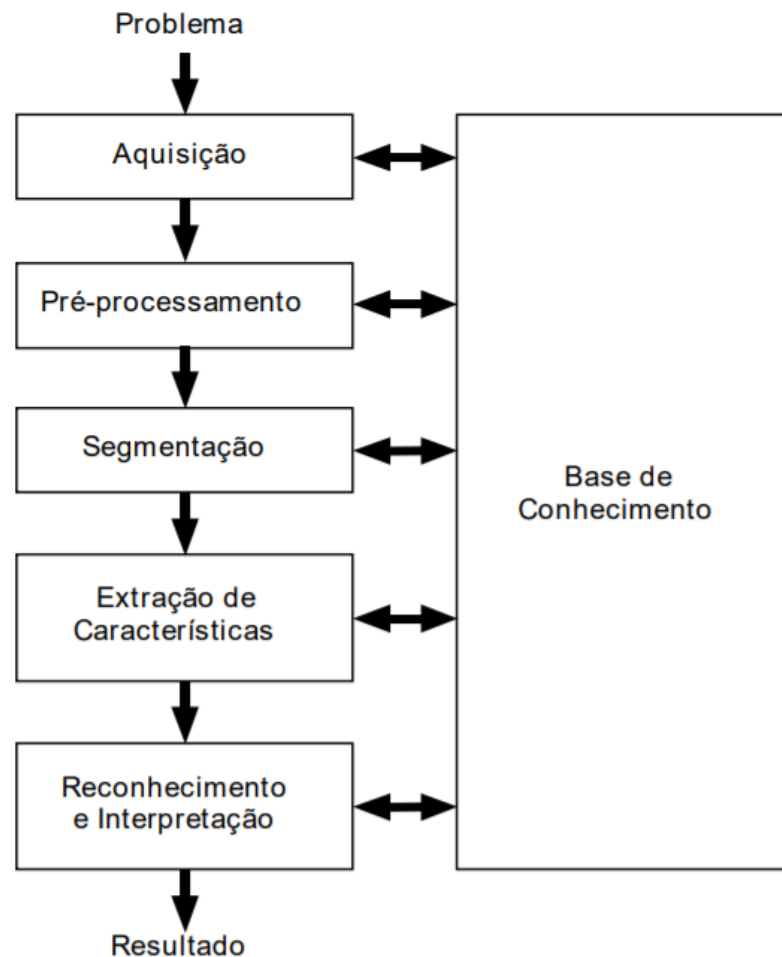


Figura 1 – O diagrama mostra todas as fases de um sistema de visão computacional, desde a definição do problema, até o resultado (FILHO; NETO, 1999).

2.2 Imagens Digitais

De acordo com Klette (2014) uma imagem digital pode ser representada por um vetor retangular de píxeis (x, y, μ) em que x e y representam as coordenadas da localização de um píxel e μ é o seu valor. Essa imagem pode ser formalmente representada pelo seguinte conjunto:

$$\Omega = \{(x, y) : 1 \leq x \leq N_{cols} \wedge 1 \leq y \leq N_{rows}\} \subset \mathbb{Z}^2$$

Ainda segundo Klette (2014), uma imagem digital possui orientação baseada num plano cartesiano, composto por eixos x e y . É possível dizer que a imagem está localizada no quarto quadrante de um plano cartesiano padrão, pois seus valores no eixo x (largura) crescem em módulo para a direita e seus valores no eixo y (altura) crescem também em módulo para baixo, partindo do ponto $(0,0)$, localizado na origem do sistema. Uma forma bastante comum para se representar a orientação de uma imagem digital é a mão esquerda, como pode ser visto na Figura 2, em que o polegar representa o eixo x , contendo as N colunas da imagem e o indicador o eixo y , que contém as N linhas.

O valor μ é comumente representado de duas maneiras: por meio de um escalar, para o caso de imagens binárias (preto e branco) e para o caso de imagens em escala de cinza ou com o uso de vetor, para as imagens coloridas.

De acordo com Klette (2014), uma imagem escalar possui valores inteiros para μ dentro do intervalo $\{0, 1, \dots, 2^a - 1\}$, em que 0 representa o preto, $2^a - 1$ representa o branco e qualquer valor intermediário é uma interpolação linear entre esses dois valores, gerando assim diferentes níveis dentro da escala de cinza. Já uma imagem binária possui apenas dois possíveis valores para seus píxeis, geralmente representados com o valor 0 para o preto e o 1 para o branco. Um valor muito utilizado para a é o número 8 e dessa maneira, é possível representar o intervalo fechado entre preto e branco com 8 bits, ou seja, 256 tons de cinza.

Ainda segundo o mesmo autor, imagens baseadas em valores vetoriais, possuem múltiplos

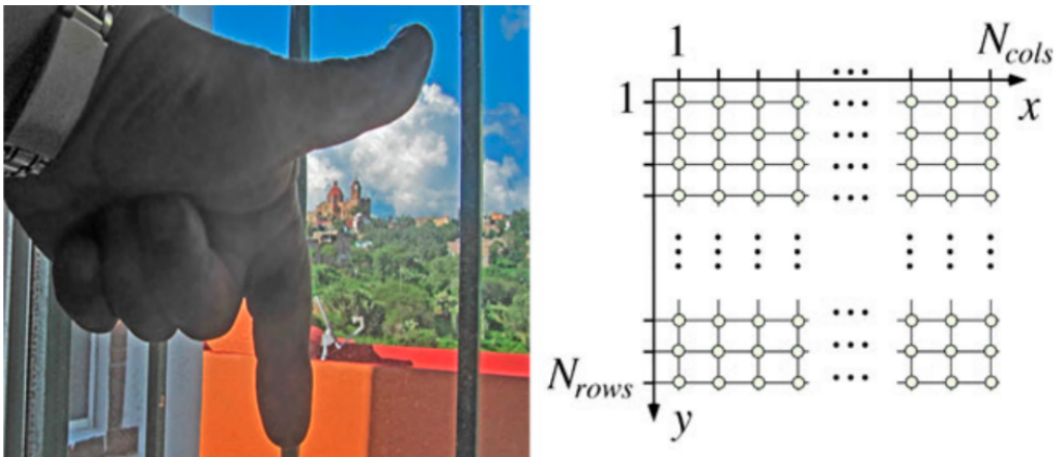


Figura 2 – No sistema de orientação da mão esquerda o polegar representa a coordenada X (colunas) enquanto o indicador representa a coordenada Y (linhas) (KLETTE, 2014).

canais ou bandas para cada valor de pixel μ . O espaço de cor RGB, considerado o espaço de cor mais comum para o ser humano, é constituído de 3 componentes, que representam as intensidades das colorações vermelha, verde e azul de determinada cena. Na [Figura 3](#) é mostrada uma imagem no espaço de cor RGB, juntamente com suas três componentes separadas, representadas em escala de cinza, pois por serem imagens de apenas um canal, cada uma das 3 imagens possui valor μ escalar, que representa a intensidade daquela componente nos píxeis da imagem original. Esta está localizada no canto esquerdo superior e ao seu lado está apresentada a componente vermelha. As componentes verde e azul estão, respectivamente, nos cantos esquerdo inferior e direito inferior. Por representarem cada uma dessas três cores primárias na imagem original, as imagens das componentes apresentam píxeis de alta intensidade em regiões que originalmente possuem forte presença de sua componente. Quanto mais vermelho tiver em determinada localização da imagem original, como na casa, mais claro será o mesmo espaço na imagem que representa a componente vermelha. É fácil observar o mesmo fenômeno para a componente azul, em que o céu, originalmente azul, se apresenta com um tom muito próximo ao branco.

De acordo com [Szeliski \(2010\)](#), as cores primárias aditivas (vermelho, verde e azul) podem ser somadas para gerar outras cores, inclusive o branco, caso as 3 sejam somadas em altas intensidades. Esse comportamento pode ser observado, por exemplo, em um projetor ou em um monitor. Esse processo ocorre devido ao próprio funcionamento do olho humano, que possui três tipos de cones (células capazes de reconhecer cores). Cada um desses cones é sensível a um determinado comprimento de onda, curto, médio e longo, o que fornece a capacidade do ser

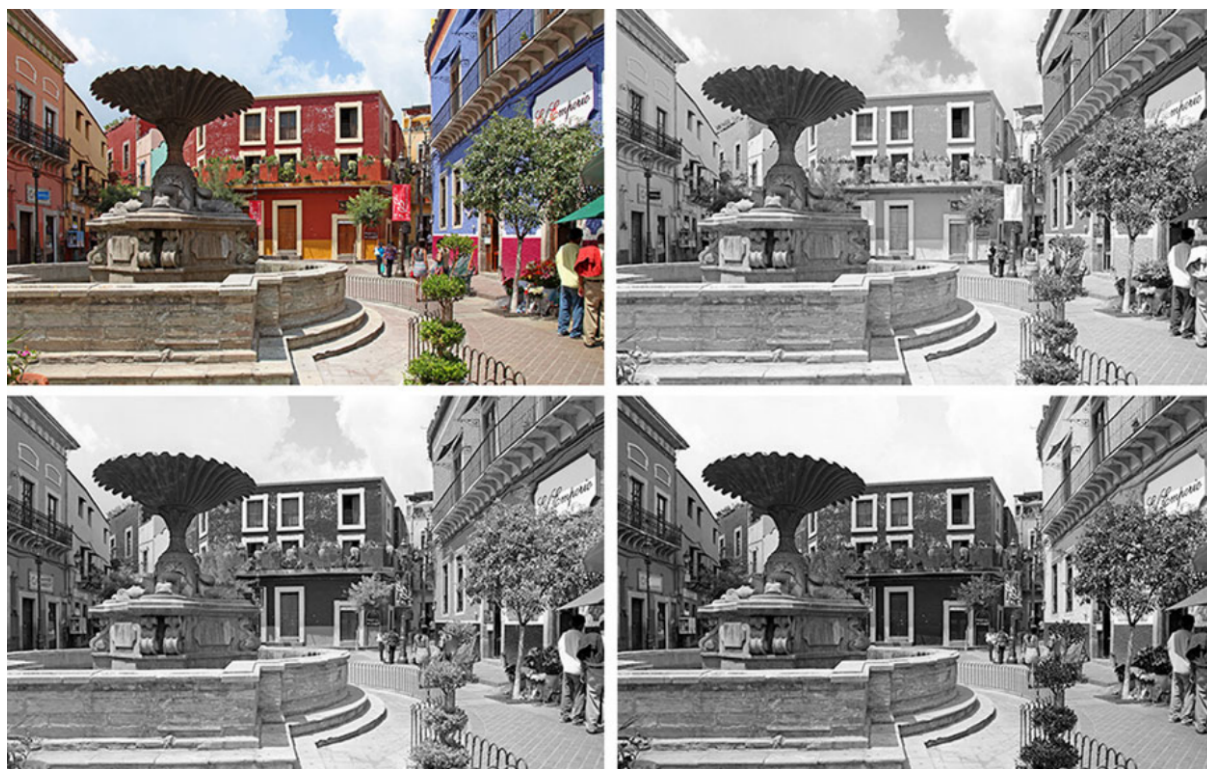


Figura 3 – Da esquerda para a direita e de cima para baixo: imagem original, componente vermelha(R), componente verde(G), componente azul(B) ([KLETTE, 2014](#)).

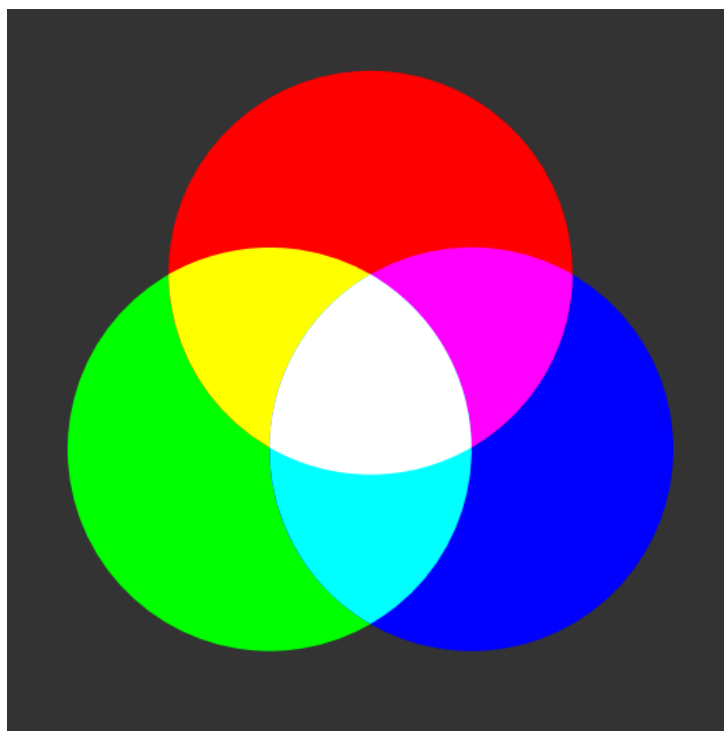


Figura 4 – Cores primárias e suas respectivas somatórias (sobreposições) (KLETTE, 2014).

humano de enxergar azul, verde e vermelho, assim como suas combinações. Essa capacidade é chamada de tricromatismo. A Figura 4 mostra as cores aditivas primárias, assim como suas combinações. O amarelo é o resultado da combinação de vermelho e verde. Já o magenta surge do vermelho e azul, enquanto o ciano é a combinação de azul e verde. O branco é formado por meio da combinação de todas as 3 cores, enquanto o preto é representado pela ausência delas.

2.2.1 Espaço de cor

Segundo Koschan e Abidi (2008), a cor não é uma dimensão física como largura ou temperatura, que pode ser medida quantitativamente em termos absolutos. Ela é, de certa forma, um fenômeno perceptivo, pois apesar de o espectro eletromagnético ser categorizado em diferentes comprimentos de onda, gerando assim um intervalo de cores visível aos olhos humanos, essas cores, ao serem absorvidas, podem ser percebidas de maneiras distintas, devido ao efeito metamérico ou metamerismo. Este fenômeno se define nos casos em que dois elementos analisados sob determinada iluminação e observador são considerados iguais, porém ao serem submetidos à variação de pelo menos uma dessas duas condições, se apresentam diferentes.

Diante da dificuldade de se rotular cores baseando-se somente no espectro eletromagnético, foi necessária a criação de um sistema representativo para as cores, a fim de se armazenar, exibir e processar imagens coloridas. Essa representação deveria ter a capacidade de atender os requisitos de algoritmos de processamento de imagem, assim como as demandas técnicas de monitores e demandas perceptivas da visão humana. Como seria praticamente impossível atender todos os requisitos ao mesmo tempo, foram desenvolvidos diversos sistemas de coordenadas de cor,

chamados de espaços de cor.

2.2.1.1 Espaço RGB

O espaço de cor RGB é um dos modelos mais comuns para representar as cores. Ele é utilizado na maioria dos algoritmos de processamento de imagens, além de ter alta correspondência com monitores, já que a combinação da emissão das luzes vermelha, verde e azul formam a combinação das demais. Ele é um espaço de cor aditivo, formado pelas componentes: vermelha, azul e verde. As demais cores são formadas pela combinação dessas três, inclusive o branco e o preto.

Uma representação bastante comum deste espaço de cor é o cubo RGB, exibido na [Figura 5](#). No cubo da esquerda é possível observar os três eixos do cubo: verde, vermelho e azul (*red, blue, green*). O cubo da figura é representado por valores de 24 bits, 8 bits para cada coordenada, ou seja, as possíveis combinações de cor variam de (0,0,0) até (255,255,255). Como na figura os valores estão normalizados, os valores vão de (0,0,0) até (1,1,1).

É possível observar que a origem do plano (0,0,0) contém o preto, em contrapartida, a outra extremidade (1,1,1) contém o branco. A diagonal do cubo que liga esses dois pontos, ou seja, a reta que contém valores do tipo (n,n,n) representa a escala de cinza.

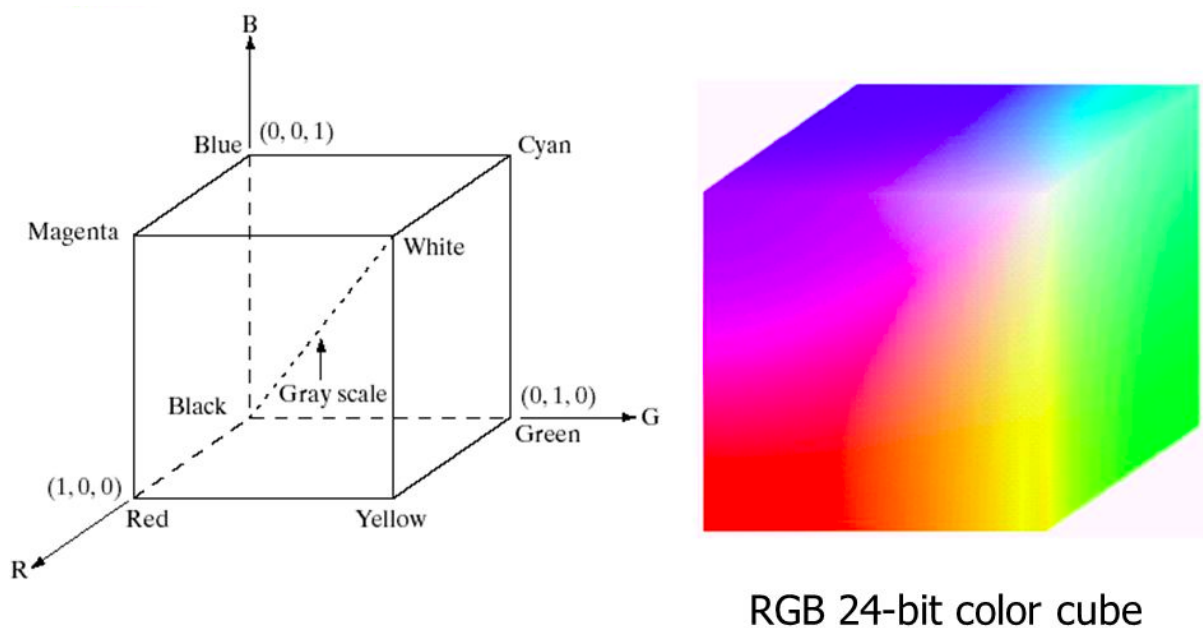


Figura 5 – Cubo RGB representando todas as cores dentro desse sistema, desde o preto (0,0,0) até o branco (1,1,1) ([ALLIE, 2017](#)).

2.2.1.2 Espaço HSV

Segundo [Rosebrock \(2021\)](#), neste espaço, diferentemente do RGB, a "luminosidade" possui sua componente específica, podendo ser isolada. No espaço RGB, para gerar determinada cor mais clara, era necessário incrementar as três componentes.

Cada componente do HSV representa determinada propriedade relacionada às cores. A componente *S(saturation)* indica o quão cinza uma cor é. Caso ela seja máxima, a cor se apresenta pura, totalmente "viva" e caso ela seja mínima, a cor se apresenta totalmente cinza (o quão próximo do branco ou preto esse cinza será, depende da componente *V*). A componente *V(value)* indica o quão luminosa uma cor é, ou seja, valores próximos de zero produzem cores próximas ao preto e valores altos produzem cores mais claras. O *H(hue)* é a componente que diz qual das cores está em análise, independente da sua pureza ou luminosidade.

É interessante observar que, assim como o cubo RGB, o cilindro HSV, exibido na [Figura 6](#), também possui uma reta em escala de cinza. É a reta vertical que passa pela origem do sistema, quando a saturação é zero e o valor varia entre mínimo e máximo. Dessa forma é possível atingir valores de puro preto, até o puro branco, passando pelos tons de cinza entre eles.

O espaço de cor HSV é muito utilizado em aplicações que utilizam visão computacional, pois em relação ao espaço de cor RGB, é mais fácil definir um intervalo para tentar identificar determinada cor utilizando o HSV. Neste trabalho inclusive, o espaço de cor HSV, dentre muitos outros analisados, foi o escolhido para ser a base da segmentação dos núcleos dos leucócitos, visto que esses elementos possuem alta saturação em relação aos componentes das demais células.

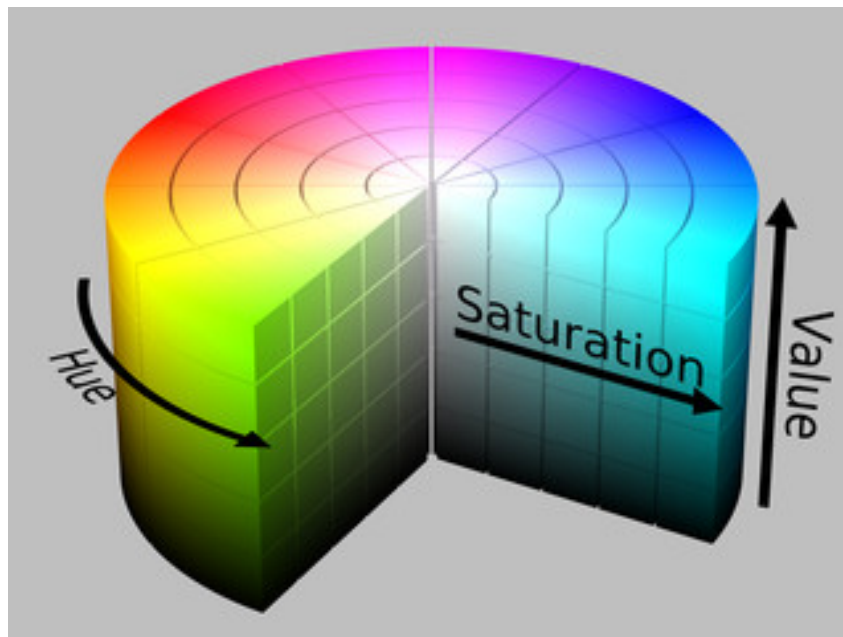


Figura 6 – Representação do espaço HSV (Hue, Saturation, Value) através de um cilindro ([OPENCV](#),).

2.3 Segmentação

Segundo Klette (2014), a segmentação de uma imagem é responsável por dividi-la em regiões, chamadas de segmentos, para futuro processamento, análise ou exibição. É possível encontrar no trabalho de Filho e Neto (1999) uma definição parecida, indicando que o objetivo da segmentação é separar, do restante da imagem, aqueles objetos que se deseja estudar, distinguindo-os dos demais. Neste trabalho, a etapa de segmentação tem o objetivo de separar os leucócitos do restante da imagem (ou cena), para que individualmente possam ser encaminhados para o restante do processo de classificação.

De acordo com Gonzalez, Woods et al. (2002) é possível destacar três principais técnicas para realizar a segmentação de imagens digitais: *Watershed*, crescimento de regiões e limiarização.

Segundo GOMES (2009), métodos baseados em limiar (*threshold*) são as abordagens mais simples para realizar a segmentação de uma imagem digital. Eles levam em consideração a diferença da intensidade entre os píxeis dos grupos pertencentes ao fundo da imagem e aqueles que compõem os elementos de interesse. O histograma da imagem geralmente é muito utilizado neste tipo de método, pois auxilia na definição do limiar de intensidade que separa os dois grupos. Geralmente a saída desse tipo de método é uma imagem binária, em que o fundo recebe valor 0 ou 1 (branco ou preto), dependendo da implementação, e os elementos de interesse recebem o valor oposto.

Já a técnica do *Watershed* consiste em dividir a imagem em diversas regiões, baseando-se em relevo e na ideia de bacias hidrográficas. Basicamente os píxeis de maior intensidade são considerados picos, enquanto píxeis de intensidades mais baixas são tratados como vales, virtualizando assim um relevo. Os vales então são "inundados", do centro para as extremidades, até que duas regiões se encontrem, gerando assim os contornos e consequentemente segmentando esses dois elementos.

As técnicas de crescimento de região, por sua vez, estipulam que determinados píxeis, dentro de regiões de interesse já conhecidas, sejam considerados sementes. A partir de um píxel semente, seus vizinhos são analisados para verificar se fazem parte do mesmo grupo que ele. Em caso positivo, esses vizinhos se tornam sementes para a próxima iteração. Quando não há mais vizinhos considerados sementes, as iterações terminam e o processo de segmentação é finalizado.

Os métodos do *Watershed* e de crescimento de regiões geralmente são mais robustos em relação ao método de limiarização, mas possuem a desvantagem de, na maioria das vezes, necessitarem de uma localização prévia para o posicionamento das sementes (ou das "fontes de água" para o caso do *Watershed*).

Para este trabalho, foram feitos alguns testes com técnicas baseadas em limiarização e técnicas baseadas em *Watershed*, sendo que o primeiro gerou resultados mais acurados. Como a maioria das imagens da base de dados apresenta fundo complexo, preenchido com diversos elementos de não interesse, o *Watershed* se mostrou uma alternativa não muito interessante, pois acabava considerando esses objetos como parte do *foreground*. Além disso, foi observado que não era necessária uma solução complexa como o *Watershed*. A imagem foi então segmentada

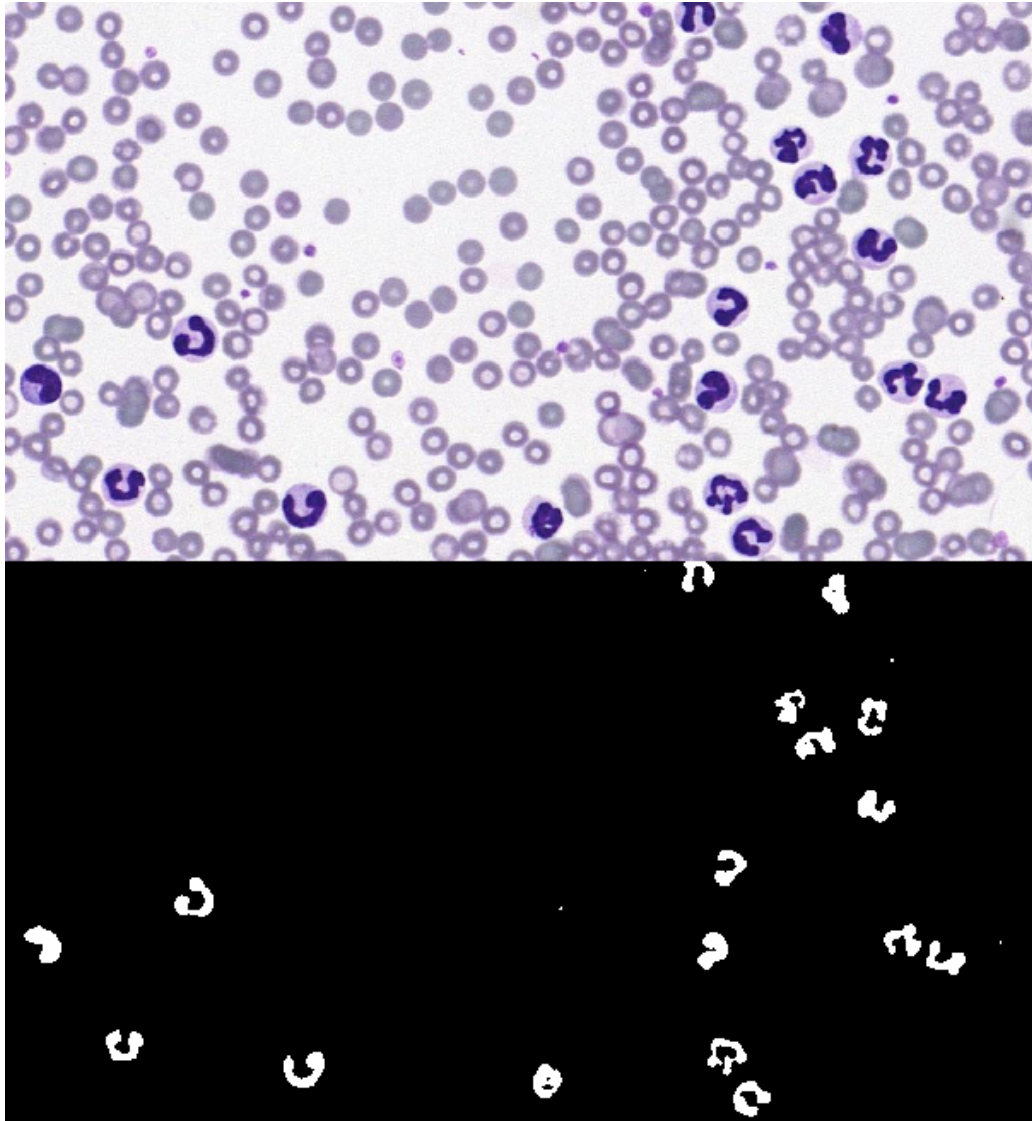


Figura 7 – Representação de uma imagem original e sua segmentação, com foco em separar leucócitos do restante da cena.

utilizando técnicas de limiarização, levando em conta a intensidade dos píxeis dos núcleos dos leucócitos da componente da saturação no espaço de cor HSV, como mostrada na [Figura 7](#).

2.4 Aprendizado de máquina

De acordo com [Chollet \(2018\)](#), inteligência artificial pode ser definida como a automação de processos manuais. Este feito é alcançado por meio da programação de regras, para que assim possa ser obtido um resultado específico.

Pode se dizer que um braço mecânico, programado para pegar um objeto de determinado lugar (em uma linha de produção, por exemplo) e soltar em outro, possui certo nível de inteligência artificial, pois ele segue as regras para as quais foi direcionado. Outro exemplo seria um sistema explicitamente programado para jogar uma partida de jogo da velha contra um ser humano.

Nesse caso, todas as possibilidades de jogadas subsequentes às jogadas do humano já estariam previamente mapeadas, ou seja, após a jogada do humano, o sistema saberia exatamente qual a melhor jogada a fazer.

Ainda conforme o autor, o aprendizado de máquina ou *machine learning*, se apresenta como uma sub área da inteligência artificial e consiste em um tipo de inteligência diferente da que foi citada no parágrafo anterior. Em vez de possuírem regras explicitamente programadas, um sistema de aprendizado de máquinas recebe uma série de estímulos ou dados de entrada, para que um padrão seja criado com eles. Este padrão é representado por um modelo generalista, que ao receber dados de entrada que nunca viu, aplica as regras aprendidas com o conjunto inicial. Isso quer dizer que não se trata mais de um sistema no qual existe o controle explicitamente programado de todas as instruções para cada possível situação. Agora o sistema aprende e aplica automaticamente uma série de regras que ele mesmo definiu, ao aprender com uma série de exemplos.

Um exemplo de aprendizado de máquina é relacionado ao problema da predição de preços de casas ([RUSHITI, 2020](#)). Neste problema se tem determinada base de dados com diversas características de diversas casas em uma região e se quer aplicar um modelo para prever o preço de casas que não se sabe o preço. Este é um exemplo de regressão que não se pode resolver utilizando regras explícitas, pois abstrair todos os parâmetros envolvidos de maneira manual para se encontrar um resultado seria uma prática inviável.

De maneira geral, os problemas resolvidos por meio de aprendizado de máquina são relacionados a reconhecimento de padrão. Grande parte desses problemas são resolvidos por meio de técnicas de aprendizado supervisionado. De acordo com [Bishop e Nasrabadi \(2006\)](#), este tipo de abordagem é caracterizada por utilizar dados de entrada rotulados, ou seja, o sistema é treinado com exemplos nos quais ele sabe a saída esperada. Normalmente este tipo de aprendizado é utilizado para classificações e regressões.

Existe também o aprendizado não supervisionado, o qual não apresenta correlação direta entre dados de entrada e saídas esperadas. Esse tipo de aprendizado é utilizado para resolver outros problemas de reconhecimento de padrões, como o agrupamento (*clustering*), em que determinado conjunto de dados precisa ser dividido em grupos, mas não necessariamente se sabe quantos grupos serão divididos e nem a quantidade de elementos em cada um.

Este trabalho propõe a solução de um problema de classificação, portanto utiliza aprendizado de máquina supervisionado. Cada dado de entrada, no caso uma imagem de um leucócito, é enviado para o sistema de classificação rotulado com sua respectiva classe: neutrófilo, linfócito, eosinófilo, basófilo ou monócito. Dessa maneira o sistema, durante a etapa de treinamento, consegue aprender as características de cada célula e generalizar o conhecimento, para que durante a fase de teste, tenha a capacidade de dizer qual a classe de uma célula branca desconhecida.

2.4.1 Redes neurais - O cérebro humano

Segundo Haykin (2007) o cérebro humano é um computador de alto desempenho, complexo, não linear e paralelo. Por meio da organização e comunicação efetiva dos neurônios, estruturas básicas da constituição desse órgão, o cérebro consegue realizar tarefas, em milissegundos, que um computador digital convencional demoraria dias para executar. Um exemplo seria o reconhecimento de padrões, como encontrar rostos familiares em uma cena mais complexa, tarefa na qual o cérebro gasta em torno de 150 milissegundos para executar.

O aprendizado do cérebro se dá através da experiência, ou seja, com o tempo, ao receber estímulos do ambiente, as conexões entre os neurônios se moldam, a fim de se adaptarem aos estímulos, visando dar uma resposta plausível para aquela situação. Este processo ocorre ao longo de vários anos, mas é crucial nos dois primeiros anos de vida.

Ainda segundo o mesmo autor, o sistema nervoso humano é constituído basicamente por 3 estágios: o primeiro é representado pelos receptores, que captam os estímulos e os transmitem, via impulsos elétricos, para o próximo estágio, a rede neural, o cérebro propriamente dito. Neste segmento a informação recebida é processada e encaminhada para o último estágio, de atuadores, no qual ações são tomadas baseadas nas informações processadas anteriormente. Um esquema deste sistema pode ser visto na Figura 8, em que as setas direcionadas para a direita representam transmissão do sinal, enquanto as voltadas para a esquerda indicam a presença de realimentação no sistema.

É interessante analisar que a velocidade de comunicação entre os neurônios humanos é cerca de 6 ordens de grandeza mais lenta do que a comunicação das portas lógicas em silício, traçando um paralelo entre o processamento do cérebro humano e de um computador digital. Para compensar esta velocidade, o cérebro humano conta com a presença de aproximadamente 10 bilhões de neurônios e aproximadamente 60 trilhões de sinapses ou conexões entre eles.

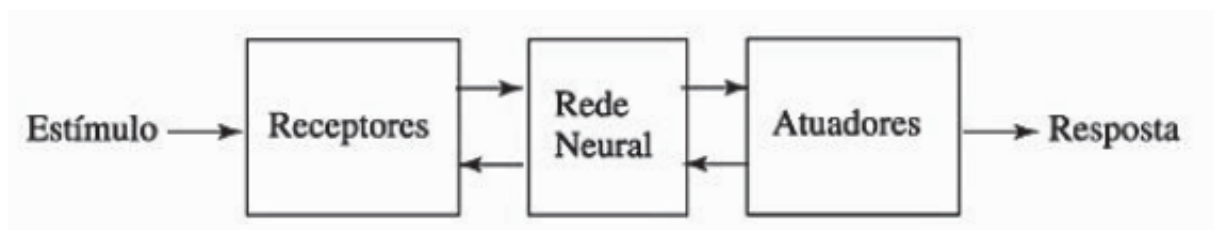


Figura 8 – Representação simplificada dos estágios do sistema nervoso humano, desde o estímulo externo, até a resposta executada (HAYKIN, 2007).

2.4.2 Redes neurais artificiais

Segundo [Silva et al. \(2017\)](#), redes neurais artificiais são estruturas formadas a partir da observação do sistema nervoso humano, principalmente o cérebro. As peças fundamentais desse tipo de estrutura são neurônios artificiais, que por sua vez foram baseados em uma versão simplificada de um neurônio humano, levando em consideração, principalmente, o modo que este gera e propaga impulsos elétricos. No trabalho de [Haykin \(2007\)](#) é dito que uma rede neural é como um processador, que faz uso de inúmeras estruturas fundamentais de processamento simples (neurônios artificiais) para armazenar conhecimento e disponibilizá-lo. Para que uma rede possa aprender então é necessário que os pesos sinápticos, as forças de conexão entre os neurônios, sejam modificados de maneira a atender determinado objetivo.

A unidade mais básica de processamento de uma rede neural é o neurônio, e um exemplo deste tipo de estrutura pode ser vista na [Figura 9](#). O conjunto $\{x_1, x_2, \dots, x_n\}$ representa a entrada, ou estímulo do sistema. O peso sináptico é representado pelo conjunto $\{w_1, w_2, \dots, w_n\}$, que por sua vez tem cada um de seus elementos multiplicado por um elemento do conjunto de entrada, cada um de acordo com sua respectiva posição. A soma ponderada das entradas passa então pelo elemento θ , conhecido como *bias*. Este elemento é responsável por definir um limiar para a ativação ou não ativação do neurônio, dependendo do valor que chegue a ele. Em seguida, representada por $g(\cdot)$, está presente a função de ativação, com a função de, ao receber o resultado u , limitar para um intervalo de valores coerente. Por fim, é possível observar a saída representada por y , para o neurônio em questão, dado determinado conjunto de entrada $\{x_1, x_2, \dots, x_n\}$, que pode ser utilizada inclusive como fonte de dado de entrada para outros neurônios subsequentes.

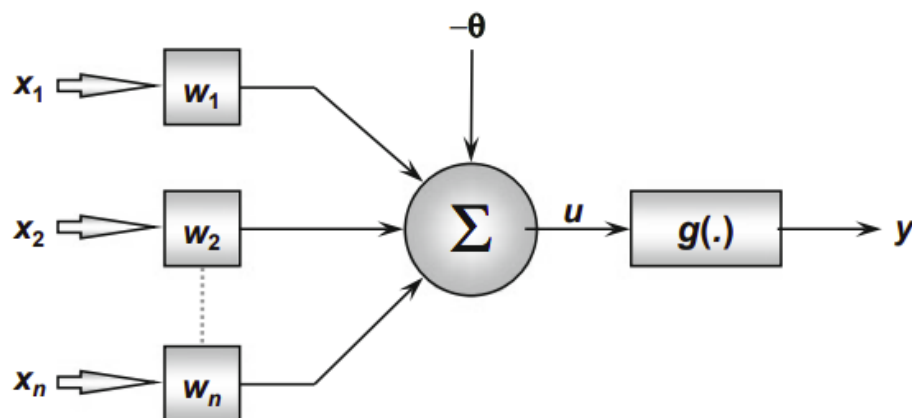


Figura 9 – Representação de um neurônio artificial, exibindo as entradas, pesos, processamento, função de ativação e saída final ([SILVA et al., 2017](#)).

2.4.2.1 Perceptron

Segundo [Géron \(2019\)](#), o perceptron, criado em 1957 por Frank Rosenblatt, é um dos tipos de arquitetura de redes neurais artificiais mais simples. É baseado em um tipo de neurônio chamado de *threshold logic unit* (TLU), que se diferencia por pouco de um neurônio artificial comum. Cada entrada está associada a um peso e sua função de ativação é a *Heaviside*, que possui a seguinte configuração:

$$heaviside(z) = \begin{cases} 0 & \text{if } z < 0, \\ 1 & \text{if } z \geq 0. \end{cases}$$

A configuração da rede TLU permite que ela seja utilizada para classificações lineares binárias simples, pois dependendo da combinação linear das entradas, aquele caso de teste se enquadra em uma classe 0 ou 1. É interessante ressaltar que o Perceptron é composto de uma única camada de TLUs, que se conecta a todas as entradas e quando todos os neurônios de determinada camada estão conectados a todos os neurônios da camada anterior é comum dizer que aquela camada é totalmente conectada (*fully connected layer*) ou *dense layer*.

2.4.2.2 Perceptron Multi Camadas

O Perceptron Multi Camadas ou *Multi Layer Perceptron* (MLP) é uma das redes neurais artificiais mais utilizadas, composta por uma camada de neurônios de entrada, uma ou mais camadas de TLUs, chamadas de camadas ocultas e uma camada de saída. Todas as camadas com exceção da camada de saída possuem um neurônio tipo *bias* e são totalmente conectadas. ([BELAN et al., 2019](#)) ([GÉRON, 2019](#))

A [Figura 10](#) mostra um exemplo de Perceptron Multi Camadas.

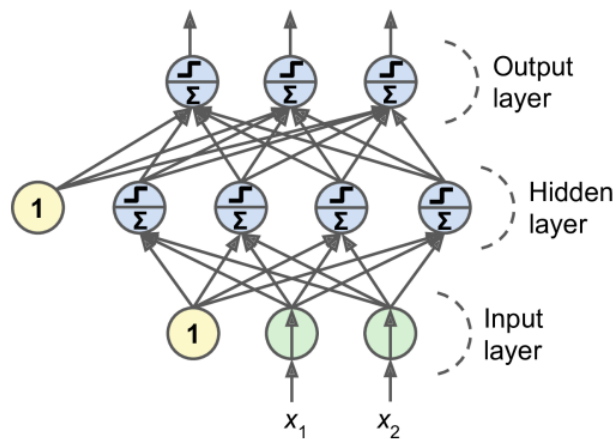


Figura 10 – Representação de um Perceptron Multi camadas. Podem ser observadas as camadas de entrada, ocultas e de saída ([GÉRON, 2019](#)).

2.4.2.3 Backpropagation

O algoritmo de *backpropagation* foi proposto em 1986 por David Rumelhart, após vários anos de dificuldade da parte dos pesquisadores em encontrar um bom modo de treinar uma MLP. Esta técnica tem como base o gradiente descendente e consegue calcular o gradiente de maneira automática em apenas duas etapas, passando uma vez pela rede para frente e outra para trás. Desse modo, o algoritmo consegue definir quais os pesos e *bias* ideais para que o erro seja minimizado, pois a cada etapa de treinamento, ou época, o valor do erro é corrigido, através da função de perda (*loss function*), que diz o quão longe o resultado atual está do esperado. Este processo então é repetido até que a rede possa convergir para uma solução ideal.

2.4.3 Redes Neurais Profundas

De acordo com [Chollet \(2018\)](#), aprendizado profundo (*deep learning*) é uma subárea do aprendizado de máquina (*machine learning*). Estruturas deste tipo podem ser compostas de dezenas ou até mesmo centenas de camadas e aprendem automaticamente ao ter contato com os dados de entrada.

Já [Albawi, Mohammed e Al-Zawi \(2017\)](#) explicitam que o termo aprendizado profundo ou rede neural profunda *Deep Neural Network* se referem a redes neurais artificiais com muitas camadas e esse tipo de técnica se tornou, nos últimos anos, uma das ferramentas mais poderosas e populares no universo do aprendizado de máquinas, já que é capaz de lidar com uma quantidade gigantesca de dados e possui a grande habilidade em reconhecer padrões.

2.4.3.1 Redes Neurais Convolucionais

Ainda segundo [Albawi, Mohammed e Al-Zawi \(2017\)](#) um dos tipos de redes neurais profundas mais conhecido é a rede neural convolucional (CNN). Este tipo de arquitetura possui inúmeras camadas, dentre elas, camadas convolucionais, camadas de *pooling* e camadas totalmente conectadas. As redes neurais convolucionais possuem alto desempenho quando se trata de aplicações que lidam com abundância de dados, como em problemas de classificação, de visão computacional e processamento de linguagem natural. Com base nesses fatos, o uso de CNNs é uma escolha interessante para resolver o problema proposto neste trabalho: classificar automaticamente leucócitos em cães utilizando técnicas de visão computacional e aprendizado de máquina.

De acordo com [Belan et al. \(2019\)](#) as redes neurais convolucionais ganharam muito espaço no reconhecimento de padrão em imagens digitais principalmente por sua habilidade de aprender filtros a serem empregados nas imagens sem a necessidade de explicitá-los. Essa característica faz com que não seja necessário ter etapas muito complexas de pré-processamento ou de separação de características a serem analisadas, o que é eficiente para a simplicidade de implementação, porém possui alto custo computacional.

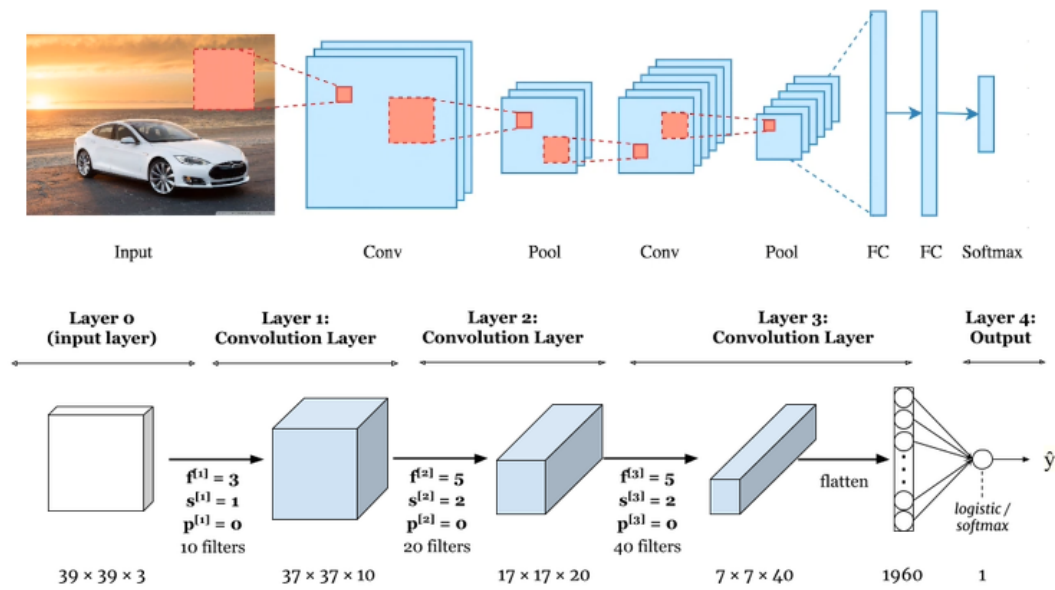


Figura 11 – Exemplo de arquitetura de uma CNN. É possível observar a entrada (figura do carro), camadas de convolução, de pooling e camadas totalmente conectadas, além da ativação, para saída final (CHAUHAN, 2021)

O mesmo autor explica que as camadas de convolução são as responsáveis por extrair automaticamente as características das imagens de entrada da rede, como contornos e bordas. Já a camada de *pooling*, também chamada de camada de agrupamento, tem como principal objetivo a redução da dimensão do objeto analisado, para que assim seja possível diminuir um pouco a complexidade do processamento, consequentemente o elevado custo computacional. Apesar de reduzir a dimensão da imagem, essas camadas não causam perda de informação. Em seguida e por fim, uma operação de *flatten* ou achatamento é realizada nos dados, para alimentar as camadas totalmente conectadas, vetores unidimensionais que possuem o papel de realizar a classificação, por meio das informações processadas pelas camadas anteriores. Por fim, por meio de determinada função de ativação, como softmax ou ReLU, a camada final é gerada, constituída das probabilidades da imagem de entrada pertencer a cada classe.

A Figura 11 exemplifica a arquitetura de uma CNN, exemplificando o que acontece nas camadas de convolução, de *pooling* e nas totalmente conectadas.

2.4.3.2 Transferência de aprendizado

A técnica de transferência de aprendizado foi desenvolvida a fim de se auxiliar na solução do problema intrínseco às redes neurais profundas: a dependência de grandes conjuntos de dados. De acordo com Tan et al. (2018) esse tipo de rede precisa de uma quantidade de dados para treinamento muito maior do que métodos clássicos de aprendizado de máquina e adquirir essa base de dados geralmente não é uma tarefa fácil, algumas vezes até inalcançável.

Conforme o trabalho de [Zhuang et al. \(2020\)](#), a transferência de aprendizado pode ter sua origem na psicologia educacional, se apoiando nas ideias de C.H. Judd, que dizem que o ato de transferir aprendizado é, na verdade, o resultado da generalização das experiências. De acordo com essa teoria, o pré-requisito para ocorrer a transferência de aprendizado é existir correspondência entre as duas tarefas. Um exemplo prático seria a maior facilidade que determinada pessoa que já toca violino teria em aprender a tocar piano, pois apesar de serem bem diferentes, os dois estão correlacionados por serem ambos instrumentos musicais, ou seja, provavelmente possuem pontos em comum a serem aprendidos.

Na transferência de aprendizado para modelos de aprendizado profundo, os pesos de um modelo previamente treinado para resolver outro problema, são utilizados para auxiliar no treinamento do modelo atual. Para isso as últimas camadas do modelo base são retiradas, para que o novo modelo, que já aprendeu características gerais, como reconhecer bordas, formas e texturas (no caso de CNNs) com as primeiras camadas, possa agora executar um aprendizado mais especializado, buscando identificar características mais refinadas das amostras da nova base de imagens apresentada a ele. Essas características aceleram o processo de treinamento, pois não há a necessidade de se executar essa etapa desde o início, então menos épocas são requeridas, além de auxiliarem em treinos que possuem base de imagens reduzida, já que o modelo anterior carrega certo nível de conhecimento, por já ter sido treinado anteriormente.

A [Figura 12](#) exemplifica o processo de transferência de aprendizado. Na figura o modelo base foi treinado utilizando a base de dados imageNet, apresentada no trabalho de [Deng et al. \(2009\)](#).

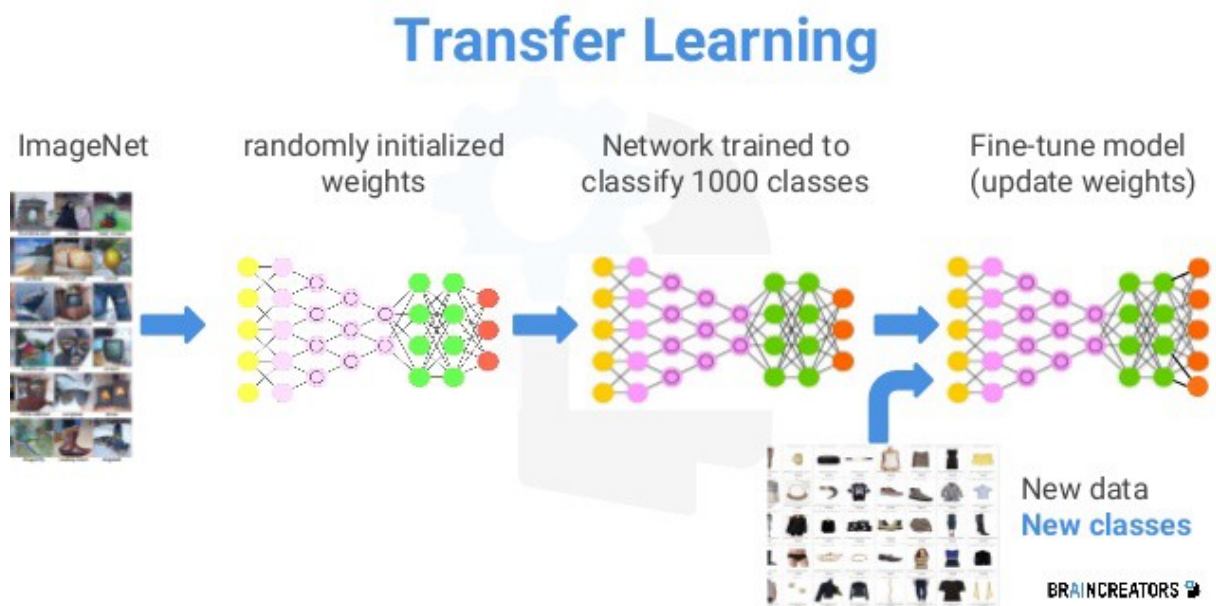


Figura 12 – Representação da técnica de transferência de aprendizado, utilizando uma rede pré-treinada para aprender sobre "outro assunto" ([RAMIAH, 2019](#)).

2.5 Conclusão

Os conceitos e técnicas apresentados neste capítulo constroem a base para a classificação automática de leucócitos em imagens sanguíneas de cães. Os espaços de cor explicitados, como HSV e YUV permitem, por meio de técnicas de limiarização, evidenciar os núcleos das células de interesse, separando esses elementos do fundo da imagem. Essa separação é fundamental para a etapa de segmentação, com o papel de garantir que os leucócitos serão devidamente localizados nas imagens e encaminhados para o modelo classificador. Este, por sua vez, ao ser construído com base em redes neurais convolucionais, dispensa a complexidade de se extrair características específicas de cada classe manualmente, graças às camadas de convolução. Por fim, técnicas de transferência de aprendizado são eficientes para mitigar o impacto de poucas imagens por classe na base de dados, utilizando conhecimento de outras redes neurais já treinadas para auxiliar na resolução do problema da classificação de leucócitos e todo o apanhado de técnicas escolhidas para compor o método de classificação deste trabalho pode ser visto em mais detalhes na [Capítulo 4](#).

3 Trabalhos Relacionados

Este capítulo é destinado a explicitar os principais pontos de trabalhos presentes na literatura, que possuam propostas similares a deste trabalho. É a apresentação do estado da arte da detecção e classificação automática de leucócitos em imagens de amostras de sangue.

Apesar deste trabalho ter seu foco em cães, a maioria dos trabalhos encontrados são voltados para a espécie humana, logo, alguns dos trabalhos exibidos nesta seção não serão especificamente sobre células caninas, entretanto, todos possuem a mesma proposta: detectar e/ou classificar células sanguíneas (primariamente leucócitos) automaticamente, utilizando técnicas de visão computacional e aprendizado de máquina. Ao se analisar os principais trabalhos encontrados, é possível dividi-los em três grandes grupos: aqueles focados em segmentação, os que priorizam classificação e aqueles que se propõem a tratar ambos os casos.

O trabalho mais relevante focado na área da segmentação foi o de [Cao, Liu e Song \(2018\)](#). O algoritmo desenvolvido foi dividido basicamente em 3 etapas: segmentação do núcleo, localização do leucócito e segmentação do leucócito. Para a segmentação do núcleo, a imagem foi transformada em HSG, pois *Hue*, *Saturation* e *Green* eram as componentes, que, dentro dos dois espaços de cores utilizados (RGB e HSI), realçavam da melhor maneira o núcleo dos leucócitos, em relação ao restante da cena. Após a transformação da imagem, um algoritmo de segmentação baseado em limiares é aplicado. O SWAM (*Stepwise averaging method*) baseia-se na intensidade média da imagem. Caso um determinado píxel tenha intensidade inferior à média, este é mantido, caso contrário, ele é alterado para zero. A média é então atualizada e ocorrem outras iterações, utilizando a mesma lógica. Para a localização das células, foram criados algoritmos que geraram ROIs (*Regions of Interest*) em volta de cada citoplasma, que são basicamente retângulos contendo todo o leucócito. O processo de segmentação das células se deu com a utilização de técnicas baseadas em *fuzzy sets* e *fuzzy divergence*, além de algoritmos utilizados no pós-segmentação, para a correção de eventuais defeitos no processo anterior, como a segmentação unitária de um ou mais leucócitos aglutinados. Para essa correção foram usados algoritmos baseados em concavidade, alcançando uma acurácia média de 93,75%.

Já na área da classificação, um dos trabalhos que mais se destacaram foi o de [Wang e Cao \(2020\)](#). Seu foco principal foi geração de um método de classificação de leucócitos (Eosinófilos, Neutrófilos, Basófilos, Linfócitos e Monócitos) utilizando CNN e, para aprimorar a classificação, um método de aumento de dados (*data augmentation*) foi utilizado, aumentando assim a base de imagens de maneira artificial, potencializando os resultados do treinamento. Inicialmente uma arquitetura de rede CNN foi construída e a base de imagens dividida em treino, validação e teste. O conjunto de teste foi reservado até que os valores de acurácia de treino e validação se tornassem satisfatórios para a avaliação. A rede foi treinada diversas vezes e sempre que os resultados se mostravam abaixo do esperado, duas possíveis ações eram tomadas: os parâmetros da rede eram otimizados, em busca de valores ótimos ou os conjuntos de treino e validação eram

alimentados com mais imagens, através das técnicas de *data augmentation*. Após o modelo ótimo ter sido encontrado, o conjunto de teste foi utilizado para avaliar sua acurácia. O processo de *data augmentation* é executado por meio da transposição de leucócitos para diferentes fundos de imagens (*backgrounds*). Algumas células de determinadas imagens são segmentadas e recortadas de seu fundo original, em seguida são coladas em outros, gerando assim novas imagens de maneira artificial. Inicialmente haviam 9300 imagens disponíveis no conjunto de treino, cada uma contendo um leucócito e seu respectivo fundo. Cada leucócito então foi copiado para outros dois fundos, gerando um total de 27900 imagens. A acurácia final do modelo alcançou 97.6%.

No trabalho de [Lin e Wang \(2018\)](#), o segundo focado em classificação, foi desenvolvido um método para reconhecer cinco tipos de leucócitos: eosinófilos, neutrófilos, basófilos, linfócitos e monócitos. O algoritmo possui três etapas principais: extração dos leucócitos, extração de características e classificação. Para a etapa de extração dos leucócitos (segmentação), foram utilizadas técnicas de transformação de distância (ou mapa de distâncias) e *GVF Snake*. Após a análise de diversos espaços de cor, foi observado que houve certa facilidade para localizar os leucócitos ao se observar a componente H do espaço de cor HSV, motivo pelo qual este foi o selecionado para compor o método. Nesta etapa ainda foi utilizado o algoritmo K-means para corrigir pequenas imperfeições na imagem e refinar a segmentação. A etapa de extração de características foi caracterizada pela aplicação de um algoritmo não paramétrico, *Mean-shift clustering*, para a composição da textura das células, característica principal utilizada para a classificação. Por fim, um classificador foi construído utilizando SVM, analisando os vetores de características, resultando em uma acurácia de 97,96%. É importante citar que uma das considerações feitas pelos autores é de que a utilização de classificação por CNN, apesar de ser utilizada com certa frequência, possui o limitante da falta de dados para treinamento. Não é trivial de se encontrar ou de se produzir uma grande quantidade de imagens de amostras sanguíneas contendo as classes de leucócitos desejadas e este é um dos principais motivos que levaram à não escolha desta técnica para este caso.

O trabalho de [Reena e Ameer \(2020\)](#) realizou experimentos tanto na área da segmentação quanto da classificação. Eles desenvolveram um método de reconhecimento e classificação de cinco classes de leucócitos (Eosinófilos, Neutrófilos, Basófilos, Linfócitos e Monócitos) utilizando aprendizado profundo tanto para a segmentação quanto para a classificação. A etapa de segmentação consiste na utilização de técnicas de segmentação semântica, em que cada píxel da imagem é categorizado. No caso do trabalho em questão, foram utilizadas duas categorias: leucócitos e fundo. Após a localização das células utilizando a técnica citada acima, são gerados recortes em volta das células encontradas e estes são encaminhados para a etapa de classificação. Nesta etapa a CNN AlexNet foi selecionada como o modelo pré-treinado a ser utilizado como base para a técnica de transferência de aprendizado, aplicada ao problema. A etapa de segmentação obteve acurácia de 98,22%, enquanto a de classificação alcançou os 98,87%.

A [Tabela 1](#) reúne os trabalhos relacionados apresentados neste trabalho, indicando o foco, as técnicas utilizadas e o resultado de cada um.

Tabela 1 – Resumo dos trabalhos relacionados

	Foco do trabalho	Técnicas utilizadas	Resultado (Acurácia)
Cao, Liu e Song (2018)	Segmentação	HSG, SWAN e <i>fuzzy</i>	93,75%
Wang e Cao (2020)	Classificação	CNN e <i>data augmentation</i> por transposição de imagem	97,60%
Lin e Wang (2018)	Classificação	Classificação por SVM	97,96%
Reena e Ameer (2020)	Segmentação e Classificação	Segmentação semântica e classificação por CNN com transferência de aprendizado	98,22% (segmentação) e 98,87% (classificação)

3.1 Conclusão

Com base nos trabalhos relacionados citados neste capítulo, foi definido que a melhor abordagem a ser seguida seria utilizar técnicas clássicas de visão computacional para detectar e segmentar os leucócitos (primariamente seus núcleos) e aprendizado profundo utilizando CNN para realizar a classificação dos leucócitos encontrados.

O fator crucial para se excluir técnicas de segmentação semântica foi a quantidade reduzida de amostras na base de dados. Além da dificuldade de se encontrar esse tipo de dado disponível publicamente, para criar uma base nova, é necessário reunir materiais, técnicas e profissionais específicos, todos voltados para a área veterinária, o que requer gasto significativo de tempo.

Apesar das CNNs também requererem quantidade relativamente elevada de exemplos para o treino do modelo, principalmente se comparado a técnicas como o SVM, pela análise da base de dados reunida, foi possível observar que existe distinção suficiente entre os tipos de leucócitos analisados, o que auxilia o modelo a convergir corretamente. É importante ressaltar ainda que as CNNs auxiliam no processo de treinamento, abstraindo algumas etapas da classificação, pois conseguem aplicar filtros e aprender características dos dados analisados de maneira automática.

4 Metodologia

Este capítulo é destinado a explicitar todo o conjunto de métodos e técnicas utilizadas ao longo deste trabalho a fim de se solucionar o problema proposto: realizar a classificação automática de leucócitos de cães por meio do uso de visão computacional e aprendizado de máquinas.

No problema de classificação de leucócitos, o objetivo é identificar e classificar, de maneira automática, essas células em uma imagem cuja cena é uma amostra sanguínea de cães. Para se ter um sistema capaz de realizar o que foi proposto, é necessário que se tenha pelo menos um modelo de rede neural treinado e uma imagem de entrada (leucócito), que ao ser analisada por ele, irá gerar como resposta a informação de qual classe pertence aquela célula. O modelo classificador, descrito na [Seção 4.2](#) foi treinado com imagens individuais de leucócitos, logo, ao testá-lo, somente imagens individuais de leucócitos devem ser passadas a ele. Para garantir tal comportamento, é necessário que a imagem de entrada, semelhante à [Figura 13](#), seja devidamente tratada, para que os leucócitos presentes nela sejam encontrados e separados automaticamente, para poderem ser passados para o classificador. Esse processo é tratado na [Seção 4.1](#).

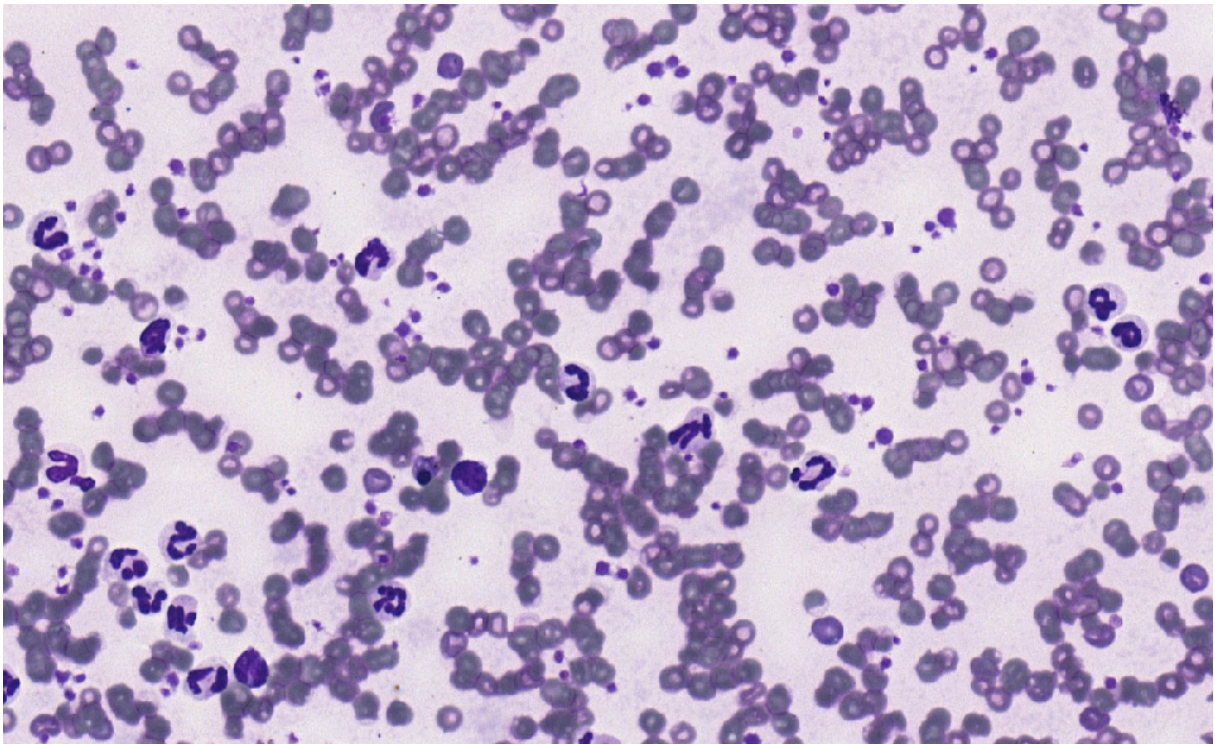


Figura 13 – Imagem de uma amostra sanguínea de cão.

4.1 Segmentação

A etapa de segmentação é de grande importância para o sistema de classificação, pois é por meio dela que os leucócitos serão localizados nas imagens de amostras sanguíneas e encaminhados para o modelo classificador.

4.1.1 Espaços de cor

Para executar o começo da etapa de segmentação, foram feitas análises com quatro dos principais espaços de cor existentes: RGB, YUV, HSV e LAB. Ao perceber que a componente S do espaço HSV proporcionava realce significativo para os núcleos dos leucócitos em relação ao restante da cena, ela foi escolhida para dar prosseguimento ao restante dos experimentos.

Além da componente S do espaço HSV, também foi utilizada a componente V do espaço YUV, com o propósito de remover restos de células, tratados por ruídos, pois não representam informações significativas para a segmentação e classificação.

A [Figura 14](#) mostra as intensidades das componentes dos quatro espaços de cor citados no parágrafo anterior. A figura mostra as componentes referentes à [Figura 13](#) e cada linha representa um espaço de cor. Da esquerda para a direita, de cima para baixo é possível observar na seguinte ordem as componentes R-G-B, Y-U-V, H-S-V e L-A-B.

Para gerar a máscara binária das imagens foi utilizada a função *threshold*, da biblioteca OpenCV. Esta função recebe como argumentos principais uma imagem de entrada, um limiar para classificação dos píxeis e um valor de atribuição aos valores que ultrapassarem o limiar. Dessa forma, todos os píxeis que apresentam valor acima do limiar, são convertidos em píxeis brancos, ressaltando assim, objetos de interesse, em contraponto com o fundo preto.

A função *threshold* foi alimentada inicialmente com a componente S das imagens (no espaço HSV), a fim de identificar aqueles elementos que possuem alto valor de saturação. Foi também utilizado o componente V das imagens (no espaço YUV) para ressaltar os restos celulares indesejados e assim remover esses pixels da imagem final.

4.1.2 Contornos dos núcleos

Como o núcleo dos leucócitos se apresenta com saturação (S) bem superior aos demais componentes encontrados na amostra sanguínea, foi possível identificar a localização destas células através da função *findContours*, também da biblioteca OpenCV, pois ela consegue encontrar pixels brancos conectados, que formam um elemento.

Apesar de ser fundamental para encontrar esses elementos, o resultado da função não é suficiente para dizer ao certo o que é um elemento válido (leucócito) ou não, portanto é necessária uma fase de refinamento.

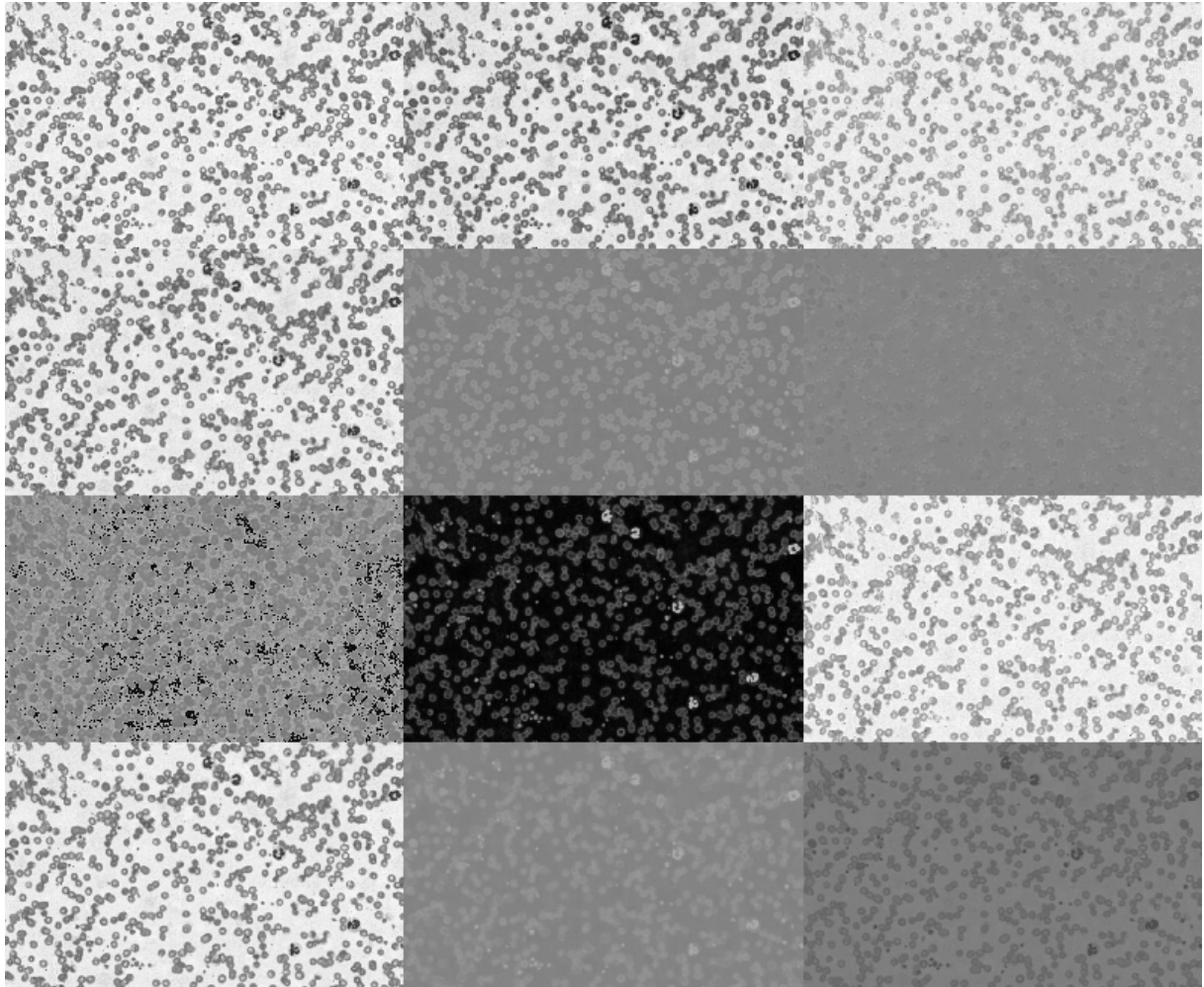


Figura 14 – Componentes de espaços de cor. Da esquerda para a direita e de cima para baixo: R-G-B, Y-U-V, H-S-V e L-A-B

4.1.3 Refinamento

Como explanado anteriormente, ao se utilizar a função *findContours* os contornos dos núcleos dos leucócitos são encontrados, mas é muito comum a permanência de ruído, pois células vermelhas e plaquetas podem apresentar pontos de saturação altos, o que contribui para o aumento dos erros, pois falsos contornos podem ser encontrados na imagem.

A fim de combater este problema, foram feitos diversos experimentos, como utilização de filtros para suavizar a imagem e técnicas de erosão, mas o resultado não foi satisfatório. Os núcleos de alguns leucócitos, principalmente neutrófilos segmentados, são formados por múltiplos elementos, geralmente conectados e, ao se utilizar as técnicas referidas anteriormente, ocorria a separação desses elementos, causando erros na segmentação, pois um único núcleo era, por muitas vezes, dividido de maneira inconsistente.

A solução final para esta situação foi fixar um valor de área em pixels para contornos válidos. Dessa maneira, todo contorno menor que esse limiar é removido da cena. Para determinar a área de cada contorno, foi utilizada a função *moments* do OpenCV, extraindo a propriedade

m00 de seu retorno.

Por fim, com os ruídos devidamente removidos da imagem, foi possível utilizar técnicas de dilatação na imagem, a fim de garantir que os elementos de um único núcleo não se separassem, dado que a saturação na junção desses elementos é naturalmente mais fraca. Foi utilizada a função *dilate* do OpenCV para a execução dessa ideia.

Na [Figura 15](#) é possível ver o resultado da segmentação sem o devido refinamento, já na [Figura 16](#) é possível ver a mesma imagem após a aplicação desse processo.

É possível observar alguns pontos de ruído na imagem sem o devido refinamento, além da fraca ligação entre as partes dos núcleos dos leucócitos realçados com os círculos na imagem. Caso não houvesse o processo de refinamento, os ruídos teriam sido considerados contornos válidos e os leucócitos citados anteriormente teriam sido divididos, cada um em dois.

4.1.4 Localização

Encontrar o núcleo da célula é crucial para determinar sua localização, pois este geralmente representa grande parte de sua estrutura, além de estar geralmente localizado ao centro da célula. Este fato tem grande importância no processo, pois a partir do centro do núcleo é gerado um recorte em volta de toda a célula.

Estes recortes podem ser vistos na imagem [Figura 17](#), destacados pelos retângulos verdes. Cada um deles representa o tipo de imagem enviada para o classificador.



Figura 15 – Segmentação de imagem de amostra sanguínea sem refinamento

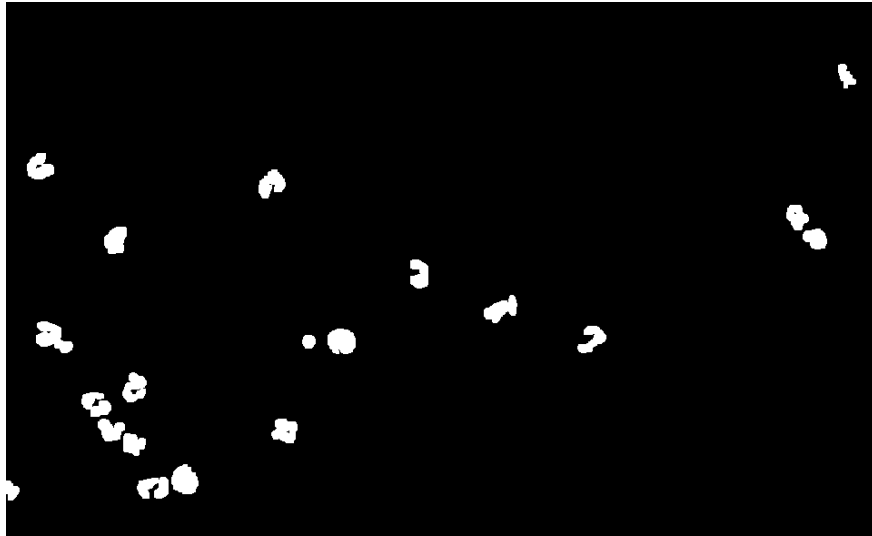


Figura 16 – Segmentação de imagem de amostra sanguínea com refinamento

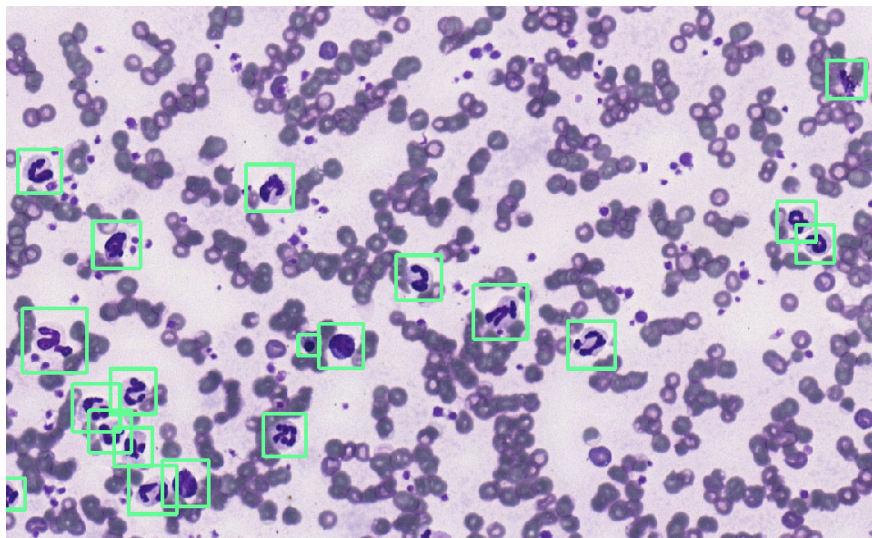


Figura 17 – Recortes nas células encontradas no processo de segmentação. Esses recortes serão as imagens de entrada do classificador.

4.2 Classificador

A etapa final do processo de classificação automática de leucócitos se dá pela utilização de um modelo classificador, que ao receber uma imagem dessa célula, irá dizer a qual classe ela pertence.

É muito comum de se ver em estudos voltados para a classificação de leucócitos, mesmo de células humanas, a utilização de cinco classes de células: linfócitos, neutrófilos, monócitos, basófilos e eosinófilos. Para este trabalho, não foi possível capturar nenhuma amostra de basófilos, por serem bem mais raros de se encontrar. Adicionalmente, foi incluída uma nova classe ao processo de classificação, os metarrubríctos. Essas células são basicamente eritrócitos imaturos e também auxiliam no diagnóstico de enfermidades.

Para o treinamento, foram selecionadas aleatoriamente, 70 imagens aleatórias de cada uma das classes (linfócitos, neutrófilos, monócitos, eosinófilos e metarrubríctos), pois esta é a quantidade da menor delas. Essa escolha foi feita para se obter um melhor balanceamento durante o treinamento.

O modelo classificador para este trabalho foi construído com base no uso de CNN, transferência de aprendizado e *data augmentation*. Para isso foi utilizada a estrutura da VGG16, rede treinada com a base de dados ImageNet (DENG et al., 2009).

A Figura 18 mostra exemplos das células de linfócitos e neutrófilos capturadas para treinamento da rede neural.

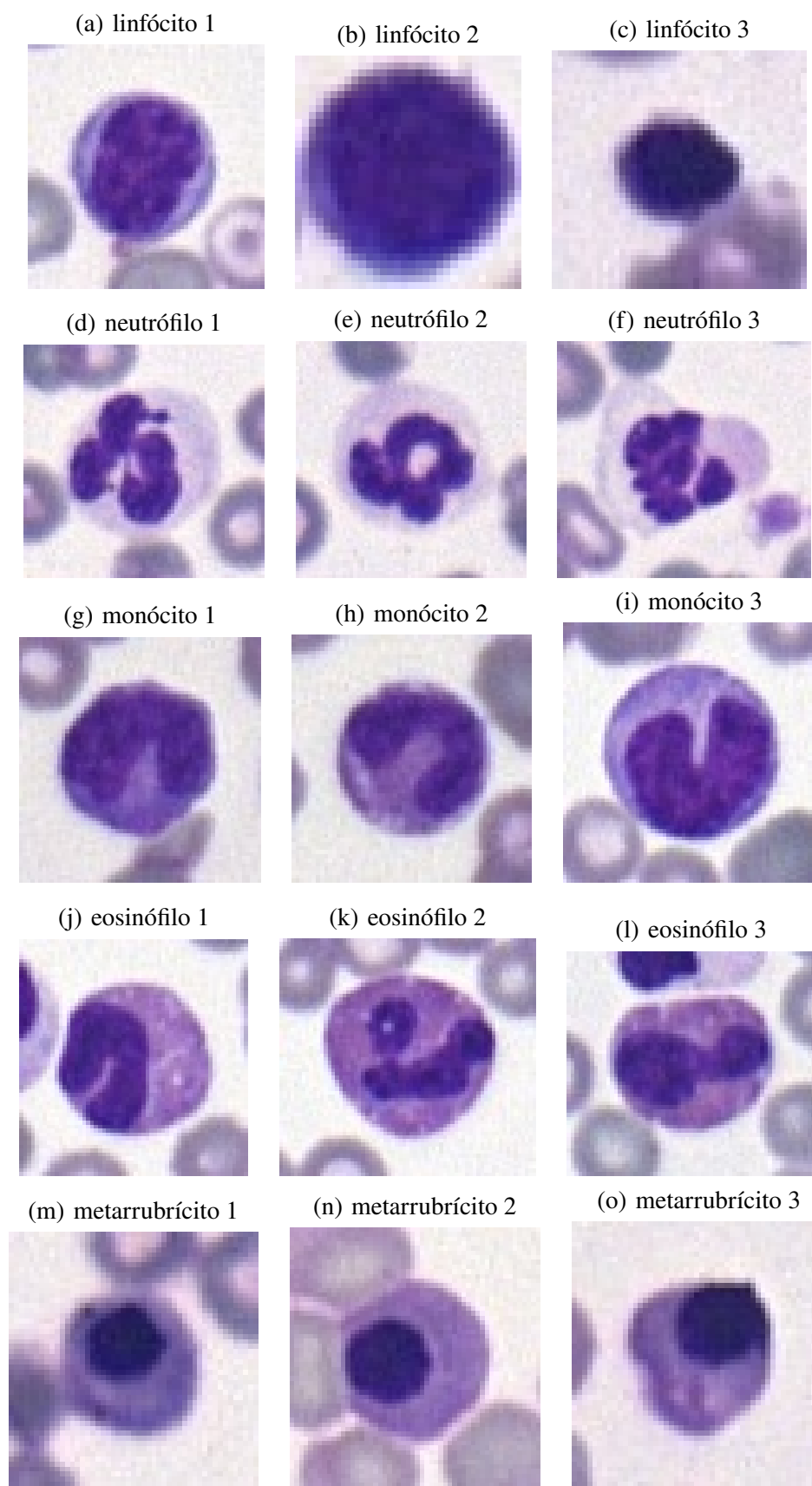


Figura 18 – Exemplos das células capturadas

5 Resultados

Este capítulo é destinado à apresentação dos resultados obtidos para a resolução do problema proposto neste trabalho, classificar automaticamente leucócitos de cães. Na [Seção 5.1](#) serão explanadas as configurações de métodos e experimentos desenvolvidos para alcançar os resultados idealizados. Serão abordados tópicos como a montagem da base de dados, métricas utilizadas, configuração dos métodos de segmentação e classificação e comparações entre métodos. Já na [Seção 5.2](#) serão apresentados os resultados obtidos com o método de detecção e classificação automática de leucócitos proposto neste trabalho e por fim, na [Seção 5.3](#), os resultados apresentados serão avaliados e discutidos.

5.1 Configuração de experimentos

Antes de apresentar os resultados é importante demonstrar as especificações de cada componente do processo da classificação automática de leucócitos caninos. Esta seção, portanto, é responsável por apresentar informações como valores dos parâmetros dos métodos de segmentação e de classificação, além de equipamentos e configurações utilizados na montagem da base de dados, possibilitando dessa maneira a reprodutibilidade dos experimentos.

5.1.1 Montagem da base de dados

Foram escaneadas e reunidas diversas imagens de amostras sanguíneas de dois cães, totalizando 433. As referidas imagens foram adquiridas e rotuladas por patologistas clínicos da Escola de Veterinária da UFMG, o que contribui para a solidez e confiança do trabalho. Cada uma das imagens possui pelo menos um leucócito, além, claro, das demais células sanguíneas, como hemácias e plaquetas. Devido ao fato de que a aquisição de imagens relevantes a partir de um esfregaço sanguíneo (aquelas amostras que apresentam diversos leucócitos) é um processo demorado e de difícil realização, pois depende de profissional habilitado e equipamentos, não foi possível, em tempo hábil, adquirir mais amostras, porém a quantidade foi suficiente para apresentar bons resultados.

As imagens foram adquiridas utilizando o *scanner* de lâminas 3D da HISTECH, como pode ser visto na [Figura 19](#) e para os esfregaços sanguíneos, geradores das imagens utilizadas neste trabalho, foi utilizado um corante chamado panótico rápido.

Em todas as imagens foram encontrados 1061 neutrófilos, 98 linfócitos, 130 monócitos, 236 eosinófilos e nenhum basófilo. Por não ser possível encontrar nenhum basófilo nas imagens escaneadas, outra classe foi incluída no processo de construção deste trabalho: os metarrubríctos. Estes elementos são eritroides jovens, ainda nucleados, que por sua vez, quando presentes na corrente sanguínea, podem indicar a presença de determinadas doenças, como anemia. No



Figura 19 – Scanner 3D utilizado para aquisição das imagens

total, foram localizados e rotulados 70 metarrubricitos. Todos os leucócitos encontrados nas imagens foram devidamente categorizados por profissionais da área veterinária, o que garante a integridade das informações.

5.1.2 Métricas de avaliação

Para avaliar os resultados do método proposto neste trabalho, foram reunidas algumas métricas de avaliação. Acurácia, precisão, revocação e F1-Score são as principais métricas utilizadas para avaliar os modelos de classificação e foram devidamente calculadas para todos os modelos gerados. Já para etapas de segmentação, a métrica utilizada foi a taxa de acertos, calculada por meio da divisão entre a quantidade de elementos segmentados e localizados corretamente pela quantidade total de elementos presentes nas imagens. Para o experimento de classificação feito com patologistas profissionais, também foi utilizada a taxa de acertos.

Abaixo é possível observar com detalhes o significado das métricas utilizadas neste trabalho:

- Acurácia: Mede a proporção entre elementos corretamente classificados e o total de elementos, com o papel de informar o desempenho geral do modelo;
- Precisão: Mede a proporção de elementos classificados corretamente como positivos em relação todas as classificações dadas como positivas;

- Revocação: Também conhecida como sensibilidade, mede a proporção dos elementos positivos corretamente classificados em relação ao total de elementos positivos;
- F1-Score: É a média harmônica entre precisão e revocação, levando em conta tanto falsos positivos quanto falsos negativos;

5.1.3 Segmentação

Para que os leucócitos pudessem ser encontrados e separados da imagem original, foi feita uma binarização baseada na componente da saturação do espaço HSV, dado que para ela, os núcleos das células de interesse apresentam um valor relativamente alto. Foi definido um limiar X por meio da função *threshold* do OpenCV para determinar se determinado elemento poderia ser considerado um núcleo válido. Isso quer dizer que, em relação à componente S do espaço de cor HSV, aqueles elementos detentores de intensidade entre X e 255 (máximo) são considerados núcleos válidos.

A fim de se remover restos celulares, elementos que poderiam atrapalhar a etapa de segmentação, foi utilizada técnica semelhante, identificando-os na imagem por meio da componente V do espaço de cor YUV. Com a definição do limiar Z , os elementos indesejados, detentores de intensidade entre Z e 255, foram removidos da cena.

Dentre os erros mais comuns encontrados no processo de segmentação, dois podem ser apontados como principais: junção e separação de núcleos.

Pode ocorrer de duas células estarem muito próximas uma da outra a ponto da técnica de dilatação expandir os contornos suficientemente para que seus núcleos se toquem. Tal fato resulta na segmentação incorreta desse elemento, inferindo que se trata de apenas uma célula, enquanto, na verdade, o caso se refere a dois elementos.

O contrário também pode ser visto no meio dos erros detectados. Em determinadas células, principalmente neutrófilos segmentados, o núcleo se apresenta dividido em diversas seções, algumas totalmente desconectadas, outras possuindo fracas conexões, que em algumas vezes não possuem saturação suficiente (acima do limiar X) para serem consideradas parte de um elemento válido. Dessa forma, um único núcleo pode ser considerado múltiplos, resultando em outro tipo de erro da segmentação.

Ainda sobre o tópico de dividir erroneamente os núcleos, tem o caso em que os elementos separados se tornam muito pequenos a ponto de não serem considerados mais como elementos nucleares válidos, eliminando completamente aquela amostra da cena.

Este exemplo pode ser visto na [Figura 20](#). A célula do canto esquerdo inferior foi totalmente removida, pois os 4 segmentos de seu núcleo, separados, são muito pequenos para formar um elemento válido. Já os dois primeiros casos podem ser vistos na imagem [Figura 21](#). Neste exemplo é possível ver um núcleo (a) sendo dividido erroneamente em dois (b), enquanto em (c) duas células se juntam para formar um único contorno.

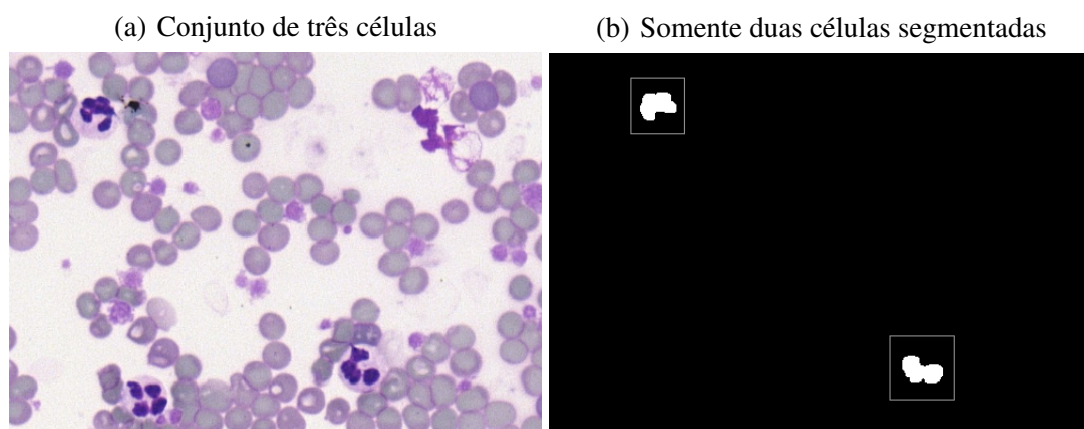


Figura 20 – Exemplo de segmentação incorreta em que uma célula é completamente eliminada da cena, por ter núcleo segmentado que se separou, não atende as especificações definidas para se enquadrar em um núcleo válido.

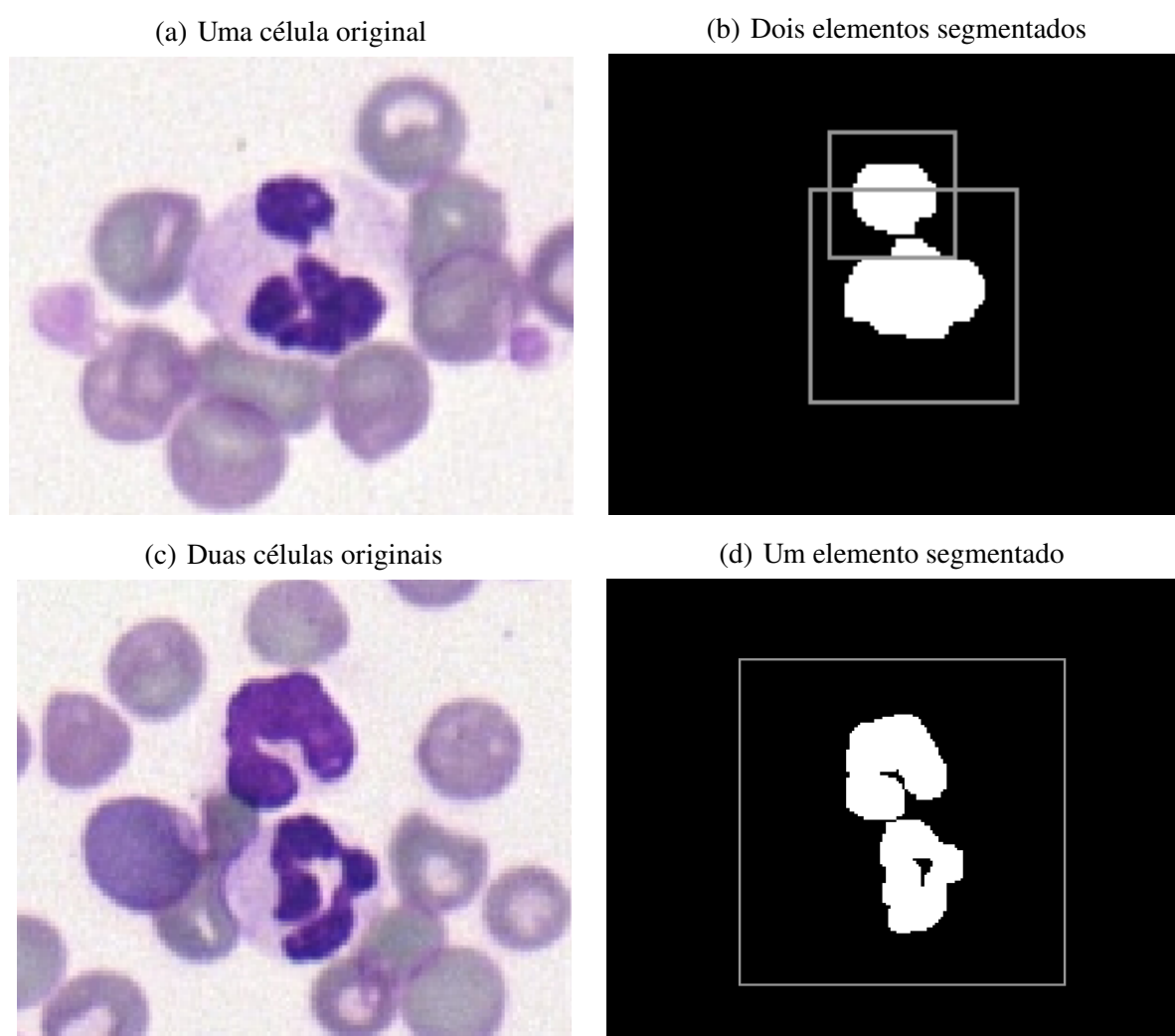


Figura 21 – Segmentação incorreta de algumas células. No primeiro caso, uma célula é considerada duas, enquanto no segundo o contrário ocorre.

5.1.4 Classificação

Para a etapa de classificação, foram consideradas cinco classes para o treinamento do modelo de rede neural: linfócitos, neutrófilos, monócitos, eosinófilos e metarrubríctos.

Foram executados ao todo, dez treinamentos distintos, cada um contendo 70 imagens de cada classe. Dentro das 70 imagens, 56 foram utilizadas para a etapa de treino e validação (construção dos modelos) e 14 foram reservadas para o teste (avaliação dos modelos criados com imagens nunca vistas por ele). A cada treinamento, as 70 imagens (assim como sua distribuição entre os conjuntos de treino, validação e teste) foram escolhidas de maneira aleatória e as imagens reservadas para teste não tiveram contato com o modelo antes da inferência, na intenção de ser remover possíveis vieses na configuração de parâmetros. Ou seja, a cada treinamento, o modelo foi alimentado por 56 imagens e as 14 separadas para teste foram as responsáveis por dar a acurácia final para o modelo em questão.

Os dez treinamentos foram executados com basicamente a mesma configuração, a única exceção foi o *data augmentation*. Cinco modelos foram treinados sem esse recurso, enquanto para os outros cinco foi utilizado o *ImageDataGenerator*, do pacote *Keras* do *Tensorflow*. A ideia do *ImageDataGenerator* é modificar sutilmente as imagens de treino a cada época, introduzindo essas pequenas modificações ao longo de toda a cadeia do treinamento dos modelos, que obtiveram resultados significativamente melhores.

A utilização do *data augmentation* na etapa de treinamento do modelo contribui positivamente por introduzir novas características a um mesmo exemplo durante a etapa de aprendizado do modelo. Mudanças sutis, como rotação, *zoom*(ampliação) e inversão da orientação nas imagens, fazem com que o modelo consiga entender que variações desse tipo poderão ocorrer, mesmo que não existam fisicamente nas imagens reservadas para o treinamento. Essa geração de exemplos sutilmente modificados para o treinamento, auxilia na generalização do modelo, pois é como se aumentasse a quantidade de exemplos diferentes para o aprendizado.

Para cada um dos dez treinamentos a principal métrica utilizada para avaliação dos modelos gerados foi a acurácia sobre o teste (aplicação do modelo treinado em imagens nunca vistas por ele durante o treinamento), também presente na grande parte dos trabalhos relacionados. As medidas de acurácia foram calculadas para os dez modelos, cinco deles contando com *data augmentation* ao longo do treinamento e cinco sem esse recurso.

5.1.5 Testes em outra bases de dados

Por compreender uma base de imagens relativamente pequena em relação ao que geralmente se vê nos estudos dessa área, dúvidas quanto à qualidade dos modelos gerados podem surgir, dado que é real a possibilidade de não generalização durante a etapa de treinamento.

A fim de diminuir esse tipo de dúvida, testes de classificação foram realizados com a base de imagens LISC. Esta base possui 5 classes: linfócitos, neutrófilos, basófilos, monócitos e eosinófilos e compreende somente células humanas, dado que bases formadas por células

de animais domésticos, mais especificamente cães, como no caso deste projeto, não foram encontradas.

Para que os testes pudessem ser devidamente coerentes, não houve alteração de nenhum parâmetro utilizado no treinamento da base de imagens deste trabalho. A distribuição entre treino, validação e teste também se manteve a mesma, aleatória e numericamente equilibrada. Foram gerados 5 modelos para a base externa, utilizando *data augmentation* e ao final, a média das acurácias foi calculada.

Foram utilizadas um total de 39 imagens para cada classe, reservando 7 delas para o teste ao fim de cada treinamento.

5.1.6 Testes com patologistas

A fim de realizar a comparação de um mesmo fluxo de trabalho realizado entre o método de classificação e profissionais capacitados, foi solicitado que 5 patologistas clínicos parceiros fizessem a classificação dos leucócitos utilizados ao longo deste trabalho. Os profissionais receberam as imagens e, apenas por meio de características visuais e conhecimento prévio adquirido, realizaram a classificação das células.

Para a classificação dos patologistas, foi considerada uma métrica mais simples de taxa de acerto, dividindo a quantidade de respostas corretas de cada patologista pela quantidade total analisada.

5.2 Resultados

Para a etapa de segmentação, os valores de limiares que trouxeram melhores resultados para a correta identificação dos núcleos dos leucócitos foram 158 para a componente S do espaço HSV e 162 para a componente V do espaço YUV. Ao todo foram utilizadas 333 imagens para essa validação, resultando em um total de 1436 amostras de células. O algoritmo conseguiu localizar corretamente 1390 células, resultando em uma taxa de acerto de 96.79%.

Para a etapa de classificação, algumas configurações foram definidas para o treinamento dos modelos. O tamanho da imagem de entrada foi definido como 56px e o tamanho do *batch* foi definido como 16. Todos os treinamentos foram executados com um total de 120 épocas a uma taxa de aprendizado de 0,000005. Para a aplicação da técnica de *data augmentation*, foram utilizados intervalo de rotação igual a 30, intervalo de deslocamento lateral e de altura iguais a 0,1, intervalo de cisalhamento igual a 0,2, intervalo de ampliação igual a 0,2, flip horizontal habilitado e modo de preenchimento como mais próximo.

Após os treinamentos e construção dos modelos, testes com imagens nunca vistas pelos modelos foram realizados e suas métricas foram devidamente calculadas.

É possível observar em [Tabela 2](#), [Tabela 3](#), [Tabela 4](#), [Tabela 5](#) e [Tabela 6](#) as matrizes de confusão referentes aos cinco treinamentos sem *data augmentation*, assim como suas respectivas métricas: acurácia, precisão, revocação e F1. Já na [Tabela 12](#) é possível analisar a

média das métricas desses cinco modelos, valor final utilizado para comparação com outros resultados.

Em [Tabela 7](#), [Tabela 8](#), [Tabela 9](#), [Tabela 10](#), [Tabela 11](#) é possível observar as matrizes de confusão referentes aos cinco treinamentos com o recurso de *data augmentation*, assim como suas respectivas métricas: acurácia, precisão, revocação e F1. Já na [Tabela 13](#) é possível analisar a média das métricas desses cinco modelos, sendo possível observar que todas as métricas obtiveram aumento significativo com a utilização das técnicas de *data augmentation*.

Tabela 2 – Modelo 1 *sem data augmentation* - Matriz de confusão e métricas

	Linfócito	Neutrófilo	Monócito	Eosinófilo	Metarrubricito
Linfócito	0,930	0,000	0,070	0,000	0,000
Neutrófilo	0,000	1,000	0,000	0,000	0,000
Monócito	0,070	0,000	0,790	0,140	0,000
Eosinófilo	0,070	0,000	0,000	0,930	0,000
Metarrubricito	0,000	0,070	0,000	0,000	0,930
		Acurácia	Precisão	Revocação	F-1
Linfócito		0,958	0,869	0,930	0,899
Neutrófilo		0,986	0,935	1,000	0,966
Monócito		0,944	0,919	0,790	0,849
Eosinófilo		0,958	0,869	0,930	0,899
Metarrubricito		0,986	1,000	0,930	0,964
Resultado geral		0,966	0,918	0,916	0,915

Tabela 3 – Modelo 2 *sem data augmentation* - Matriz de confusão e métricas

	Linfócito	Neutrófilo	Monócito	Eosinófilo	Metarrubricito
Linfócito	1,000	0,000	0,000	0,000	0,000
Neutrófilo	0,000	0,930	0,070	0,000	0,000
Monócito	0,000	0,000	0,790	0,210	0,000
Eosinófilo	0,000	0,000	0,140	0,860	0,000
Metarrubricito	0,000	0,000	0,000	0,000	1,000
		Acurácia	Precisão	Revocação	F-1
Linfócito		1,000	1,000	1,000	1,000
Neutrófilo		0,986	1,000	0,930	0,964
Monócito		0,916	0,790	0,790	0,790
Eosinófilo		0,930	0,804	0,860	0,831
Metarrubricito		1,000	1,000	1,000	1,000
Resultado geral		0,966	0,919	0,916	0,917

Tabela 4 – Modelo 3 *sem data augmentation* - Matriz de confusão e métricas

	Linfócito	Neutrófilo	Monócito	Eosinófilo	Metarrubricito
Linfócito	1,000	0,000	0,000	0,000	0,000
Neutrófilo	0,000	0,930	0,070	0,000	0,000
Monócito	0,000	0,000	0,930	0,070	0,000
Eosinófilo	0,000	0,000	0,210	0,790	0,000
Metarrubricito	0,000	0,000	0,000	0,000	1,000
		Acurácia	Precisão	Revocação	F-1
Linfócito		1,000	1,000	1,000	1,000
Neutrófilo		0,986	1,000	0,930	0,964
Monócito		0,930	0,769	0,930	0,842
Eosinófilo		0,944	0,919	0,790	0,849
Metarrubricito		1,000	1,000	1,000	1,000
Resultado geral		0,972	0,937	0,930	0,931

Tabela 5 – Modelo 4 *sem data augmentation* - Matriz de confusão e métricas

	Linfócito	Neutrófilo	Monócito	Eosinófilo	Metarrubricito
Linfócito	1,000	0,000	0,000	0,000	0,000
Neutrófilo	0,000	0,930	0,000	0,070	0,000
Monócito	0,000	0,000	0,860	0,140	0,000
Eosinófilo	0,070	0,000	0,070	0,860	0,000
Metarrubricito	0,000	0,000	0,000	0,000	1,000
		Acurácia	Precisão	Revocação	F-1
Linfócito		0,986	0,935	1,000	0,966
Neutrófilo		0,986	1,000	0,930	0,964
Monócito		0,958	0,925	0,860	0,891
Eosinófilo		0,930	0,804	0,860	0,831
Metarrubricito		1,000	1,000	1,000	1,000
Resultado geral		0,972	0,933	0,930	0,930

Tabela 6 – Modelo 5 *sem data augmentation* - Matriz de confusão e métricas

	Linfócito	Neutrófilo	Monócito	Eosinófilo	Metarrubricito
Linfócito	0,790	0,000	0,000	0,000	0,210
Neutrófilo	0,000	1,000	0,000	0,000	0,000
Monócito	0,000	0,000	0,790	0,140	0,070
Eosinófilo	0,000	0,000	0,210	0,790	0,000
Metarrubricito	0,000	0,000	0,000	0,000	1,000
		Acurácia	Precisão	Revocação	F-1
Linfócito		0,958	1,000	0,790	0,883
Neutrófilo		1,000	1,000	1,000	1,000
Monócito		0,916	0,790	0,790	0,790
Eosinófilo		0,930	0,849	0,790	0,819
Metarrubricito		0,944	0,781	1,000	0,877
Resultado geral		0,950	0,884	0,874	0,874

Tabela 7 – Modelo 1 **com data augmentation** - Matriz de confusão e métricas

	Linfócito	Neutrófilo	Monócito	Eosinófilo	Metarrubricito
Linfócito	0,930	0,000	0,000	0,000	0,070
Neutrófilo	0,070	0,860	0,000	0,000	0,070
Monócito	0,000	0,000	1,000	0,000	0,000
Eosinófilo	0,000	0,000	0,210	0,790	0,000
Metarrubricito	0,000	0,000	0,000	0,000	1,000
	Acurácia	Precisão	Revocação	F-1	
Linfócito	0,972	0,930	0,930	0,930	
Neutrófilo	0,972	1,000	0,860	0,925	
Monócito	0,958	0,826	1,000	0,905	
Eosinófilo	0,958	1,000	0,790	0,883	
Metarrubricito	0,972	0,877	1,000	0,935	
Resultado geral	0,964	0,928	0,916	0,915	

Tabela 8 – Modelo 2 **com data augmentation** - Matriz de confusão e métricas

	Linfócito	Neutrófilo	Monócito	Eosinófilo	Metarrubricito
Linfócito	1,000	0,000	0,000	0,000	0,000
Neutrófilo	0,000	1,000	0,000	0,000	0,000
Monócito	0,000	0,000	1,000	0,000	0,000
Eosinófilo	0,000	0,000	0,000	1,000	0,000
Metarrubricito	0,000	0,000	0,000	0,000	1,000
	Acurácia	Precisão	Revocação	F-1	
Linfócito	1,000	1,000	1,000	1,000	
Neutrófilo	1,000	1,000	1,000	1,000	
Monócito	1,000	1,000	1,000	1,000	
Eosinófilo	1,000	1,000	1,000	1,000	
Metarrubricito	1,000	1,000	1,000	1,000	
Resultado geral	1,000	1,000	1,000	1,000	

Tabela 9 – Modelo 3 **com data augmentation** - Matriz de confusão e métricas

	Linfócito	Neutrófilo	Monócito	Eosinófilo	Metarrubricito
Linfócito	0,930	0,000	0,000	0,070	0,000
Neutrófilo	0,000	0,930	0,070	0,000	0,000
Monócito	0,000	0,000	1,000	0,000	0,000
Eosinófilo	0,000	0,000	0,070	0,930	0,000
Metarrubricito	0,000	0,000	0,000	0,000	1,000
	Acurácia	Precisão	Revocação	F-1	
Linfócito	0,986	1,000	0,930	0,964	
Neutrófilo	0,986	1,000	0,930	0,964	
Monócito	0,972	0,877	1,000	0,935	
Eosinófilo	0,972	0,930	0,930	0,930	
Metarrubricito	1,000	1,000	1,000	1,000	
Resultado geral	0,983	0,961	0,958	0,958	

Tabela 10 – Modelo 4 **com** *data augmentation* - Matriz de confusão e métricas

	Linfócito	Neutrófilo	Monócito	Eosinófilo	Metarrubricito
Linfócito	0,860	0,000	0,000	0,070	0,070
Neutrófilo	0,000	1,000	0,000	0,000	0,000
Monócito	0,000	0,000	1,000	0,000	0,000
Eosinófilo	0,000	0,000	0,070	0,930	0,000
Metarrubricito	0,000	0,000	0,000	0,000	1,000
	Acurácia	Precisão	Revocação	F-1	
Linfócito	0,972	1,000	0,860	0,925	
Neutrófilo	1,000	1,000	1,000	1,000	
Monócito	0,986	0,935	1,000	0,966	
Eosinófilo	0,972	0,930	0,930	0,930	
Metarrubricito	0,986	0,935	1,000	0,966	
Resultado geral	0,983	0,960	0,958	0,957	

Tabela 11 – Modelo 5 **com** *data augmentation* - Matriz de confusão e métricas

	Linfócito	Neutrófilo	Monócito	Eosinófilo	Metarrubricito
Linfócito	0,930	0,000	0,000	0,000	0,070
Neutrófilo	0,000	1,000	0,000	0,000	0,000
Monócito	0,000	0,000	1,000	0,000	0,000
Eosinófilo	0,000	0,000	0,000	1,000	0,000
Metarrubricito	0,000	0,000	0,000	0,000	1,000
	Acurácia	Precisão	Revocação	F-1	
Linfócito	0,986	1,000	0,930	0,964	
Neutrófilo	1,000	1,000	1,000	1,000	
Monócito	1,000	1,000	1,000	1,000	
Eosinófilo	1,000	1,000	1,000	1,000	
Metarrubricito	0,986	0,935	1,000	0,966	
Resultado geral	0,994	0,987	0,986	0,986	

Tabela 12 – Média das métricas dos modelos sem *data augmentation*

Acurácia	Precisão	Revocação	F-1
0,965	0,918	0,913	0,913

Tabela 13 – Média das métricas dos modelos com *data augmentation*

Acurácia	Precisão	Revocação	F-1
0,985	0,967	0,964	0,963

Quanto aos testes com a base de dados LISC, é possível observar na [Tabela 14](#) a matriz de confusão e os resultados das médias das métricas por classe dos cinco modelos treinados para esta base de imagens.

Este trabalho obteve média de acurácia igual a 99,6% para a classificação dos leucócitos da base LISC, enquanto em outros trabalhos, como no de [Gupta et al. \(2019\)](#), a acurácia média alcançada foi de 97,3%. Já no trabalho de [Reena e Ameer \(2020\)](#), foi obtida uma acurácia média de 98,87%. Portanto, com base no que foi apresentado é possível dizer que, em relação à base LISC, o trabalho obteve resultados comparáveis ao estado da arte.

A [Tabela 15](#) representa o resultado obtido pelos patologistas profissionais e por meio deles podemos analisar que o método proposto neste trabalho obteve desempenho superior, atingindo acurácia de 98,5%.

Além da relação entre acurácia e taxa de acerto dos métodos, o quesito tempo é de grande importância. Enquanto os profissionais necessitaram de horas para concluir o trabalho, o método proposto alcançou tempo médio de 76,4 segundos para segmentar e classificar 333 imagens, contendo aproximadamente 1400 leucócitos, o que resulta em 4,4 inspeções por segundo. Esse resultado foi obtido utilizando o ambiente do Google colab, contando com um processador Intel(R) Xeon(R) CPU @ 2.20GHz e 13.3 GB de memória RAM disponíveis. Já ao utilizar aceleração de hardware via GPU, paralelizando a inferência dos modelos de predição, o tempo médio foi de 3,7 segundos, o que resulta em 90 inspeções por segundo. A GPU utilizada para este teste foi a Nvidia Tesla T4.

Tabela 14 – Média das métricas - 5 treinamentos com augmentation - Base LISC

	Acurácia	Precisão	Revocação	F-1
Linfócito	0,994	1,000	0,972	0,986
Neutrófilo	1,000	1,000	1,000	1,000
Monócito	0,994	0,973	1,000	0,986
Eosinófilo	0,994	1,000	0,972	0,986
Metarrubricito	0,994	0,973	1,000	0,986
Resultado geral	0,996	0,989	0,989	0,989

Tabela 15 – Taxa de acerto dos profissionais

Profissionais	Taxa de acerto
Patologista 1	0,937
Patologista 2	0,946
Patologista 3	0,929
Patologista 4	0,922
Patologista 5	0,935
Média	0,934

5.3 Análise e discussão dos resultados

Com base nos resultados analisados nas seções anteriores, é possível observar que as técnicas de segmentação empregadas neste trabalho foram suficientes para alcançar nível próximo aos principais trabalhos encontrados na literatura, alcançando uma taxa de acertos de 96,76%. O trabalho de [Cao, Liu e Song \(2018\)](#), focado exclusivamente em segmentar os leucócitos, alcançou acurácia média de 93,75%, já [Reena e Ameer \(2020\)](#), alcançaram uma acurácia de 98,22% para a etapa de segmentação.

Em relação à classificação, é possível observar que os resultados obtidos se aproximam das referências tomadas como base para a construção deste trabalho, que alcançou uma acurácia média de 98,5%, enquanto no trabalho de [Wang e Cao \(2020\)](#), que contava com uma base de imagens de quase 30000 exemplares e com utilização de técnicas avançadas de *data augmentation*, foi obtida uma acurácia de 97,6%. Já [Reena e Ameer \(2020\)](#) obtiveram acurácia de 98,87%, ao utilizarem técnicas de aprendizado profundo e transferência de aprendizado, enquanto [Lin e Wang \(2018\)](#), utilizando um classificador baseado em SVM, alcançaram 97,96%

Por fim, testes de classificação em outra base de dados (LISC) se mostraram superiores aos encontrados na literatura e em relação aos métodos manuais, qualidade e desempenho também se mostraram superiores, quando comparados aos resultados fornecidos pelos patologistas profissionais.

6 Conclusão

Assim como na medicina humana convencional, a medicina veterinária possui grande interesse nos resultados que o avanço tecnológico computacional pode gerar. Tomando como foco específico o diagnóstico e tratamento de doenças relacionadas a animais domésticos, a preocupação com a qualidade e rapidez dos procedimentos é fundamental. Muitos desses animais, como cães e gatos, são parte fundamental da rotina doméstica de milhões de lares ao redor do mundo, se tornando, de certa maneira, parte dessas famílias, dado que os vínculos afetivos e emocionais construídos ao longo de suas vidas se tornam substancialmente fortes. Portanto, hoje é de grande importância que, por meio da tecnologia, seja possível minimizar os erros e maximizar a qualidade dos processos e métodos utilizados para diagnosticar e tratar diversas doenças em animais domésticos.

Com base nesse contexto, este trabalho propôs a construção de um método de detecção e classificação automática de leucócitos presentes em amostras de sangue de cães, a fim de auxiliar patologistas clínicos no diagnóstico de doenças detectáveis por meio desse material, como anemia.

Diferentemente dos métodos convencionais, a análise proposta neste trabalho, é feita com o uso de técnicas de visão computacional atreladas ao aprendizado de máquina. Dessa forma, imagens de amostras sanguíneas são analisadas automaticamente pelo software, que por sua vez detecta os principais leucócitos presentes naquela cena, informando o usuário o tipo de cada uma das células encontradas.

A resposta do método pode dar ao especialista, em poucos segundos, informações relevantes para o diagnóstico de doenças sem todo o esforço humano envolvido no processo manual de identificação, contagem e classificação das células, acelerando o processo, o que é de grande relevância para os animais em espera de diagnóstico, além de aumentar a própria qualidade de vida dos profissionais envolvidos no processo, que conseguem realizar com menos esforço e morosidade, a mesma quantidade de trabalho. Em relação à segmentação, o método proposto obteve taxa de acerto de 96,76%, localizando corretamente 1390 leucócitos presentes nas amostras sanguíneas analisadas, de um total de 1436. Já em relação à classificação dessas células, a acurácia média obtida para os modelos gerados foi de 98,5%, o que fica próximo dos trabalhos relacionados e do estado da arte atualmente. Além da qualidade, é importante destacar o desempenho, pois o método proposto consegue classificar milhares de leucócitos contidos em centenas de imagens em poucos segundos, variando de 4 a 90 inspeções por segundo, dependendo do *hardware* utilizado no processo.

6.1 Próximos passos

Para a construção do método foi necessária a produção de uma base de dados com mais de 400 imagens de amostras sanguíneas de cães, o que gerou mais de 1500 exemplares de leucócitos. Ao se analisar os trabalhos similares desenvolvidos atualmente, é possível analisar que essa base ainda se trata de uma coleção pequena, dado que vários dos trabalhos contam com dezenas de milhares de exemplares. Como algumas células são mais raras de se encontrar do que outras (basófilos são muito raros, enquanto neutrófilos são abundantes), é um desafio alcançar um número satisfatório de exemplares de todas as classes. É importante ressaltar que quanto mais exemplares de células disponíveis, maiores as chances de construir modelos mais acurados, pois mais informações estariam disponíveis para que eles possam aprender cada vez mais as diferentes características dos leucócitos.

Além do desafio da base de dados, há espaço para melhorias em relação à segmentação das células. Neste trabalho foram encontradas algumas dificuldades quanto à aglutinação incorreta de duas células ou separação de uma única célula. Um método de correção focado em resolver essas duas situações pode ser desenvolvido para que a qualidade dessa etapa seja melhorada, consequentemente melhorando todo o processo de classificação.

Referências

- ALBAWI, S.; MOHAMMED, T. A.; AL-ZAWI, S. Understanding of a convolutional neural network. In: IEEE. **2017 international conference on engineering and technology (ICET)**. [S.l.], 2017. p. 1–6. Citado na página 28.
- ALLIE. **How we see colors and common notations of colors**. 2017. Disponível em: <<https://allielearning.wordpress.com/2017/10/19/how-we-see-colors-and-common-notations-of-colors/>>. Acesso em: 30 de janeiro de 2022. Citado na página 20.
- ALZUBAIDI, L. et al. Deep learning models for classification of red blood cells in microscopy images to aid in sickle cell anemia diagnosis. **Electronics**, Multidisciplinary Digital Publishing Institute, v. 9, n. 3, p. 427, 2020. Citado na página 11.
- BELAN, P. A. et al. Sistema de visão computacional para inspeção da qualidade de grãos de feijão. Universidade Nove de Julho, 2019. Citado 2 vezes nas páginas 27 e 28.
- BISHOP, C. M.; NASRABADI, N. M. **Pattern recognition and machine learning**. [S.l.]: Springer, 2006. v. 4. Citado na página 24.
- CAO, H.; LIU, H.; SONG, E. A novel algorithm for segmentation of leukocytes in peripheral blood. **Biomedical Signal Processing and Control**, Elsevier, v. 45, p. 10–21, 2018. Citado 3 vezes nas páginas 32, 34 e 53.
- CHAUHAN, A. **Understand CNN Basics with a Keras Example in Python**. 2021. Disponível em: <<https://pub.towardsai.net/understand-cnn-basics-with-a-keras-example-in-python-c1fd6c449935>>. Acesso em: 07 de fevereiro de 2022. Citado na página 29.
- CHOLLET. **Deep Learning With Python**. [S.l.]: Manning, 2018. Citado 2 vezes nas páginas 23 e 28.
- DENG, J. et al. Imagenet: A large-scale hierarchical image database. In: **2009 IEEE Conference on Computer Vision and Pattern Recognition**. [S.l.: s.n.], 2009. p. 248–255. Citado 2 vezes nas páginas 30 e 40.
- FILHO, O. M.; NETO, H. V. **Processamento digital de imagens**. [S.l.]: Brasport, 1999. Citado 3 vezes nas páginas 15, 16 e 22.
- GÉRON, A. **Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems**. [S.l.]: "O'Reilly Media, Inc.", 2019. Citado na página 27.
- GOMES, D. d. F. **Visão computacional e segmentação de imagens por discriminação de textura**. Dissertação (Mestrado) — Universidade Federal de Pernambuco, 2009. Citado na página 22.
- GONZALEZ, R. C.; WOODS, R. E. et al. **Digital image processing**. [S.l.]: Prentice hall Upper Saddle River, NJ, 2002. Citado 2 vezes nas páginas 15 e 22.
- GUPTA, D. et al. Optimized binary bat algorithm for classification of white blood cells. **Measurement**, Elsevier, v. 143, p. 180–190, 2019. Citado na página 52.

- HAYKIN, S. **Redes neurais: princípios e prática**. [S.l.]: Bookman Editora, 2007. Citado 2 vezes nas páginas 25 e 26.
- KLETTE, R. **Concise computer vision**. [S.l.]: Springer, 2014. Citado 5 vezes nas páginas 15, 17, 18, 19 e 22.
- KOSCHAN, A.; ABIDI, M. **Digital color image processing**. [S.l.]: John Wiley & Sons, 2008. Citado na página 19.
- KUAN, D.-H. et al. A microfluidic device for simultaneous extraction of plasma, red blood cells, and on-chip white blood cell trapping. **Scientific reports**, Nature Publishing Group, v. 8, n. 1, p. 1–9, 2018. Citado na página 11.
- LIN, L.; WANG, W. A robust leukocyte recognition method based on multi-scale regional growth and mean-shift clustering. **Journal of Algorithms & Computational Technology**, SAGE Publications Sage UK: London, England, v. 12, n. 3, p. 208–216, 2018. Citado 4 vezes nas páginas 12, 33, 34 e 53.
- MEIJDEN, P. E. van der; HEEMSKERK, J. W. Platelet biology and functions: new concepts and clinical perspectives. **Nature Reviews Cardiology**, Nature Publishing Group, v. 16, n. 3, p. 166–179, 2019. Citado na página 11.
- OPENCV. **Thresholding Operations using inRange**. Disponível em: <https://docs.opencv.org/3.4/da/d97/tutorial_threshold_inRange.html>. Acesso em: 02 de fevereiro de 2022. Citado na página 21.
- RAMIAH, M. **Deep Learning- using ResNets for Transfer Learning**. 2019. Disponível em: <<https://madhuramiah.medium.com/deep-learning-using-resnets-for-transfer-learning-d7f4799fa863>>. Acesso em: 10 de fevereiro de 2022. Citado na página 30.
- REENA, M. R.; AMEER, P. Localization and recognition of leukocytes in peripheral blood: A deep learning approach. **Computers in Biology and Medicine**, Elsevier, v. 126, p. 104034, 2020. Citado 5 vezes nas páginas 11, 33, 34, 52 e 53.
- ROSEBROCK, A. **OpenCV Color Spaces (cv2.cvtColor)**. 2021. Disponível em: <<https://www.pyimagesearch.com/2021/04/28/opencv-color-spaces-cv2-cvtColor/>>. Acesso em: 30 de janeiro de 2022. Citado na página 21.
- RUSHITI, B. **Classical Machine Learning — Supervised Learning Edition**. 2020. Disponível em: <<https://medium.com/swlh/classical-machine-learning-7efc6674fca1>>. Acesso em: 04 de fevereiro de 2022. Citado na página 24.
- SANTOS, L.; LEME, F. Comparação entre metodologias analíticas para a realização de hemogramas na medicina veterinária. **Revista VZ Em Minas**, n. 148, p. 29–37, 2021. Citado 2 vezes nas páginas 11 e 12.
- SARASWAT, M.; ARYA, K. Automated microscopic image analysis for leukocytes identification: A survey. **Micron**, Elsevier, v. 65, p. 20–33, 2014. Citado na página 12.
- SILVA, I. N. D. et al. Artificial neural networks. **Cham: Springer International Publishing**, Springer, v. 39, 2017. Citado na página 26.
- STUART, M. J.; NAGEL, R. L. Sick-cell disease. **The Lancet**, Elsevier, v. 364, n. 9442, p. 1343–1360, 2004. Citado na página 11.

- SZELISKI, R. **Computer vision: algorithms and applications**. [S.l.]: Springer Science & Business Media, 2010. Citado na página 18.
- TAN, C. et al. A survey on deep transfer learning. In: SPRINGER. **International conference on artificial neural networks**. [S.l.], 2018. p. 270–279. Citado na página 29.
- WANG, Y.; CAO, Y. Human peripheral blood leukocyte classification method based on convolutional neural network and data augmentation. **Medical physics**, Wiley Online Library, v. 47, n. 1, p. 142–151, 2020. Citado 3 vezes nas páginas 32, 34 e 53.
- YILDIRIM, M.; CINAR, A. C. Classification of white blood cells by deep learning methods for diagnosing disease. **Rev. d’Intelligence Artif.**, v. 33, n. 5, p. 335–340, 2019. Citado na página 11.
- ZHAO, J. et al. Automatic detection and classification of leukocytes using convolutional neural networks. **Medical & biological engineering & computing**, Springer, v. 55, n. 8, p. 1287–1301, 2017. Citado na página 12.
- ZHUANG, F. et al. A comprehensive survey on transfer learning. **Proceedings of the IEEE**, IEEE, v. 109, n. 1, p. 43–76, 2020. Citado na página 30.