



CEMEP

CENTRO MUNICIPAL DE ENSINO PROFISSIONALIZANTE

OSMAR PASSARELLI SILVEIRA

Avenida Brasil, 330, Vila Bressani, Paulínia, São Paulo

CEP – 13140496 – Telefone (19)38749457 – e-mail: sec.cemep@gmail.com

Felipe Mauricio Mundin

Maria Clara dos Santos

PLAQUET.AI:

Inteligência Artificial para segurança hematológica infantil

PAULÍNIA
2024



Felipe Mauricio Mordin

Maria Clara dos Santos

PLAQUET.AI:

Inteligência Artificial para segurança hematológica infantil

O presente relatório é orientado pela professora Dra. Eliza de Oliveira Cardoso e co-orientado pelo professor Me. Diogo Pelaes Franco Pereira para submissão em feiras científicas.

Orientador: Eliza de Oliveira Cardoso

Coorientador: Diogo Pelaes Franco Pereira

PAULÍNIA
2024



RESUMO

A contagem de plaquetas é uma etapa essencial para o diagnóstico de diversas doenças hematológicas, especialmente em crianças com leucemia, nas quais a ausência de transfusões pode aumentar o risco de óbito em até 67% (Gonçalves, 2019). A demora na obtenção desses resultados pode comprometer a intervenção clínica em casos críticos. Diante desse problema, o projeto tem como objetivo desenvolver um modelo de Inteligência Artificial (IA) baseado em *Deep Learning* para a contagem automática de plaquetas em amostras sanguíneas, oferecendo uma solução rápida e precisa. A justificativa para o uso de IA e Visão Computacional reside na necessidade de um método acessível, de baixo custo e eficiente para regiões com limitações de infraestrutura laboratorial. Como aponta Antonello (2017), a Visão Computacional atua como os “olhos” da máquina, permitindo análises automatizadas com alta precisão. A metodologia adotada inclui o uso de Python e dos *frameworks* Keras e Tensor Flow para a construção do modelo, que, a partir de dados obtidos em amostras reais e por meio de visitas ao Centro de Pesquisa Boldrini, é treinado para identificar e contar as lâminas de forma automatizada. Os resultados preliminares indicam uma alta precisão do modelo, o que pode contribuir significativamente para a prevenção de hemorragias graves e para a melhoria do tratamento em pacientes pediátricos. Esperamos que a implementação desse modelo em laboratórios reduza o tempo de diagnóstico e permita intervenções médicas mais rápidas e eficazes, principalmente em áreas de difícil acesso.

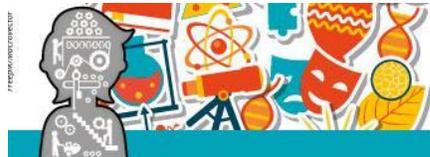
Palavras-Chave: Inteligência Artificial, Plaquetas, Leucemia.



ABSTRACT

Platelet counting is an essential step in diagnosing various hematological diseases, especially in children with leukemia, where the absence of transfusions can increase the risk of mortality by up to 67% (Gonçalves, 2019). Delays in obtaining these results may hinder timely clinical intervention in critical cases. To address this problem, the project aims to develop a Deep Learning-based Artificial Intelligence (AI) model for the automated counting of platelets in blood samples, providing a fast and accurate solution. The justification for using AI and Computer Vision lies in the need for an accessible, low-cost, and efficient method for regions with limited laboratory infrastructure. As noted by Antonello (2017), Computer Vision functions as the "eyes" of the machine, enabling highly precise automated analyses. The methodology includes using Python and frameworks such as Keras and TensorFlow to build the model, which is trained to identify and count platelets automatically using real sample data and insights from visits to the Boldrini Research Center. Preliminary results indicate high model accuracy, which could significantly contribute to preventing severe hemorrhages and accelerating treatment in pediatric patients. It is expected that the implementation of this model in laboratories will reduce diagnostic time and allow faster and more effective medical interventions, particularly in remote areas.

Key-Words: Artificial Intelligence, Platelets, Leukemia.



LISTA DE ILUSTRAÇÕES

Figuras

Figura I: Etapas do KDD Process, (MOURA,2011).....12

Figura II: Esquema de detecção YOLO (“Deep Learning:Detecção de Objetos em Imagens”, [s.d.].....13

Figura III: interface do modelo treinado com 15 épocas.....18

Figura IV: interface do modelo treinado com 30 épocas.....19



SUMÁRIO

1. INTRODUÇÃO.....	7
1.1. Problema.....	7
1.2. Justificativa.....	7
2. OBJETIVOS DO TRABALHO.....	9
2.1 Objetivos gerais.....	9
2.2 Objetivos específicos.....	9
3 RELEVÂNCIA DO TRABALHO.....	10
4 DESENVOLVIMENTO.....	11
4.1 Metodologia.....	11
4.2 Materiais e Ferramentas.....	12
4.4 Barreiras e Limitações.....	15
5 DESENVOLVIMENTO DO MODELO DE IA.....	16
5.1 Arquitetura do Modelo.....	16
5.2 Treinamento e Validação.....	17
6 RESULTADOS E CAMINHOS FUTUROS.....	18
7 CONCLUSÃO.....	19
8 REFERÊNCIAS.....	20



1. INTRODUÇÃO

1.1. Problema

À medida que a tecnologia avança, surge a necessidade de aprimorar técnicas diagnósticas para que se tornem mais eficientes e eficazes no campo da medicina. No caso de exames hematológicos, torna-se essencial que os métodos empregados acompanhem esses avanços. Foi identificado, através de revisões bibliográficas, que a contagem de plaquetas (também chamadas de trombócitos) ainda depende de métodos manuais ou de máquinas pesadas e caras, como a Sysmex XN 550, o que dificulta seu acesso e praticidade. Diante desse contexto, questiona-se: como é possível melhorar os métodos de contagem de plaquetas existentes?

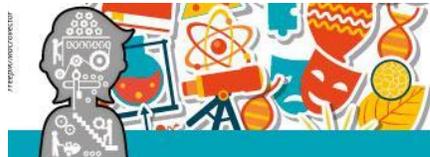
Apesar da existência de técnicas automatizadas realizadas por máquinas hematológicas, a contagem de plaquetas ainda pode ser feita manualmente, utilizando métodos como o Esfregaço Sanguíneo, a Contagem de Hemocítômetro e o Método de Fônio, sendo os dois primeiros muito comuns na identificação de células em geral. Embora esses métodos apresentem resultados semelhantes aos obtidos por máquinas, eles são desvantajosos quando há necessidade de rapidez nos diagnósticos, conforme sugerido por Schweirigert e colaboradores (2010).

Em crianças com leucemia, especialmente nos casos de Leucemia Linfóide Aguda (LLA) e Leucemia Mieloide Aguda (LMA), a produção de células sanguíneas é prejudicada, resultando em uma baixa contagem de plaquetas, essenciais para a coagulação. A contagem inadequada ou lenta dessas células pode acarretar complicações graves, incluindo hemorragias que, em casos extremos, podem levar ao óbito. Além disso, a lacuna de tempo entre a coleta e o resultado da contagem plaquetária pode agravar o quadro clínico dos pacientes, atrasando a medicação necessária e aumentando a dependência de transfusões de sangue.

1.2. Justificativa

A Inteligência Artificial (IA) e, mais especificamente, a Visão Computacional, têm o potencial de transformar esse cenário. O uso de IA está presente no cotidiano de maneira discreta, como na aplicação de filtros em redes sociais ou no reconhecimento de gestos para desbloquear smartphones. Conforme apontado por Antonello (2017), dar "visão" a uma máquina é um avanço significativo, permitindo que ela reconheça objetos e imagens com precisão. No contexto da medicina diagnóstica, isso pode ser aplicado à identificação de células em imagens microscópicas.

Dessa forma, através de *Machine Learning* e *Deep Learning*, é possível treinar máquinas para



identificar e contar plaquetas em amostras de sangue de forma automatizada, rápida e com alta precisão, oferecendo uma solução de baixo custo e acessível. Isso é especialmente relevante para pacientes com leucemia, pois a rapidez e a precisão na contagem de plaquetas podem evitar quadros hemorrágicos graves, melhorando significativamente o tratamento dessas crianças.

Além disso, a Confederação Nacional de Municípios (CNM) alerta que, durante os períodos festivos e de férias, o Brasil sofre uma redução de 20% a 30% nos estoques de sangue dos hemocentros, o que pode agravar a situação de pacientes que necessitam de transfusões. Portanto, este projeto busca desenvolver um sistema que forneça monitoramento constante e rápido dos níveis de plaquetas, contribuindo para um diagnóstico mais ágil e eficaz.



2. OBJETIVOS DO TRABALHO

2.1 Objetivos gerais

Explorar a aplicação da Visão Computacional e do *Deep Learning* na criação de um sistema informatizado que realizará a contagem automatizada das plaquetas e aferirá a sua quantidade no sangue, diminuindo o custo benefício da contabilização de trombócitos, melhorando a acurácia e dando agilidade aos resultados.

2.2 Objetivos específicos

Temos como principais objetivos específicos:

- Identificar e analisar potenciais obstáculos ou impedimentos presentes no processo diagnóstico relacionado à detecção de baixos níveis de plaquetas;
- Realizar uma especialização sobre as bases da estruturação de Inteligências Artificiais, ou seja, fundamentos sobre Visão Computacional e Deep Learning;
- Encontrar um *dataset* (Banco de Imagens) de imagens adequado para a construção e treinamento de um algoritmo de processamento de imagens;
- Desenvolver um sistema capaz de realizar a contagem automatizada de trombócitos em amostras sanguíneas por meio da aplicação de técnicas de Inteligência Artificial como o *Deep Learning* e a Visão Computacional;
- Aprimorar a acurácia e celeridade dos processos, utilizando métodos e tecnologias avançadas, como o *Dropout* para evitar *Overfitting* na rede neural, com o intuito de garantir resultados mais precisos e rápidos;
- Utilizar o sistema para antecipar o diagnóstico de quadros hemorrágicos, com o objetivo de prevenir a progressão do quadro clínico para estágios mais graves;
- Otimizar relações entre custo e benefício associada à implementação de sistemas de contagem automatizada, visando maximizar a eficiência operacional e minimizar os custos envolvidos;
- Testar o sistema rigorosamente para avaliar a eficácia, robustez e confiabilidade do sistema desenvolvido, garantindo sua adequação aos padrões de desempenho esperados e alcançar uma boa acurácia.



3 RELEVÂNCIA DO TRABALHO

O projeto se destaca por sua capacidade de atender regiões remotas, como comunidades ribeirinhas, onde grandes máquinas automatizadas para contagem de plaquetas são inacessíveis. Assim, a solução proposta com IA e Visão Computacional é viável para que um microscopista treinado, usando apenas um dispositivo portátil, possa realizar análises precisas de amostras. Com efeito, isso leva a tecnologia de ponta para áreas onde equipamentos avançados são impraticáveis.

Além disso, o projeto oferece um excelente custo-benefício. Máquinas automatizadas são caras e muitas clínicas pequenas ou hospitais em áreas desfavorecidas não têm condições de adquiri-las. Portanto, nossa solução democratiza o acesso à contagem de plaquetas, disponibilizando uma alternativa acessível, especialmente para populações vulneráveis, como crianças com leucemia, cuja condição demanda monitoramento constante. Dessa forma, o projeto pode desempenhar um papel essencial na prevenção de complicações graves, como hemorragias.

Ademais, a automação e escalabilidade do sistema são vantagens fundamentais. O uso da IA permite processar um grande número de amostras, aumentando a eficiência, mesmo com recursos limitados. Assim sendo, em áreas com poucos profissionais de saúde, a solução pode suprir a demanda de análise laboratorial e contribuir no acompanhamento de pacientes em larga escala.

Por fim, a facilidade de implementação da solução é notável. Como resultado, com equipamentos simples de laboratório, como um microscópio, e um dispositivo com software de IA, a metodologia pode ser aplicada em diversas configurações. Logo, trata-se de uma proposta adaptável, que pode beneficiar tanto locais de infraestrutura limitada quanto profissionais que necessitam de ferramentas acessíveis e eficientes para realizar diagnósticos críticos.



4 DESENVOLVIMENTO

4.1 Metodologia

O presente trabalho foi dividido em um conjunto de etapas, cada qual com um importante papel para o desenvolvimento do projeto.

Iniciou-se a busca e coleta de artigos acadêmicos a fim de definir o problema central. Após sua identificação, fomos em busca de dados que comprovassem a eficácia dos métodos automatizados sobre os manuais. Também realizamos um curso da plataforma *APICE* sobre métodos estatísticos para ter uma base sobre como analisar e utilizar as informações coletadas.

A partir dessas etapas se deu a pesquisa em relação à Visão Computacional, onde realizamos um curso sobre o tema na plataforma *SENAI Play* para compreender como estruturar algoritmos implementando a IA nos mesmos. Também buscamos por *datasets* de imagens enviando e-mails para instituições e contatos externos.

Inserimos o modelo KDD (*Knowledge Discovery in Databases*) como base metodológica para estruturar o ciclo de desenvolvimento da solução de IA. O modelo KDD é composto por cinco etapas principais: seleção, pré-processamento, transformação, mineração de dados e interpretação dos resultados. Primeiramente, na fase de seleção, reunimos e escolhemos as imagens e lâminas de esfregaços sanguíneos para compor o *dataset* de treinamento e validação. Em seguida, no pré-processamento, utilizamos técnicas de manipulação de imagens com OpenCV para melhorar a qualidade e adequação das imagens ao nosso modelo de IA. A fase de transformação envolveu a normalização e ajuste das imagens para padronizar os dados de entrada. Na etapa de mineração de dados, foi aplicado o modelo YOLO, utilizando *Deep Learning* para a detecção de plaquetas. Finalmente, os resultados foram avaliados em termos de precisão e tempo de processamento, o que nos permitiu ajustar o modelo e identificar os pontos de melhoria para aumentar a acurácia.

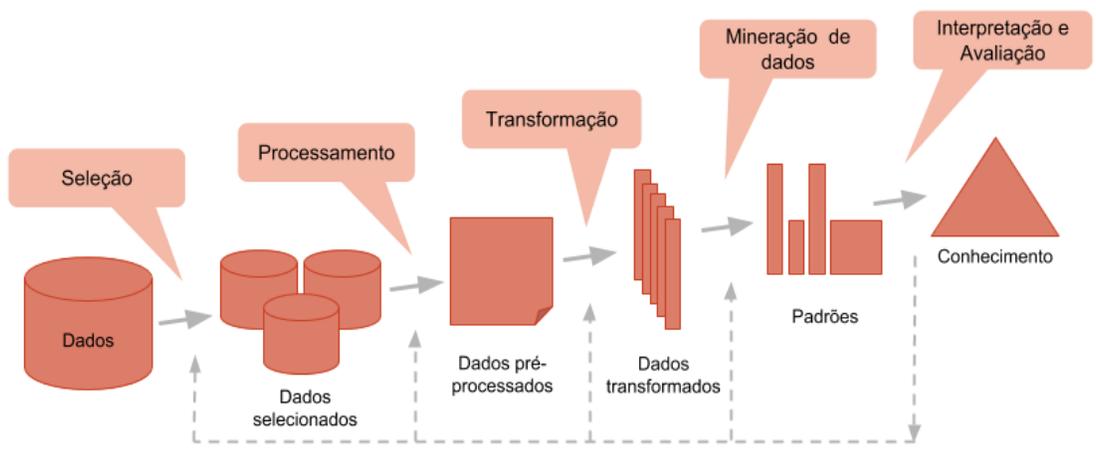


Figura I: Etapas do KDD Process, (MOURA,2011).

Em sequência, agendamos uma visita técnica ao Centro de Pesquisa Boldrini na cidade de Campinas, com o objetivo de aplicar um questionário aos profissionais da área hematológica para compreendermos como as contagens automáticas são realizadas.

Visando compreender mais minuciosamente a construção de um software de IA, iniciamos o curso de programação em Python e Inteligência Artificial fornecido pela *Samsung Innovation Campus*, onde foi possível ter os conhecimentos sobre as principais bibliotecas de Python como Pytorch e Scikit-Learn e quais caminhos seguir na estruturação do modelo.

Tendo toda esta base, foi possível instaurar a construção do sistema como um todo, reunindo o algoritmo e desenvolvendo uma interface gráfica para, em conjunto, formar o produto final da pesquisa que é o software que realizará a contagem automatizada dos trombócitos utilizando a Inteligência Artificial como principal ferramenta para isso.

Por fim, planeja-se realizar testes com o software e analisar seus resultados definindo assim sua acurácia identificando também possíveis melhorias.

4.2 Materiais e Ferramentas

Para a execução deste projeto, uma série de materiais e ferramentas tecnológicas foram essenciais, permitindo a implementação eficaz da solução proposta com IA e Visão Computacional. A seguir, detalhamos os principais itens utilizados:

- **Linguagem de Programação Python:** A escolha de Python como linguagem principal se deve à sua ampla utilização em projetos de IA e processamento de imagens. Sua simplicidade e extensiva coleção de bibliotecas otimizadas para *machine learning* foram cruciais para a

viabilidade do projeto.

- **Ambientes de Desenvolvimento:** Para facilitar o desenvolvimento e a execução do código, utilizamos diferentes ambientes:
 - **Visual Studio Code:** Utilizado para o desenvolvimento local, oferece uma interface robusta com suporte a diversas extensões que auxiliam na programação em Python.
 - **Google Colab:** Permitiu a execução de notebooks na nuvem, facilitando o treinamento de modelos com maior poder computacional sem a necessidade de recursos locais avançados.
 - **Jupyter Notebook:** Amplamente utilizado para desenvolvimento de IA, este ambiente permite visualizações interativas e a prototipação rápida de código.
- **Arquiteturas de Detecção de Objetos:** YOLO (*You Only Look Once*), uma das principais arquiteturas usadas para detecção de objetos, possibilitando a identificação rápida e eficiente de plaquetas em imagens microscópicas. Sua velocidade e precisão são ideais para análises que exigem agilidade. OpenCV (*Open Source Computer Vision Library*), biblioteca fundamental para o processamento de imagens, usada para manipular e preparar as imagens de células sanguíneas antes de serem analisadas pelo modelo de IA.

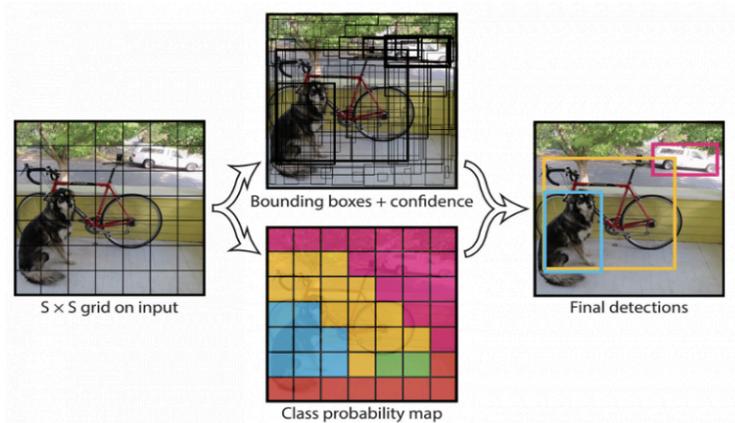


Figura II: Esquema de detecção YOLO (“Deep Learning:Detecção de Objetos em Imagens”, [s.d.]

- **Bibliotecas e Componentes do Python:** O uso de diversas bibliotecas foi indispensável para a construção e treinamento do modelo de IA:
 - **Keras:** Uma API de alto nível que facilita a construção de redes neurais, permitindo prototipação rápida e resultados eficientes.
 - **NumPy, Pandas e SciPy:** Utilizadas para manipulação e análise de dados, essas bibliotecas permitem o processamento eficiente de grandes volumes de dados, como



imagens e informações relacionadas às plaquetas.

- **Tensor Flow:** A base para o treinamento e deploy do modelo de deep learning, sendo fundamental para o desenvolvimento de redes neurais convolucionais (CNNs) que analisam as imagens de plaquetas.
- **Matplotlib:** Utilizada para visualização gráfica dos resultados, facilitando a análise do desempenho do modelo e a interpretação de dados de treinamento.
- **Scikit-Learn:** Biblioteca essencial para a implementação de algoritmos de machine learning e para as fases de pré-processamento e avaliação do modelo.
- **Sites para Levantamento Bibliográfico:** Para fundamentar o embasamento teórico do projeto, foram consultados:
 - **Google Acadêmico:** Amplamente utilizado para encontrar artigos acadêmicos, estudos de caso e referências científicas.
 - **SciELO:** Uma importante base de dados de artigos e publicações científicas, essencial para validar métodos e técnicas aplicadas em áreas de saúde e IA.
- **Dataset de Imagens de Células Sanguíneas:** O dataset de imagens desempenhou um papel central no treinamento do modelo de IA, fornecendo os dados necessários para que o sistema aprendesse a identificar e contar as plaquetas de forma autônoma e precisa. Para garantir a diversidade e a robustez dos dados, foram utilizados três conjuntos principais de imagens, obtidos de fontes renomadas e acessíveis.

O primeiro dataset utilizado foi o BCCD Dataset, que contém mais de 300 imagens rotuladas de células sanguíneas. Essas imagens foram essenciais nas fases iniciais do treinamento, permitindo ao modelo reconhecer características básicas das plaquetas e de outras células sanguíneas em amostras microscópicas. O fato de o dataset ser bem organizado e já rotulado facilitou o processo de adaptação do modelo para detecção automática.

Além disso, utilizamos um conjunto especializado proveniente do Cancer Moonshot Biobank, intitulado Cancer Moonshot Biobank – Acute Myeloid Leukemia (CMB-AML). Esse dataset, acessado através do repositório The Cancer Imaging Archive, inclui 45 lâminas de esfregaços sanguíneos de pacientes com leucemia mieloide aguda (LMA). Este conjunto de dados é especialmente relevante para o projeto, pois oferece imagens reais de casos clínicos, permitindo que o modelo seja treinado em condições que simulem a prática médica. Por fim, um terceiro dataset foi obtido através da plataforma Kaggle, o Blood Cells Image Dataset, que



contém mais de 10.000 imagens de células sanguíneas. Este extenso conjunto de dados foi fundamental para validar o modelo em larga escala, garantindo que ele fosse exposto a uma variedade maior de casos e condições, aumentando a robustez e a generalização dos resultados.

4.4 Barreiras e Limitações

Durante o desenvolvimento do modelo de detecção de plaquetas, algumas barreiras e limitações foram identificadas. Primeiramente a disponibilidade de dados rotulados de qualidade apresentou-se como um desafio, visto que imagens médicas específicas, como as de plaquetas, são escassas e, muitas vezes, carecem da diversidade necessária para garantir uma generalização adequada do modelo.

Ademais, o processo de detecção em tempo real pode ser prejudicado pela capacidade computacional disponível. Embora o modelo YOLO seja eficiente, ele exige hardware de alto desempenho para realizar previsões, o que é essencial em aplicações clínicas que demandam alta precisão e baixa latência.

Por fim, a captura das imagens, principalmente em locais de difícil acesso, também representa uma dificuldade. Essa limitação afeta a qualidade das imagens, comprometendo a capacidade do modelo de detectar plaquetas com precisão.



5 DESENVOLVIMENTO DO MODELO DE IA

5.1 Arquitetura do Modelo

O modelo desenvolvido para a detecção de plaquetas é baseado na arquitetura YOLO, uma abordagem inovadora para a detecção de objetos em imagens. Essa técnica permite que o modelo identifique e localize objetos em tempo real, o que é fundamental para aplicações que exigem precisão e rapidez, como a análise de imagens médicas.

Inicialmente, a configuração do modelo é definida em um arquivo YAML, que especifica os parâmetros essenciais para o treinamento. Nesse arquivo, o caminho para o conjunto de dados é indicado, permitindo que o modelo acesse as imagens necessárias para aprender a identificar plaquetas. O conjunto de dados é dividido em duas partes: uma para treinamento e outra para validação. Dessa maneira, o modelo é capaz de evitar a memorização excessiva das imagens de treinamento e, ao mesmo tempo, generaliza suas habilidades para novos dados.

Além disso, o arquivo YAML também define o número de classes que o modelo deve reconhecer — neste caso, uma única classe: as plaquetas. Com isso, o processo de treinamento é simplificado, pois o modelo se concentra em aprender as características específicas dessa classe. Ademais, parâmetros como o número de épocas e o tamanho do lote são configurados, influenciando a eficiência e a eficácia do treinamento. A escolha desses parâmetros é crítica para equilibrar a precisão e o tempo de processamento.

Durante a execução, o modelo carrega os pesos previamente treinados, possibilitando a detecção em novas imagens. Assim que uma imagem é fornecida, o modelo processa os dados e retorna as detecções, indicando a localização e a confiança de cada plaqueta identificada. Essa abordagem melhora a precisão da detecção e proporciona um feedback rápido ao usuário, o que é essencial em cenários clínicos.

A integração da arquitetura YOLO com uma configuração detalhada no arquivo YAML resulta em um sistema robusto e eficiente para a detecção de plaquetas. Consequentemente, essa solução facilita a análise de imagens médicas, ao mesmo tempo que contribui para avanços no diagnóstico e pesquisa.



5.2 Treinamento e Validação

O treinamento do modelo de detecção de plaquetas foi realizado com um conjunto de dados meticulosamente organizado, composto por imagens rotuladas de plaquetas. A estrutura dos dados, portanto, é crucial para o sucesso do treinamento. Dessa forma, o conjunto foi dividido em duas partes principais: um conjunto de treinamento e outro de validação.

O conjunto de treinamento, por sua vez, é utilizado para ensinar o modelo a reconhecer e localizar plaquetas. Esse conjunto contém uma variedade de imagens que capturam diferentes condições e contextos em que as plaquetas podem aparecer. Essa diversidade é fundamental, pois permite que o modelo aprenda a generalizar suas habilidades em vez de apenas memorizar as imagens. Para garantir o acesso a esses dados, o arquivo de configuração YAML especifica o caminho para as imagens, organizando o processo de treinamento de forma eficiente.

Por outro lado, o conjunto de validação é composto por imagens que o modelo não viu durante o treinamento. Esse conjunto é utilizado para avaliar o desempenho do modelo em dados novos e não rotulados. A avaliação contínua com o conjunto de validação é essencial para monitorar a capacidade do modelo de generalizar o que aprendeu, ajudando a evitar o *overfitting* — situação em que o modelo se ajusta excessivamente aos dados de treinamento e falha em reconhecer novas imagens.

Além disso, a organização dos dados no arquivo YAML facilita o acesso e a utilização durante o treinamento. Nele, o caminho raiz para o conjunto de dados é especificado, juntamente com os caminhos relativos para as pastas de treinamento e validação. Essa estrutura permite que o modelo acesse rapidamente as imagens necessárias para cada etapa do processo.

Durante o treinamento, o modelo utiliza algoritmos de aprendizado de máquina para ajustar seus pesos internos com base nas imagens fornecidas. Ele analisa as características visuais das plaquetas e aprende a identificar padrões que as diferenciam de outros elementos presentes nas imagens. A cada época, o desempenho do modelo é avaliado com o conjunto de validação, o que possibilita ajustes contínuos e melhorias em sua precisão.

Após o término do treinamento, o modelo é testado em um conjunto de imagens não vistas anteriormente, a fim de verificar sua capacidade de detecção. Essa fase final é crucial, pois garante que o modelo não apenas aprendeu a reconhecer plaquetas, mas também pode aplicar esse conhecimento em situações do mundo real. Como resultado, obtém-se um modelo robusto, capaz de detectar plaquetas com alta precisão, contribuindo significativamente para a análise de imagens médicas e para o avanço das pesquisas na área.

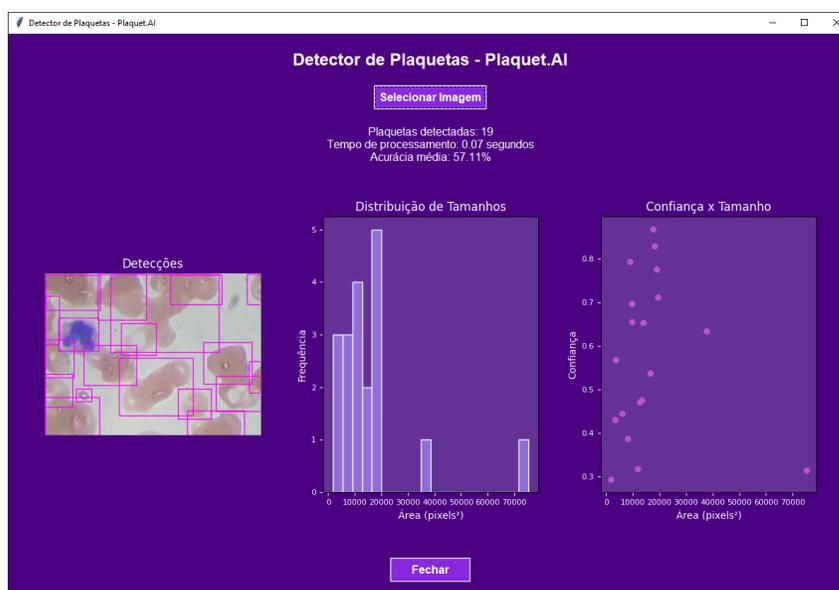


6 RESULTADOS E CAMINHOS FUTUROS

O protótipo do modelo desenvolvido foi treinado com um conjunto significativo de dados, incluindo 364 imagens rotuladas, mais de 10.000 imagens não rotuladas para validação, bem como 45 lâminas de alta resolução de esfregaços sanguíneos. Apesar disso, o modelo alcançou uma precisão entre 50% e 60%, o que indica ainda uma margem de erro na detecção de plaquetas. Entretanto, a execução dos processos ocorre de maneira eficiente, visto que o tempo de processamento é igual ou inferior a um segundo. Portanto, embora o desempenho do modelo ainda precise ser otimizado, sua capacidade de processamento rápido pode ser um diferencial importante para futuras versões.

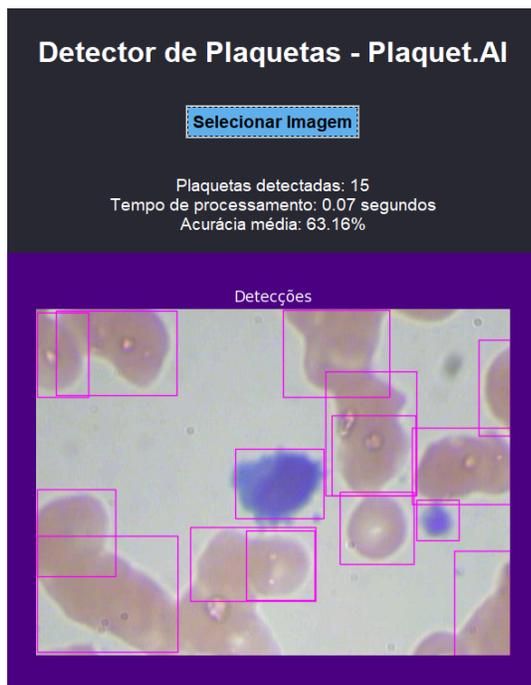
Os resultados obtidos evidenciam que o modelo necessita de ajustes para melhoria da acurácia. Nesse sentido, uma possibilidade é aumentar o conjunto de imagens rotuladas, enriquecendo o processo de aprendizado do modelo. Além disso, técnicas de balanceamento e refinamento de dados não rotulados, como o uso de pseudo-rotulação ou aprendizado semi-supervisionado, podem ser exploradas. Outro caminho seria a implementação de arquiteturas mais complexas, com o objetivo de aumentar a precisão sem comprometer o tempo de resposta.

Figura III: interface do modelo treinado com 15 épocas.



Fonte: Elaborado pelos autores.

Figura IV: interface do modelo treinado com 30 épocas.



Fonte: Elaborado pelos autores.

7 CONCLUSÃO

Este trabalho desenvolveu um modelo de IA para a contagem automatizada de plaquetas, buscando otimizar o processo em regiões remotas e para populações vulneráveis. A hipótese de criar uma solução rápida foi confirmada, visto que o modelo realiza as detecções em até um segundo. Entretanto, a acurácia de 50% a 60% indica que o desempenho precisa ser aprimorado para garantir maior confiabilidade nos diagnósticos.

Portanto, a principal necessidade para melhorias está na ampliação do conjunto de dados rotulados, além da implementação de técnicas como aprendizado semi-supervisionado e o ajuste dos parâmetros do modelo. Conclui-se que, o sistema é promissor, mas ainda requer ajustes para atingir o nível de precisão necessário para uso clínico seguro e eficaz, principalmente em áreas de difícil acesso.



8 REFERÊNCIAS

ANTONELLO, Ricardo . Introdução a Visão Computacional com Python e OpenCV. Luzerna: Instituto Federal Catarinense, 2017.

ÁVILA, F.; OLIVEIRA, L. S. DE. Comparação da contagem de plaquetas entre o método de Fônio e automação. Atas de Ciências da Saúde (ISSN 2448-3753), v. 3, n. 3, 2015.

BCCD Dataset [Dataset]. Dataset Ninja. <https://datasetninja.com/bccd>.

Cancer Moonshot Biobank. (2023). Cancer Moonshot Biobank – Acute Myeloid Leukemia (CMB-AML) (Version 4) [Dataset]. The Cancer Imaging Archive. <https://doi.org/10.7937/PCTE-6M66>

CAROLINA, A. et al. A Methodology for Determining Hematimetric Indices Using High Accuracy Algorithms. [s.l: s.n.]. Disponível em: <<https://lcv.fee.unicamp.br/images/BTSym-17/Papers/68211.pdf>>. Acesso em: 21 jun. 2024.

COMAR, Samuel R. ; DANCHURA , Heloisa S. M. ; SILVA, Paulo H. . Contagem de plaquetas: avaliação de metodologias manuais e aplicação na rotina laboratorial. Curitiba: Associação Brasileira de Hematologia e Hemoterapia, 2009.

DANIELTEOFILO2004. KDD – Knowlegde Discovery in Database. Disponível em: <<https://danielteofilo.wordpress.com/2015/02/16/kdd-knowlegde-discovery-in-database/>>.

Deep Learning: Detecção de Objetos em Imagens. Disponível em: <<https://lapix.ufsc.br/ensino/visao/visao-computacionaldeep-learning/deteccao-de-objetos-em-imagens/>>.

FARIAS, Mariela Granero; CASTRO , Simone Martins . Diagnóstico Laboratorial das leucemias linfóides agudas. 2. ed. Rio de Janeiro: Jornal brasileiro de patologia e medicina laboratorial, 2003. 91-98 p. v. 40.

GONÇALVES, Thaiana. Análise comparativa das metodologias de contagem manual e automatizada de plaquetas em plaquetopenias, 2019.

LEITE, C. A. L.; NILSON JUNIOR, S. M.; MIRANDA, S. M. Comparação entre a contagem de plaquetas pelos métodos manual e automatizado. NewsLab, v. 81, p. 106- 114, 2007.

MARENGONI, M.; STRINGHINI, S. Tutorial: Introdução à Visão Computacional usando OpenCV. Revista de Informática Teórica e Aplicada, [S. l.], v. 16, n. 1, p. 125–160, 2010. DOI: 10.22456/2175-2745.11477. Disponível em: https://seer.ufrgs.br/index.php/rita/article/view/rita_v16_n1_p125. Acesso em: 31 jan. 2024.

MONTALVÃO, Raquel. Municípios Doadores alerta para baixa nos estoques de sangue, em período de férias. Confederação Nacional de Municípios, 2019. Disponível em: <https://www.cnm.org.br/comunicacao/noticias/municipios-doadores-alerta-para-baixa-nos-estoques-d-e-sangue-em-periodo-de-ferias>. Acesso em: 31 jan. 2024.



MOURA, K. Ciclo de vida dos dados #1. Disponível em:
<<https://medium.com/@kvmoura/kdd-process-9b8e3062142>>.

SCHWEIRGERT, Augusto ; REZENDE, Fábio Hosoi ; FANTONI, Denise Tabacchi; MOROZ, Ludmila Rodrigues. Avaliação da contagem plaquetária pelo contador automático QBC Vet Autoread® comparado com estimativa em esfregaço sanguíneo e contagem em hemocítômetro. 4. ed. Londrina: Semina: Ciências Agrárias, 2010. 1001-1008 p. v. 27.

SEVERINO, Antonio Joaquim. Metodologia do Trabalho Científico. 23ª ed. Cortez Editora, 2007.

SILVA, Drielly Fernandes . Leucemia Mielóide Aguda. São José do Rio Preto: Academia de Ciência e Tecnologia, 2012.

SILVA, Matusalem Dalla Lana Da ; NARDIN, Jeanine Marie . AVALIAÇÃO MANUAL DO NÚMERO DE PLAQUETAS EM PACIENTES ONCOLÓGICOS QUE APRESENTEM TROMBOCITOPENIA. Curitiba: Cadernos da Escola de Saúde (ISSN 1984-7041), 2011. 60-73 p.

UNCLESAMULUS. (2023). Blood Cells Image Dataset [Dataset]. Kaggle.
<https://www.kaggle.com/datasets/unclesamulus/blood-cells-image-dataset>