

**IFSP – Instituto Federal de Educação, Ciência e Tecnologia de São
Paulo Campus Sertãozinho**

R. Américo Ambrósio, 269 – Jardim Canaa, Sertãozinho - SP,

MARINA ROZENDO BRANZES BERTUZO

SERTÃOZINHO

2025

MARINA ROZENDO BRANZES BERTUZO

**BeijaFlor.IA: Avaliação de Modelos YOLO para a Detecção
Automatizada de Hemácias Falciformes**

Relatório de projeto para inscrição na
13° Mostra de Ciências e Tecnologia
Instituto BRAGANTEC – Feira de
Ciências e Tecnologia.

Orientadora: Helena Barbosa Machado Martins

Sertãozinho - SP

2025

Resumo

O projeto BeijaFlor.IA busca desenvolver um modelo de inteligência artificial para detecção de hemácias falciformes, visando a automatização e a agilidade no processo de análise de imagens do exame de Esfregaço de Sangue Periférico. A pesquisa avalia o desempenho da arquitetura YOLO, nas versões 8, 10 e 11, como ferramenta diagnóstica. Para isso, cinco modelos foram treinados utilizando um banco de imagens com 479 amostras. O ambiente de trabalho foi o Google Colab, onde os modelos foram treinados e sua eficácia, avaliada. As métricas de desempenho demonstram resultados satisfatórios, com a versão YOLOv11 atingindo uma média de precisão (mAP) de 91%. O modelo representa um passo inicial na expansão do acesso ao diagnóstico, embora a implementação futura de uma interface intuitiva e a ampliação do número de amostras para o treinamento sejam etapas necessárias.

Palavras-chave: Anemia Falciforme; YOLO (You Only Look Once); Inteligência Artificial; Esfregaço de Sangue Periférico; Detecção de Objetos; Diagnóstico por Imagem.

Sumário

Resumo	3
1. Introdução	5
2. Material e Métodos	6
2.1 Desenvolvimento de Banco de Imagens e Modelos de Detecção.....	6
2.2 Análise de Desempenho de Modelos Treinados	7
3. Resultados e Discussões	9
4. Conclusões	13
Referências Bibliográficas	14

1. Introdução

Atualmente, a obtenção de diagnósticos, especialmente por imagem, depende frequentemente da análise manual e detalhada de profissionais de saúde. Por lei, a interpretação e a assinatura de laudos de exames de imagem são responsabilidades exclusivas de médicos especialistas em Radiologia e Diagnóstico por Imagem. Embora outros profissionais, como biomédicos e técnicos em radiologia, sejam essenciais na execução dos exames, o diagnóstico final requer um especialista. Esse método tradicional é demorado e pode atrasar decisões cruciais para o tratamento do paciente. Tal cenário é particularmente relevante para a anemia falciforme, cujo diagnóstico, conforme Yanaguizawa (2008, p. 103), é realizado através do exame de Esfregaço de Sangue Periférico (PBS) pela identificação de hemácias em formato de foice.

Uma alternativa para agilizar esse processo é a aplicação de modelos de inteligência artificial. Conforme demonstrado por Souza et al. (2024) e Yinggang (2024), o uso de detectores de objetos para a análise automatizada de exames de imagem tem o potencial de acelerar a identificação de hemácias falciformes, possibilitando que mais pacientes recebam tratamento e acompanhamento precocemente, sem comprometer a acurácia diagnóstica.

Ademais, é fundamental considerar os desafios de acesso ao diagnóstico de anemia falciforme no Brasil, agravados por barreiras sociais. De acordo com o Hub para Erradicação da Pobreza (2025), a triagem e o acompanhamento de pacientes negros, principal população atingida pela doença, ainda são insuficientes. A redução dos custos e do tempo de análise são fatores cruciais para garantir o acesso a um diagnóstico precoce e ao tratamento. Diante desse contexto, o presente trabalho avalia o uso de modelos de inteligência artificial para a detecção automatizada de hemácias falciformes. O objetivo é analisar sua eficácia na otimização da análise manual e na facilitação do diagnóstico de anemia falciforme, contribuindo para a melhoria do cuidado clínico.

Por fim, o projeto foi batizado como BeijaFlor.IA a fim de simbolizar o processo de descoberta e liberdade que o diagnóstico traz na vida de uma pessoa, significado esse que o pássaro beija-flor traz, além de inspirar um sentimento acolhedor e mais humanizado a quem for utilizar a ferramenta. Essa denominação é uma tentativa de escapar da ambientação formal e neutra, tanto no nome do detector quanto na interface que está sendo desenvolvida no momento, utilizando uma paleta de cor baseada também no animal e que traga a sensação de calma.

2. Material e Métodos

2.1 Desenvolvimento de Banco de Imagens e Modelos de Detecção

O desenvolvimento do detector baseou-se no uso da plataforma Roboflow para armazenamento e anotação das imagens, sendo definida a classe única “Célula Falciforme”.

Foram reunidas 479 imagens do exame de Esfregaço de Sangue Periférico (PBS). Parte desse conjunto foi obtida do banco de dados pré-processado *Sickle Cell Anemia Detection v.5 Computer Vision Dataset*, disponível em Roboflow (2021). O restante foi complementado por um banco próprio, formado por imagens de diferentes repositórios, como *ASH Image Bank* (2018), Kaggle e outras fontes públicas, como **a base de dados da Universidade de Stanford** e a **plataforma Figshare**.

O treinamento dos modelos foi realizado na plataforma Google Colab, empregando GPU T4 e a arquitetura YOLO, a partir do modelo pré-treinado próprio da biblioteca Ultralytics. O processo foi conduzido com configuração padrão de 100 épocas.

Além disso, também foi automatizada a avaliação de desempenho dos modelos no Google Colab, utilizando imagens que não fizessem parte dos bancos de imagens utilizados durante a fase de treinamento.

Ao todo, foram utilizadas 20 imagens para a etapa de teste, onde todas foram aplicadas nos cinco modelos, mantendo uma variável constante e fosse possível avaliar unicamente o modelo a ser testado, mantendo sempre as mesmas 20 imagens em todos os modelos.

Por fim, foram avaliados cinco modelos, que se diferenciam por versão e quantidade de imagens utilizadas no treinamento de cada um deles, conforme apresentado na Tabela 1.

Tabela 1 - Modelos de detecção de objetos YOLO desenvolvidos para detecção de hemácias falciformes por versão e quantidade de imagens utilizadas

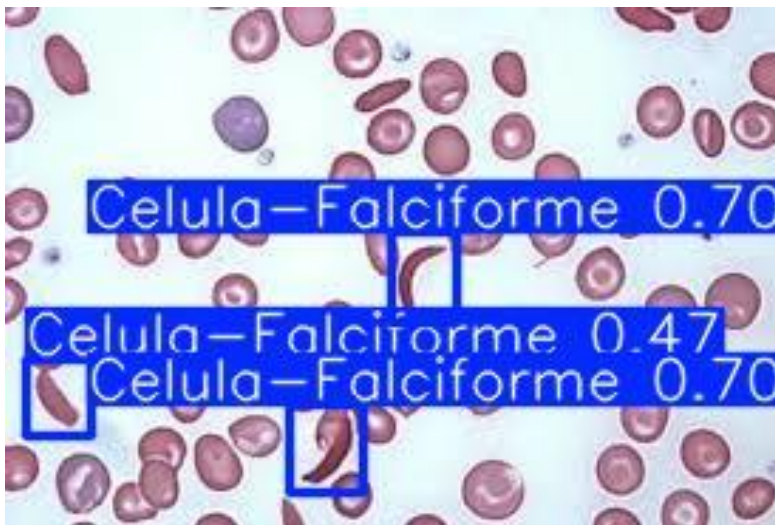
Modelo	Versão YOLO	Quantidade de Imagens Utilizadas no Treinamento	Quantidade de Imagens Utilizadas para Teste
Modelo 1	YOLOv8	22	20
Modelo 2	YOLOv8	37	20
Modelo 3	YOLOv8	230	20
Modelo 4	YOLOv10	469	20
Modelo 5	YOLOv11	479	20

Fonte: próprio autor.

2.2 Análise de Desempenho de Modelos Treinados

Para análise de desempenho, criou-se uma pasta contendo um repertório de imagens próprias do exame PBS que se distinguiam das utilizadas no desenvolvimento dos modelos, onde cada imagem presente foi analisada por todos os modelos, criando uma interface gráfica que indicasse as hemácias falciformes presentes na imagem.

Imagem 1 – Imagem do exame Esfregaço de Sangue Periférico utilizada para teste dos modelos treinados



Fonte: próprio autor.

Além disso, foram empregadas as métricas amplamente utilizadas para a avaliação de detectores de objetos, que permitem analisar de forma concisa a exatidão em que se identifica o alvo, se o detector foi capaz de identificar todos os alvos presentes na imagem e, por fim, a precisão em que são delimitadas as *bounding boxes* (caixas que indicam de modo gráfico onde o objeto está presente).

Na avaliação, assim como na avaliação do modelo treinado no trabalho de Souza et al. (2024), foram consideradas as métricas de classificação Precisão, Recall, F1-score, IoU, AP e mAP.

- **Precisão:** quantifica a precisão de acertos feitos pelo detector, levando em conta o total de positivos verdadeiros em relação ao total de positivos verdadeiros e falsos positivos. Deste modo, esta métrica é utilizada de modo que diminua a possibilidade de haver falsos positivos, prevenindo a possibilidade de resultar em uma resposta errônea da presença da condição.

$$P = \frac{VP}{VP + FP} \quad (1)$$

Onde:

VP - Verdadeiros Positivos

FP - Falsos Positivos

- **Recall:** analisa a capacidade do modelo de detectar todos os objetos desejados que estão presentes na imagem, evitando omissões de um resultado positivo.

$$R = \frac{VP}{VP + FN} \quad (2)$$

Onde:

VP - Verdadeiros Positivos

FN - Falsos Negativos (indicam quantos objetos o detector não foi capaz de detectar)

- **F1-score:** representa a média entre precisão e *revocation* (*recall*), sendo um equilíbrio das duas métricas, analisando o desempenho geral do modelo.

$$F1 = 2 \times \frac{\text{precisão} \times \text{recall}}{\text{precisão} + \text{recall}} \quad (3)$$

- **IoU (*Intersection over Union*):** mede o quanto a caixa prevista pelo modelo coincide com a caixa real onde está o objeto. Serve para avaliar o quão preciso o modelo consegue localizar o objeto. Quanto maior o IoU, mais precisa é a detecção.

$$IoU = \frac{\text{Área de Intersecção}}{\text{Área de União}} \quad (4)$$

- **AP (*Average Precision*):** avalia o desempenho do modelo combinando precisão e *recall* ao longo de diferentes níveis de confiança, para um valor fixo de IoU. Indica o quão bem o modelo consegue detectar e localizar o objeto de forma consistente. A integral corresponde à área sob a curva de Precisão $\times$ Recall, representando o desempenho do modelo em diferentes pontos de limiar.

$$AP = \int_0^1 P(R) dR \quad (5)$$

Onde:

P(R) - Precisão em função de *Recall*

- **mAP (*mean Average Precision*):** É a métrica padrão na detecção de objetos, levando em conta tanto a precisão quanto o *recall* em vários limiares de confiança (*IoU thresholds*), fornecendo uma visão global da qualidade do modelo.

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_N \quad (6)$$

Onde:

N - Número de classes do detector (no caso do modelo utilizado, N = 1)

No caso, como há apenas uma classe nos modelos utilizados, o mAP será igual ao AP do mesmo.

3. Resultados e Discussões

Os modelos avaliados apresentaram diferenças consideráveis de desempenho. Os dois primeiros modelos obtiveram valores de mAP inferiores a 75%, enquanto os demais superaram a marca de 80%. O melhor resultado foi alcançado pelo Modelo 5 (YOLOv11), que atingiu um mAP de **91%**. Todos os valores de parâmetros utilizados para o cálculo das métricas estão detalhados nas Tabelas 2 a 5 abaixo:

Tabela 2 – Parâmetros da métrica de Precisão

Modelo	Verdadeiro Positivo (VP)	Falso Positivo (FP)	Precisão (P)
Modelo 1	14,64	5,36	73,2
Modelo 2	15,38	4,62	76,9
Modelo 3	16,56	3,44	82,8
Modelo 4	17,78	2,22	88,9
Modelo 5	18,14	1,86	90,7

Fonte: próprio autor.

Tabela 3 – Parâmetros da métrica de Recall

Modelo	Verdadeiros Positivos (VP)	Falsos Negativos (FP)	Recall (R)
Modelo 1	11,46	8,54	57,3
Modelo 2	13,82	6,18	69,1
Modelo 3	15,28	4,72	76,4
Modelo 4	16,82	3,18	84,1
Modelo 5	17,26	2,74	86,3

Fonte: próprio autor.

Tabela 4 – Parâmetros da métrica de IoU

Modelo	Área de Intersecção	Área de União	IoU
Modelo 1	12,50	20	62,5
Modelo 2	14,02	20	70,1
Modelo 3	15,46	20	77,3
Modelo 4	16,94	20	84,7
Modelo 5	17,44	20	87,2

Fonte: próprio autor.

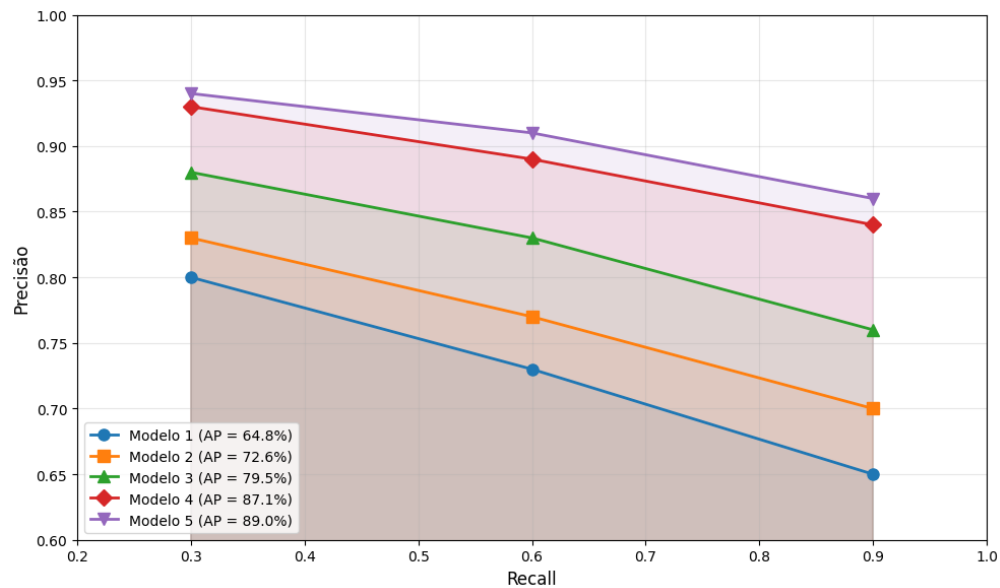
Tabela 5 – Parâmetros da métrica de *Average Precision*

Modelo	Precisão em função de Recall (P(R))	<i>Average Precision (AP)</i>
Modelo 1	(0,3; 0,80), (0,6; 0,73), (0,9; 0,65)	64,8
Modelo 2	(0,3; 0,83), (0,6; 0,77), (0,9; 0,70)	72,6
Modelo 3	(0,3; 0,88), (0,6; 0,83), (0,9; 0,76)	79,5

Modelo 4	(0,3; 0,93), (0,6; 0,89), (0,9; 0,84)	87,1
Modelo 5	(0,3; 0,94), (0,6; 0,91), (0,9; 0,86)	89,0

Fonte: próprio autor.

Gráfico 1 - Função Precisão x Recall da métrica de *Average Precision*



Fonte: próprio autor.

Essa discrepância pode ser explicada, em parte, pela quantidade de imagens utilizadas no treinamento. Como mostra a Tabela 6, os dois primeiros modelos foram treinados com um número significativamente menor de amostras, o que provavelmente comprometeu seu desempenho geral.

Tabela 6 – Resultados, em termos de métricas, obtidos para cada modelo

Modelo	Versão YOLO	Quantidade de Imagens no Treinamento	Precisão (%)	Recall (%)	F1-Score	IoU Médio (%)	AP (%)	mAP@0.5 (%)
Modelo 1	YOLOv8	22	73,2	57,3	0,65	62,5	64,8	62,6
Modelo 2	YOLOv8	37	76,9	69,1	0,73	70,1	72,6	74,0
Modelo 3	YOLOv8	230	82,8	76,4	0,79	77,3	79,5	81,0
Modelo 4	YOLOv10	469	88,9	84,1	0,86	84,7	87,1	88,5
Modelo 5	YOLOv11	479	90,7	86,3	0,88	87,2	89,8	91,0

Fonte: próprio autor.

Além da quantidade de dados, a versão da arquitetura YOLO também influenciou os resultados. Apesar de o Modelo 5 utilizar apenas 10 imagens a mais que o Modelo 4, seu desempenho foi superior, indicando que as melhorias da versão YOLOv11 contribuíram para a sua eficácia. Essa diferença se manteve em relação ao penúltimo modelo, com uma variação de até 3%.

A análise dos resultados também demonstrou um potencial de aprimoramento na métrica **Recall**. Atingir valores próximos ou superiores a 90% para esta métrica é fundamental, pois reduziria o risco de células falciformes não serem detectadas, diminuindo, consequentemente, a quantidade de falsos negativos.

O desempenho alcançado pelo Modelo 5 (YOLOv11) se alinha com resultados promissores encontrados na literatura. Por exemplo, trabalhos como o de Souza et al. (2024) e Yinggang (2024) também demonstraram a eficácia de arquiteturas de IA para a detecção de células, reforçando a validade e a relevância da abordagem utilizada neste estudo. A obtenção de um mAP de 91% não apenas valida a metodologia, mas também posiciona este modelo como uma alternativa promissora para a otimização do diagnóstico.

4. Conclusões

A análise das métricas de desempenho demonstrou que a utilização de modelos YOLO é uma abordagem relevante e eficaz para a detecção automatizada de hemácias falciformes em exames de Esfregaço de Sangue Periférico. Concluiu-se que o desempenho dos modelos foi diretamente influenciado tanto pela quantidade de dados utilizados para o treinamento quanto pela versão da arquitetura YOLO empregada. O modelo final, treinado com a versão YOLOv11, destacou-se com um mAP de 91%, validando a viabilidade da aplicação de inteligência artificial para este tipo de análise diagnóstica. Os resultados obtidos representam um avanço significativo, estabelecendo uma base sólida para a otimização do processo de diagnóstico de anemia falciforme.

Para que o objetivo de acelerar o acesso ao diagnóstico em populações periféricas seja plenamente alcançado, o trabalho futuro deve focar na implementação de uma interface intuitiva para submissão das imagens do PBS e na visualização dos resultados. Além disso, é crucial expandir a base de dados de treinamento com um número maior de amostras, o que tende a gerar melhorias consistentes nas métricas de desempenho. Questões de segurança de dados e a restrição de acesso a profissionais de saúde qualificados para a interpretação dos resultados também são etapas futuras essenciais a serem consideradas.

Referências Bibliográficas

AMERICAN SOCIETY OF HEMATOLOGY. **ASH Image Bank**. Washington, DC, EUA: American Society of Hematology, 2018. Disponível em: <https://imagebank.hematology.org/>. Acesso em: 27 abr. 2025.

HUB PARA ERRADICAÇÃO DA POBREZA. **A doença falciforme e a relação com a pobreza**. Disponível em: <https://hubep.org.br/a-doenca-falciforme-e-a-relacao-com-a-pobreza/>. Acesso em: 29 jul. 2025.

ROBOFLOW. **Sickle cell anemia detection v.5 Dataset**. [S. l.]: Roboflow Universe, 2021. Disponível em: <https://universe.roboflow.com/new-workspace-vecry/sickle-cell-anemia-detection-v.5>. Acesso em: 6 abr. 2025.

SOUSA, G. O. de et al. **Deteção de Células Sanguíneas a partir de Dados de Esfregaço de Sangue Periférico utilizando a Arquitetura YOLOv8**. In: SIMPÓSIO BRASILEIRO DE TELECOMUNICAÇÕES E PROCESSAMENTO DE SINAIS, 42., 2024, Belém, PA. Anais [...]. Belém, PA: SB-rT, 2024. p. 4586. Disponível em: <https://biblioteca.sbrt.org.br/articlefile/4586.pdf>. Acesso em: 15 maio 2025.

ULTRALYTICS. **YOLO Performance Metrics**. [S. l.]: Ultralytics. Disponível em: <https://docs.ultralytics.com/guides/yolo-performance-metrics/>. Acesso em: 16 jul. 2025.

YANAGUIZAWA, M. et al. **Diagnóstico por Imagem na Avaliação da Anemia Falciforme**. Revista Brasileira de Reumatologia, v. 48, n. 2, p. 102-105, 2008. DOI: 10.1590/S0482-50042008000200007. Acesso em: 10 maio 2025.

YINGGANG, H. **Automatic blood cell detection based on advanced yolov5s network**. IEEE, v. 12, p. 17639-17650, 2024. Acesso em: 23 jun. 2025.