



**INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO PIAUÍ
CAMPUS FLORIANO
CURSO ANÁLISE E DESENVOLVIMENTO DE SISTEMA**

Iago Amorim Avelino

**IMPLEMENTAÇÃO DE UM MODELO DE DEEP LEARNING PARA DETECÇÃO DE
RETINOPATIA DIABÉTICA APLICADO EM UMA BASE DE IMAGENS DE RETINA**

FLORIANO

2025

IAGO AMORIM AVELINO

**IMPLEMENTAÇÃO DE UM MODELO DE DEEP LEARNING PARA DETECÇÃO DE
RETINOPATIA DIABÉTICA APLICADO EM UMA BASE DE IMAGENS DE RETINA**

Trabalho de Conclusão de Curso (artigo científico) apresentado como exigência parcial para obtenção do diploma do Curso de Tecnólogo em Análise e Desenvolvimento de Sistemas do Instituto Federal de Educação, Ciência e Tecnologia do Piauí, Campus Floriano.

Orientador: Prof. Dr. Willamys Rangel
Nunes de Sousa

Co-orientador: José Joaquim de Moura
Ramos

FLORIANO

2025

IMPLEMENTAÇÃO DE UM MODELO DE DEEP LEARNING PARA DETECÇÃO DE RETINOPATIA DIABÉTICA APLICADO EM UMA BASE DE IMAGENS DE RETINA

Iago Amorim Avelino¹
Willamys Rangel Nunes de Sousa²
José Joaquim de Moura Ramos³

RESUMO

O seguinte estudo baseia-se no desenvolvimento de técnicas automatizadas que sejam capazes de identificar a presença de retinopatia diabética em imagens de retina. Para isso foi utilizado de técnicas de aprendizado profundo, um subcampo de estudo da área de inteligência artificial, onde foram treinados 5 (cinco) modelos utilizando a arquitetura de rede neural convolucional *ResNet18*. Desses cinco modelos, apenas um focava em detectar olhos doentes em qualquer estágio da doença, enquanto os outros quatro foram treinados para identificar em cada estágio da doença. Ao final do treinamento dos modelos que detectaram a doença por estágio foi montado uma *pipeline*, com os quatro modelos que detectasse por estágio, para que pudesse identificar se a imagem de retina apresentava ou não sinais da doença. Todos os modelos apresentaram uma acurácia acima de 90%, entretanto, o modelo que comparava olhos saudáveis com olhos com o estágio inicial da doença, a especificidade desse modelo foi inferior aos outros, indicando poucas presenças de características da doença na imagem. Para trabalhos futuros fica-se a sugestão de aumento do conjunto de dados e treinar modelos com outras arquiteturas no mesmo conjunto.

Palavras-chave: retinopatia diabética; redes neurais convolucionais; aprendizado profundo; diabetes mellitus.

ABSTRACT

The following study is based on the development of automated techniques that are capable of identifying the presence of diabetic retinopathy in retinal images. For this, deep learning was used, a subfield of study in the area of artificial intelligence, where 5 (five) models were trained using the ResNet18 convolutional neural network architecture. Of these five models, only one focused on detecting diseased eyes at any stage of the disease, while the other four were trained to identify at each stage of the disease. At the end of training the models that detected the disease by stage, a pipeline was set up, with the four models that detected by stage, so that it could identify whether or not the retinal image showed signs of the disease. All models showed an accuracy above 90%, however, the model that compared healthy eyes with eyes with the initial stage of the disease, the specificity of this model was lower than the others, indicating few presence of disease characteristics in the image. For future work, we suggest increasing the data set and training models with other architectures on the same set.

Keywords: diabetic retinopathy; convolutional neural networks; deep learning; diabetes mellitus.

¹ Graduando em Tecnólogo em Análise e Desenvolvimento de Sistemas. Instituto Federal do Piauí. caflo.2023114tads0011@aluno.ifpi.edu.br.

² Doutor em Geologia (UFC). Instituto Federal do Piauí. rangelnunes@ifpi.edu.br.

³ Doutor em Computação (UDC). Universidade da Coruña. joaquim.demoura@udc.es

Data de aprovação: 09/07/2025 (Data de apresentação)

1 INTRODUÇÃO

Segundo a Organização Mundial de Saúde (OMS), em 2018, existiam mais de 422 milhões de pessoas com a *diabetes mellitus* (Casarin *et al.*, 2022). O Brasil ocupa o 6º lugar entre os países do mundo com mais pessoas diabéticas, sendo que uma em cada três pessoas que possuem diabetes, não sabem que têm diabetes (Sociedade Brasileira de Diabetes, 2024).

Uma das consequências da *diabetes mellitus* é a retinopatia diabética, uma das principais causas de comorbidade em pacientes com a doença, e, uma das maiores causas de perda de visão no mundo, afetando um em cada três diabéticos (Bortoli *et al.*, 2022; Franco *et al.*, 2022). Por a diabetes ser uma doença sem cura, as suas consequências podem acontecer em qualquer momento da vida.

Portanto, o diagnóstico precoce da retinopatia diabética é fundamental para evitar seu desenvolvimento ou, caso ocorra, para controlá-la e minimizar os danos (Casarin *et al.*, 2022). A retinopatia diabética é dividida em dois tipos principais: a não proliferativa e a proliferativa. A retinopatia diabética não proliferativa, por sua vez, é classificada nos estágios leve, moderado e severo. O diagnóstico é realizado pelo oftalmologista que, em casos com grande volume de pacientes, pode utilizar imagens de fundo de olho feitas por retinografia colorida (Kanadani *et al.*, 2021). Além disso, a carga de trabalho na classificação desses exames pode ser reduzida por meio de inteligência artificial, o que contribui para a diminuição de custos e tempo (Bortoli *et al.*, 2022).

Estudos prévios já aplicaram métodos automáticos para a detecção de retinopatia diabética utilizando aprendizado profundo (Wan, Liang, Zhang, 2018). Sendo comprovado que a utilização de diversos métodos de aprendizado profundo para a detecção da retinopatia em imagens de fundo de olhos tem uma boa eficácia (Tsiknakis *et al.* 2021). Ademais, há utilização de métodos automáticos somente para a identificação e segmentação de lesões da retinopatia diabética em imagens emitidas por retinógrafos (Sánchez *et al.* 2004). Mesmo assim, ainda há poucos métodos automatizados para a identificação e classificação dos estágios da doença (Alyoufi, Shalash, Abulkhair, 2024).

Com base nisso, este estudo tem como objetivo facilitar o diagnóstico da retinopatia diabética em imagens de fundo de olho dos pacientes, através da implementação de um modelo de aprendizado profundo que reconheça padrões característicos da doença. Permitindo, assim, ajudar os oftalmologistas no seu dia a dia a terem um melhor diagnóstico nos exames realizados.

2 REFERENCIAL TEÓRICO

2.1 *Diabetes mellitus*

De acordo com o Ministério da Saúde, a *diabetes mellitus* (DM), ou só diabetes, é uma doença causada pela produção insuficiente ou má absorção de insulina, hormônio responsável pela quebra das moléculas de glicose presentes no sangue, transformando-as em energia para o organismo (Brasil, 2024).

Devido a essa disfunção, a pessoa portadora da doença passa a apresentar um aumento nas taxas de glicemia, que em altas taxas pode causar complicações para outros órgãos, como coração, olhos e rins, e em casos mais extremos, a diabetes pode levar à morte (Casarin *et al.*, 2022).

Segundo a Sociedade Brasileira de Diabetes (2024), 10.2% da população brasileira, aproximadamente 20 milhões de pessoas, possuem diabetes. Dentre eles, 90% apresentam diabetes do *tipo 2*, quando o organismo não consegue usar adequadamente a insulina que produz, ou não produz insulina suficiente para controlar a taxa de glicemia. Enquanto isso, os outros 5% a 10% do total apresentam diabetes do *tipo 1*, que ocorre em indivíduos com predisposição genética, nos quais o sistema imunológico erroneamente ataca as células beta do pâncreas, responsáveis pela produção de insulina.

A DM se perdura por toda a vida e, caso não seja bem acompanhada mantendo um controle dos índices glicêmicos, pode gerar algumas complicações, como retinopatia, nefropatia, neuropatia e, em alguns dos casos, podem estar acompanhadas de doenças crônicas (Casarin *et al.*, 2022).

A retinopatia diabética (RD) é uma microangiopatia que afeta os vasos da retina, os quais são vulneráveis aos elevados níveis glicêmicos (Bastos; Oliveira; Jorge, 2022). Ela pode levar à alterações na visão, desde manchas ou pontos cegos até à cegueira total, sendo que ela é uma das principais causas de cegueira em todo o mundo (Melo *et al.*, 2019).

No Brasil, a incidência da RD é de 24% a 39% da população diabética, sendo estimado que tenha uma prevalência de 2 milhões de casos. Após 20 anos de doença, estima-se que 90% dos diabéticos de tipo 1 (DM1) e 60% dos do tipo 2 (DM2) terão algum grau de RD (Brasil, 2021).

O tempo de duração da DM é um agravante dos fatores de risco da RD. Estima-se que cerca de 15 anos depois do diagnóstico da doença, pacientes de DM tipo 1 já apresentam algum grau de retinopatia, e 5 anos após isso 50% desses pacientes já desenvolvem RD proliferativa. Enquanto isso, a maioria dos pacientes com DM tipo 2 após 20 anos de doença possuem alguns sinais de retinopatia, sendo menos comum nos não dependentes de insulina (Silva *et al.*, 2023).

A RD é classificada clinicamente em dois tipos de estágios principais, o não-proliferativo (RDNP) e o proliferativo (RDP), sendo diferenciadas essencialmente pela presença ou não do surgimento de neovaso (Melo *et al.*, 2019; Bortoli *et al.*, 2022). Na RDNP a causa da perda da capacidade visual é resultado apenas dos edemas e isquemias maculares, caracterizada por acontecer somente na retina e não ir além da membrana limitante interna. Enquanto na RDNP não ocorre um surgimento de neovasos, na RDP ocorre o surgimento deles, além da hemorragia pré-retiniana e a hemorragia vítrea (Bastos; Oliveira; Jorge, 2022).

Ademais, como a RD é causada principalmente pela hiperglicemia, métodos para reduzir, controlar e monitorar os altos índices de glicemia podem mitigar as consequências da RD a longo prazo. “Com triagem adequada, bom controle glicêmico e pressórico e intervenção precoce com terapias cirúrgicas e farmacológicas, a perda visual grave pode ser evitada em muitos pacientes” (Franco *et al.*, 2022).

Segundo o Ministério da Saúde (Brasil, 2021), é recomendado a utilização dos métodos de oftalmoscopia binocular indireta e biomicroscopia de fundo. Ademais, a busca inicial da doença pode ser realizada com exame de fundo de olho com ou sem dilatação e retinografias simples ou coloridas, que de acordo com a *American Academy of Ophthalmology* (AAO) compõe o padrão-ouro para detecção e classificação de RD (Kanadani *et al.*, 2021; Bortoli *et al.* 2022).

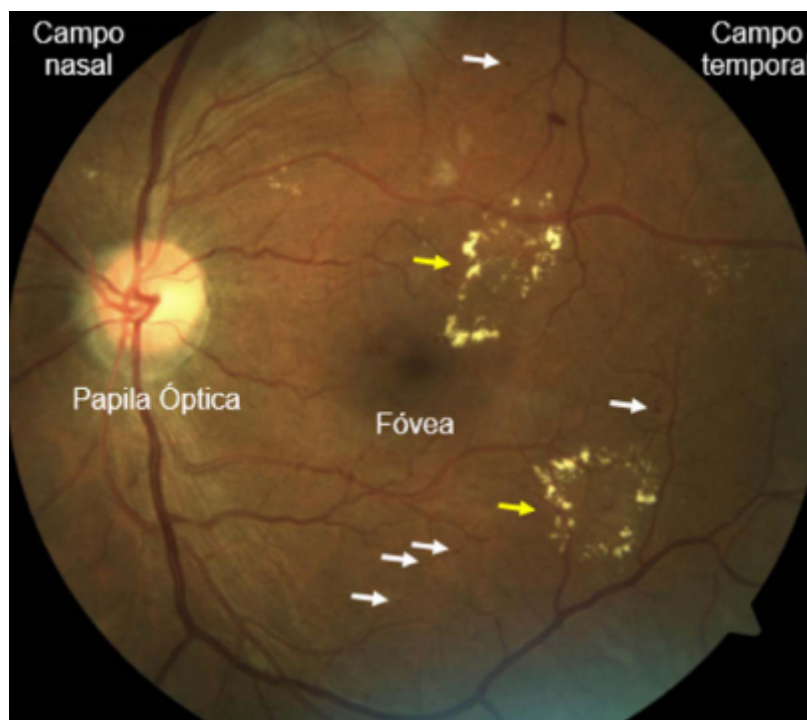
Quando há um grande número de pacientes, os exames de fundo de olho realizados por oftalmologistas treinados podem ser substituídos pelo exame de retinografia colorida, realizada por técnico ortoptista qualificado, sob coordenação de um médico oftalmologista. O uso do exame de retinografia colorida está sendo muito

utilizado na telemedicina, permitindo uma análise mais abrangente de possíveis casos de RD e evitando possíveis perdas da acuidade visual de pacientes em longas filas de espera (Kanadani *et al.*, 2021; Zanotto *et al.*, 2020).

A RD é detectada pela presença de microaneurismas, exsudatos duros e suaves, e hemorragias. Microaneurismas aparecem como pequenos pontos ($< 125 \mu\text{m}$) vermelhos arredondados, devido às fraquezas dos vasos. Os exsudatos duros aparecem como manchas amarelas brilhantes na retina causadas por vazamento de plasma, e, os exsudatos suaves aparecem como manchas brancas ovais ou arredondadas na retina causada pelo inchaço da fibra nervosa. As hemorragias aparecem como grandes manchas na retina ($> 125 \mu\text{m}$), sendo bem maiores que os microaneurismas, e sem um formato específico (Alyoubi; Shalas; Abulkhair, 2020).

Na Figura 1, pode-se notar a presença de microaneurismas, indicados pelas setas brancas, e a presença de exsudatos suaves, indicados pelas setas amarelas.

Figura 1 - Retinografia de um paciente com RDNP



Fonte: Bortoli *et al* (2022)

A triagem da RD vem se provando muito útil para o controle da doença em países com programas ativos para o rastreamento dela. Ademais, a utilização da retinografia proporcionou um grande aumento na quantidade de pacientes

atendidos, reduzindo assim a superlotação em centro ambulatoriais e grandes filas de espera, permitindo portanto uma redução no número de pessoas cegas por RD (Bortoli *et al.*, 2022).

Além disso, com o avanço das tecnologias houve um crescente aumento no número de métodos de análises automatizadas. As pesquisas de técnicas de aprendizado de máquina nos campos da medicina tiveram um crescimento exponencial nas últimas décadas (Paixão *et al.*, 2022). Os métodos automatizados de diagnósticos de RD poupam tempo e dinheiro, além de serem menos propensos a erros e não necessitarem de muitos esforços comparado com a utilização de análises manuais (Alyoubi; Shalas; Abulkhair, 2020).

2.2 Aprendizado de Máquina e Aprendizado Profundo

Aprendizado de Máquina (do inglês *Machine Learning* - ML) é a ciência (e a arte) da programação de computadores para que eles possam aprender com os dados (Géron, 2019). É um ramo do estudo de Inteligências Artificiais (IA) e que por meio de técnicas computacionais de algoritmos e aprendizado automatizado, a máquina seja capaz de aprender por conta própria através dos dados (Monard, Baranauskas, 2003; Mitchell, 1997).

O principal ponto do ML é o reconhecimento e aprendizado de padrões em conjunto de dados. Com esse aprendizado de padrões, a máquina consegue se ajustar aos dados que está recebendo. Diferente de programas de computadores ou algoritmos normais, que possuem uma hierarquia de tomada de decisões e tarefas completamente pré-definidas e imutáveis, os algoritmos de ML conseguem, com base nos padrões extraídos dos dados, ajustar os seus parâmetros para otimizar a performance de decisões (Alpaydin, 2021).

Os principais tipos de ML são definidos com base na quantidade de supervisão que recebem durante o treinamento. Os tipos são supervisionados, não-supervisionados, semi supervisionados e por reforço. No aprendizado supervisionado os dados de entradas já contém os rótulos, ou *labels*, com os resultados desejados. Algoritmos como regressão linear e logística, *k-nearest neighbours*, Máquinas de Vetores de Suporte (SVM), árvores de decisão e florestas aleatórias, são de aprendizado supervisionado (Alpaydin, 2021; Géron. 2019; Zhou, 2021).

Enquanto isso, o aprendizado não supervisionado não recebe os dados com as *labels* na entrada. Sendo utilizado em algoritmos de *clustering*, visualização e redução de dimensionalidade e aprendizado por associação. O aprendizado semi supervisionado é uma união dos dois tipos anteriores, uma parte dos dados entram rotulados e a outra parte não (Alpaydin, 2021; Géron, 2019; Zhou, 2021).

No aprendizado por reforço, o sistema de aprendizado avalia o ambiente em que está inserido e vai aprendendo com as recompensas das ações que realiza nesse ambiente. O resultado de cada ação pode lhe dar uma recompensa positiva, se for a decisão certa, ou uma recompensa negativa, se for a errada (Alpaydin, 2021; Géron, 2019; Zhou, 2021).

Indo mais a fundo, dentro do aprendizado de máquina há o modelo matemático de Redes Neurais Artificiais (RNAs), que foram criadas baseando-se nas estruturas neurais humanas. As redes neurais são o centro do Aprendizado Profundo, ou *Deep Learning*, que é um subcampo de estudo dentro do estudo de ML, onde busca-se o desenvolvimento de redes neurais profundas, ou seja, com grandes quantidades de camadas ocultas (Géron, 2019).

As RNAs são um conjunto de camadas interligadas compostas por neurônios artificiais, onde um neurônio de uma camada é ligado a outro neurônio de outra camada. Cada conexão entre neurônios é ponderada por valores reais que são chamados de pesos (Bezerra, 2016). O processamento das informações é feito nos neurônios, que recebem as informações de neurônios de camadas anteriores e envia o resultado do processamento para os neurônios das próximas camadas (Géron, 2019).

As arquiteturas de RNAs apresentam três partes principais: a camada de entrada, a camada de saída e as camadas ocultas que ficam entre a camada de entrada e a de saída. A camada de entrada recebe e prepara a informação, fazendo uma preparação linear, depois ela envia a informação para as camadas ocultas onde são processadas, e por último, são enviadas para a camada de saída (Data Science Academy, 2022).

Para se treinar uma arquitetura de RNA é preciso expor, várias vezes, a arquitetura ao conjunto de dados necessários para resolver o problema. A cada iteração em que a arquitetura “vê” o conjunto de dados completo é chamado de época. Ao final de cada época é utilizado funções de perda para identificarem a

diferença entre as previsões feitas pela a arquitetura e os valores reais do treinamento (Géron, 2019; Data Science Academy, 2022).

Com o resultado das perdas, através de algoritmos de *backpropagation* (ou retropropagação), os pesos de cada neurônio é recalculado, otimizando assim a rede e atingindo assim valores mínimos no gradiente de perda. Ou seja, aumentando a precisão das previsões da rede (Data Science Academy, 2022).

No final da década de 1950, considerada a primeira fase da IA, já se buscava o desenvolvimento de IAs para aplicações como tradução automática de idiomas e participação em jogos. Nessa época, foram desenvolvidos programas capazes de vencer jogadores principiantes (Campos, 2020). Atualmente, mais de meio século depois, as máquinas estão sendo treinadas para diversos outros tipos de aplicações, para dirigir carros, visão computacional no reconhecimento de rostos e objetos, sistemas financeiros com aplicações na bolsa de valores, entre muitos outros (Ludermir, 2021).

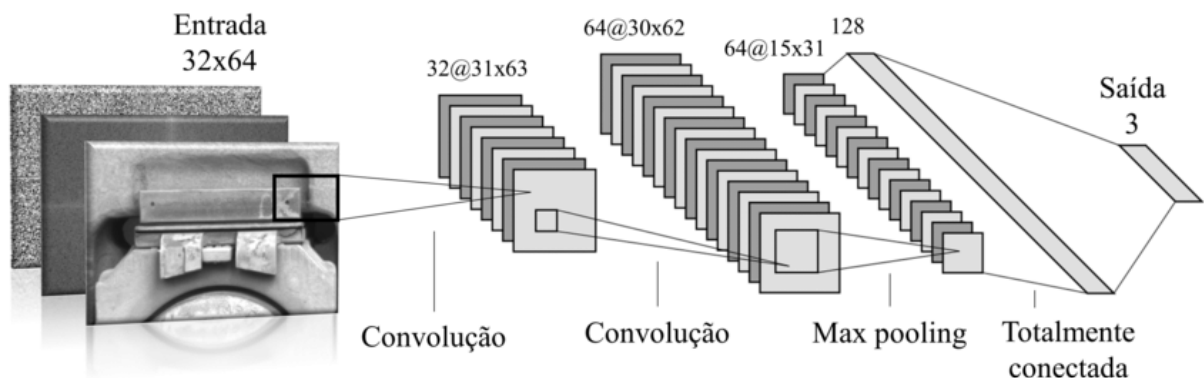
No campo da medicina vem se explorando muito a capacidade do ML de reconhecer padrões e fazer análises preditivas para poder predizer diagnósticos em diferentes áreas da saúde. Uma RNA com a eficácia semelhante a um comitê de 21 dermatologistas conseguiu distinguir, entre 129.000 casos, lesões dermatológicas benignas e malignas. Na área da cardiologia, no diagnóstico de imagens cardíacas como ecocardiograma e tomografia computadorizada e apresentaram boa acurácia em correlacionar alterações estruturais com a etiologia e fisiopatologia de doenças cardiovasculares (Paixão, *et al.* 2022).

O processamento e extração de informações importantes de imagens registradas por quaisquer meios, está dentro do campo da ciência responsável pela visão de uma máquina, ou seja, visão computacional. Este campo de estudo é recente e busca conseguir replicar o sentido da visão completa em uma máquina. Para isso utilizam métodos de aprendizado, como redes neurais, SVM, Modelos Escondidos de Markov, entre outros (Milano; Honorato, 2010).

Entretanto, as redes neurais tradicionais têm uma grande luta com a complexidade computacional necessária para o processamento de imagens. Dentro do DL, há arquiteturas de redes focadas no recebimento de imagens na entrada de dados, como as Redes Neurais Convolucionais (RNCs), principalmente utilizadas na área de reconhecimento de padrões de imagens (O'Shea; Nash, 2015)

As RNCs são arquiteturas de RNAs capazes de aprenderem e identificarem padrões independente do tamanho, escala, translação, rotação e transformações nas imagens (O'Shea; Nash, 2015), sendo muito utilizadas e mais eficientes que outros métodos de reconhecimento de imagens na medicina (Alyoubi; Shalas; Abulkhair, 2020).

Figura 2 - Representação da camadas de uma RNC



Fonte: Rocha, *et al* (2019)

As RNCs apresentam três tipos principais de camadas: as camadas convolucionais, camadas de *pooling* e camadas totalmente conectadas. As camadas convolucionais determinam a saída de cada neurônio que estão ligados à uma região local da imagem, a camada de *pooling* faz a redução dimensional do valor de entrada, diminuindo assim a quantidade de parâmetros da função de ativação, e, as camadas totalmente conectadas realizam a produção de pontuação para cada classe, com base nas ativações, para serem usadas na classificação (O'Shea; Nash, 2015).

Ao aumentar a complexidade de rede neurais adicionando camadas à elas, acabam por passar por um problema conhecido como *vanishing gradient*, onde os neurônios das camadas iniciais acabam por ter uma taxa de aprendizado inferior ao das camadas posteriores (Data Science Academy, 2022). Para lidar com este problema em RNCs, desenvolveu-se às *Residual Neural Networks*, ou *resnets*, que possuem como principal componente os blocos residuais (He et al, 2015).

Os blocos residuais são caracterizados por uma conexão de atalho que adiciona diretamente o valor da entrada do bloco à saída das camadas convolucionais subsequentes dentro do próprio bloco. Essa arquitetura permitiu que

as ResNets alcançassem uma otimização substancialmente superior em comparação com arquiteturas concorrentes, mesmo ao empregar uma profundidade de camadas significativamente maior, o que foi evidenciado por seu desempenho notável no extenso conjunto de dados da ImageNet (He *et al.*, 2016).

2.3 Trabalhos Relacionados

Wan (2018), em seu trabalho “*Deep convolutional neural networks for diabetic retinopathy detection by image classification*”, afirma que a detecção automática de diabetes tem trazido muitos benefícios, incluindo poder realizar um diagnóstico precoce da doença. Nele utiliza-se diferentes tipos de arquiteturas de RNCs, a *AlexNet*, *VggNet-s*, *VggNet-16*, *VggNet-19*, *GoogLeNet* e *ResNet*, para a detecção da RD em um *dataset* público de imagens de retinografias. Onde a *VggNet-s*, com hiperparâmetros ajustados, apresentou uma maior precisão, uma sensibilidade mediana e uma maior especificidade que as outras arquiteturas. Entretanto o estudo foi limitado a uma única base de dados e não detectou diretamente lesões.

Ademais, Khalifa *et al* (2019) reforça a eficiência de arquiteturas como *ResNet*, *AlexNet* e *VggNet*, apresentando modelos com uma alta acurácia e sensibilidade utilizando destas arquiteturas e da base de dados *Kaggle APTOS 2019*. Mohanty *et al* (2023) também apresenta uma alta acurácia ao utilizar a arquitetura *DenseNet-121*, na mesma base de dados, em conjunto com técnicas de pré-processamento, como o desfoque gaussiano.

No estudo “*Deep learning for diabetic retinopathy detection and classification based on fundus images: A review*” de Tsiknakis (2021), que faz uma revisão de detecção de RD utilizando *deep learning*, a *GoogLeNet* apresentou uma boa especificidade para a classificação em duas classes no mesmo *dataset* utilizado por Wan (2018). Ademais, arquiteturas customizadas de RNCs alinhadas com árvores de decisões, permitiram ótimos desempenhos também na classificação de duas classes tanto neste *dataset* e também em outros. Aumentando a quantidade de classe para 5, permitindo cobrir todos os estágios da doença, há uma queda nos resultados, que pode permitir evitar o sobreajuste nos datasets.

Por outro lado, no estudo “*Diabetic retinopathy detection through deep learning techniques: A review*” (Alyoubi; Shalas; Abulkhair, 2020), constatou-se que os resultados estão fortemente ligados ao tamanho da base de dados utilizada para

treinar, além de qualidade das imagens e equilíbrio na quantidade de imagens por classes. Dos 59% dos trabalhos analisados pelo autor, que utilizam mais de uma base de dados, utilizaram mais de uma base de dados públicos.

Em complemento, os trabalhos que desenvolveram sua própria arquitetura de RNC tiveram uma precisão mais notável do que outros que utilizaram de arquiteturas prontas. A maioria dos estudos analisados classificaram os dados em apenas duas classes, uma que indicava a presença de RD e outra que não, portanto, não identificando os estágios da doença. E, apenas uma pequena parte da minoria realizou a classificação em estágios e os tipos de lesões.

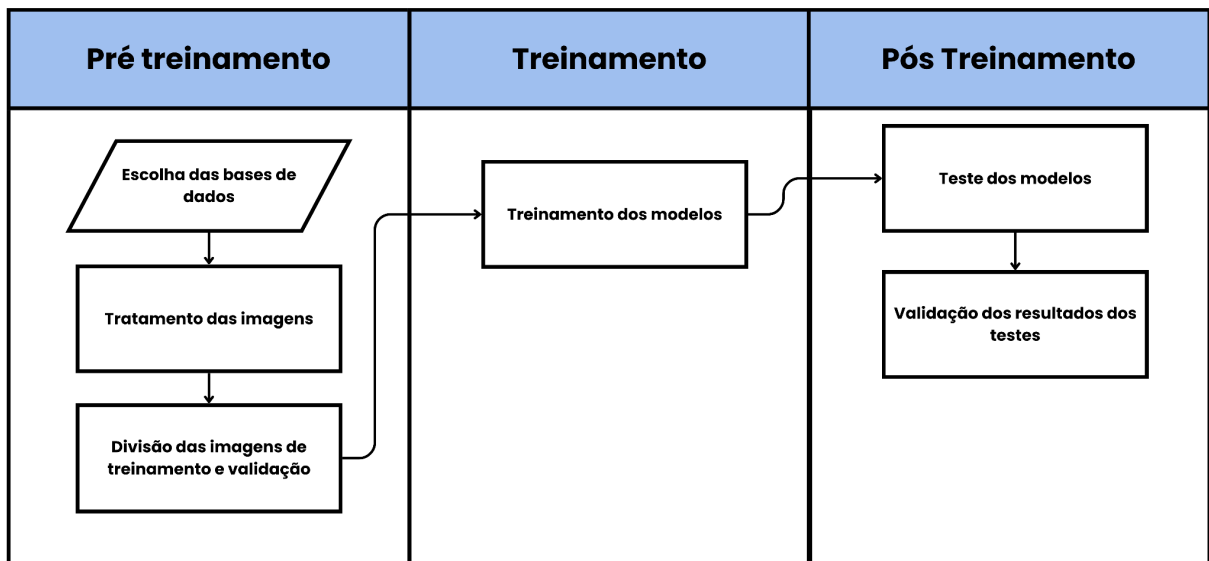
3 METODOLOGIA

Este estudo possui uma metodologia experimental quantitativa, pois visa manipular variáveis como o tamanho dos dados para treinamentos, pré-processamento de imagens e ajustes no modelo, a fim de levantar informações que serão posteriormente analisadas com métodos estatísticos

Neste sentido, um estudo exploratório caracteriza-se pela criação de um ambiente controlado para o surgimento de fatos que identifiquem alguma correlação entre eles, e, um estudo de abordagem quantitativa busca traduzir números em informações classificadas (Rodrigues, 2007).

Assim, o processo metodológico proposto nesta pesquisa será constituído de 3 (três) etapas: pré-treinamento, treinamento e pós treinamento, conforme apresentadas na Figura 3.

Figura 3 - Fluxograma do processo metodológico



Fonte: Elaborada pelo autor (2025)

Conforme apresentado na Figura 3, na primeira etapa da metodologia, foi realizado a escolha das bases de dados desta pesquisa. O conjunto de imagens escolhido para o treinamento do modelo foi o conjunto de dados públicos *Kaggle APTOS 2019 Blindness Detection*.

O *Kaggle APTOS 2019 Blindness Detection* é uma base de dados feita pelo hospital indiano *Aravind Eye Hospital* com imagens de fundo de olho de pessoas de áreas rurais. O conjunto possui 5590 imagens, entretanto utilizou-se apenas 3662 imagens, pois estavam devidamente classificadas em 5 classes diferentes, utilizando como base a Escala de Gravidade da Doença Retinopatia Diabética (Brasil, 2021).

Cada imagem é rotulada com um número de 0 a 4, cada número representa uma das classes, tendo 0 como uma imagem que não tem retinopatia diabética, 1 como retinopatia diabética não proliferativa leve, 2 como retinopatia diabética não proliferativa moderada, 3 como retinopatia diabética não proliferativa severa e 4 como retinopatia diabética proliferativa.

Em seguida, houve a divisão do conjunto de dados em treinamento, validação e teste. O conjunto foi dividido em dois subconjuntos, um com 20% das imagens, que foram utilizadas para teste do modelo, e, outro com 80% das imagens, no qual foi repartido em 20% para validação e o restante para treinamento. Na Tabela 1 é apresentado a quantidade final da distribuição de imagens nos conjuntos.

Tabela 1 - Quantidade de imagens para cada etapa

Classes	Quantidade de imagens para treinamento	Quantidade de imagens para validação	Quantidade de imagens para testes
Classe 0	1011	433	361
Classe 1	208	88	74
Classe 2	560	239	200
Classe 3	108	46	39
Classe 4	166	70	59
Total	2053	876	733

Fonte: Elaborada pelo autor (2025)

Após a etapa de divisão da base de dados, ocorreu a etapa de treinamento do modelo. Nesta etapa executou-se o treinamento de 5 modelos diferentes. O primeiro modelo consistiu em um classificador binário entre olho saudável e olho doente, utilizando todas as imagens que apresentavam alguma classificação de doença contra as de olhos sem doença.

Os outros 4 treinamentos basearam-se em treinar um classificador binário que comparasse um olho normal com cada estágio da doença, ou seja, um modelo para distinguir olhos normais de olhos com estágio 1, outro para estágio 2, outro para estágio 3 e outro para estágio 4 da doença.

Todos os modelos foram treinados com a arquitetura de RNC *resnet 18*, pré-treinada com os pesos do conjunto de dados *imagenet*; com as camadas iniciais congeladas, mantendo apenas a camada 4 e a de saída descongeladas, para a extração de características. Ademais, utilizou-se a função de perda *Binary Cross Entropy*, com ponderação de classes para lidar com o desbalanceamento das mesmas, e o otimizador *Adam* com uma taxa de aprendizado de $1e-5$. Foram aplicados *batches* de 32 imagens, funções de *Early Stopping*, com paciência de 15 épocas para o perda de validação, e *Reduce Learning Rate On Plateau* com um fator de 0.1 e paciência de 5 épocas.

Posteriormente, realizou-se a etapa de pós-treinamento, ou, avaliação do modelo. Nela foi feito o teste de cada modelo com as imagens restantes das suas respectivas classes e posteriormente a avaliação de cada modelo. As métricas que foram utilizadas são: precisão, sensibilidade e especificidade (Wan, Liang, Zhang, 2018; Monard, Baranauskas. 2003).

$$Acurácia = \frac{TN + TP}{(TN + TP + FN + FP)} \quad (1)$$

$$Sensibilidade = \frac{TP}{(TP + FN)} \quad (2)$$

$$Especificidade = \frac{TN}{(TN + FP)} \quad (3)$$

A acurácia identifica a taxa de acertos do modelo, verificando quantas das imagens totais o modelo classificou corretamente. As métricas de sensibilidade e especificidade examinam, respectivamente, a quantidade de imagens positivas que foram classificadas como positivas e a quantidade de imagens negativas que foram classificadas como negativas.

Neste estudo, as métricas de sensibilidade e especificidade são vitais para a avaliação do modelo, pois, identificar as taxas de verdadeiros positivos e verdadeiros negativos permitem que a classificação de falsos negativos seja reduzida. Porventura, um paciente que possua a doença, receba uma classificação de saudável, terá sérias complicações no seu futuro, podendo ter toda a sua visão prejudicada.

4 RESULTADOS

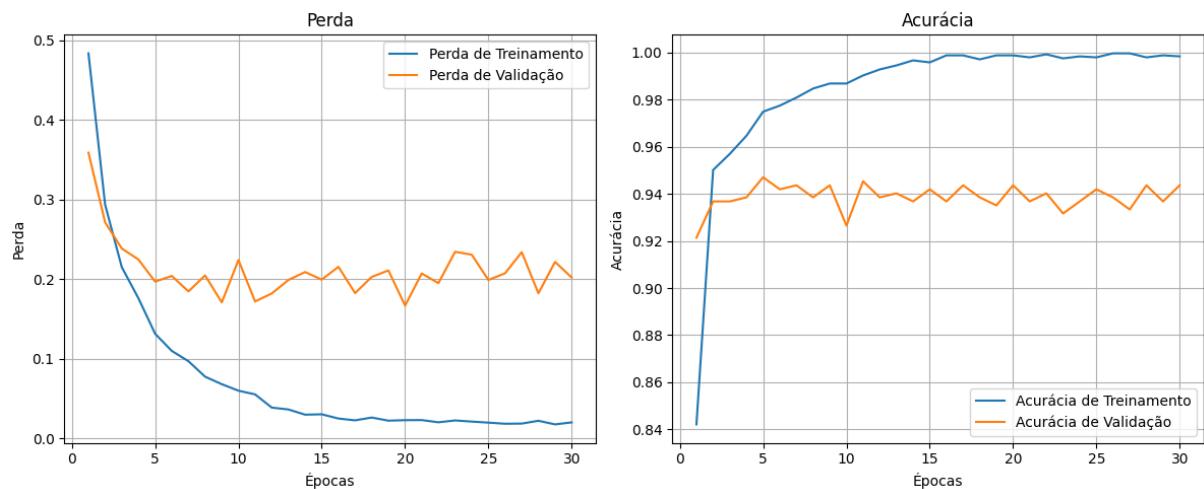
Nesta subseção são abordados os resultados dos 5 (cinco) modelos treinados, e, posteriormente, executados no conjunto de dados de testes separados previamente. Nos processos de treinamento é evidenciado os gráficos de acurácia e perda para avaliar o aprendizado dos modelos. Com a aplicação dos modelos nos conjuntos de testes é evidenciado as métricas de sensibilidade e especificidade.

4.1 Olho Saudável vs Olho Doente

Durante o treinamento do modelo de olho saudável vs olho doente, apresentado na Figura 4, pode-se observar uma curva crescente de acurácia de

treinamento, aumentando no decorrer de cada época. Entretanto, a acurácia de validação não acompanha essa crescente, se mantendo sempre num intervalo de 92% a 94%. Já nas épocas iniciais é possível identificar um sinal de sobreajuste do modelo, pois a perda de validação para de acompanhar o treinamento, e, apresenta algumas flutuações nos valores.

Figura 4 - Gráficos do treinamento e validação do modelo de classificação entre olhos saudáveis e olhos doentes



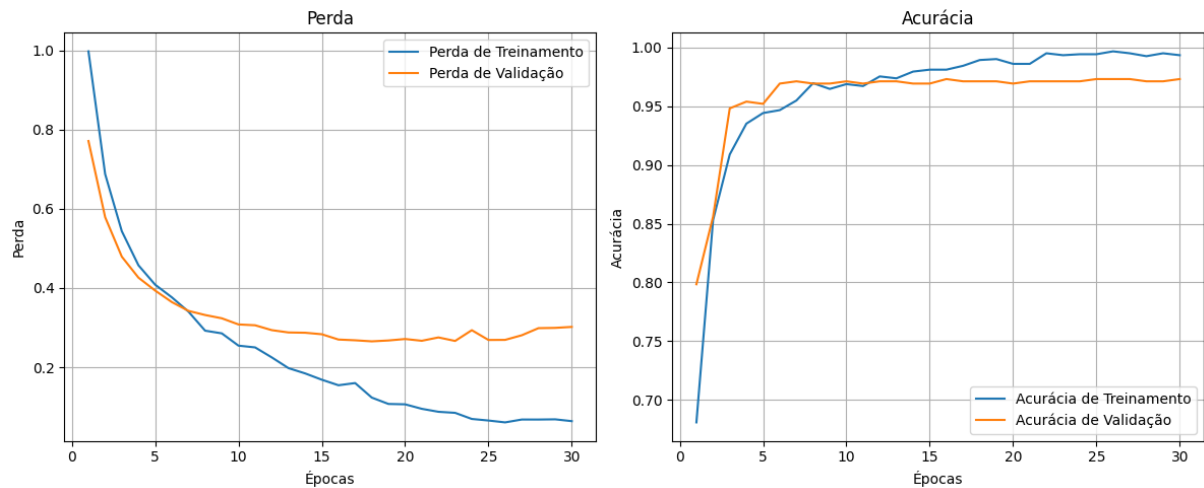
Fonte: Elaborada pelo autor (2025)

A função de *Early Stopping* salva o modelo na época 20, mas já com altos sinais de sobreajuste. O modelo salvo ao ser colocado para executar as imagens de testes obteve uma acurácia de 95.4%, sensibilidade de 95.7% e especificidade de 95.2%. Classificando apenas 16, das 372, imagens com olhos doentes do conjunto de teste como olhos saudáveis.

4.2 Olho Saudável vs Cada Estágio da Doença

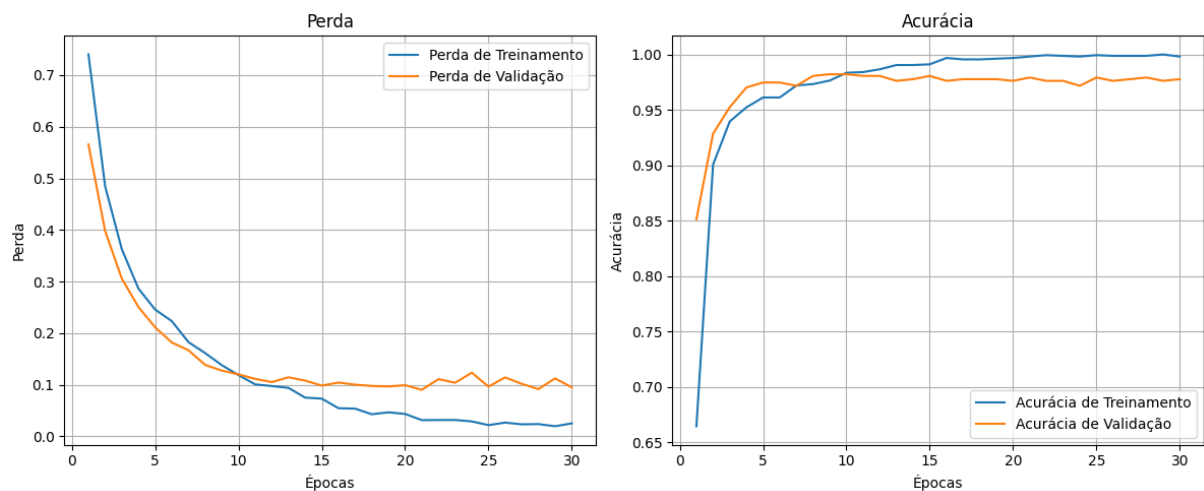
As Figuras 5, 6, 7 e 8 mostram a curva de perda e acurácia de treinamento do seu respectivo modelo. Todos os modelos, por volta da terceira à quinta época conseguem convergir em uma alta acurácia, atingindo um platô, enquanto a perda continua a decair.

Figura 5 - Gráficos do treinamento e validação do modelo de classificação entre olhos saudáveis e olhos no estágio inicial



Fonte: Elaborada pelo autor (2025)

Figura 6 - Gráficos do treinamento e validação do modelo de classificação entre olhos saudáveis e olhos no estágio moderado

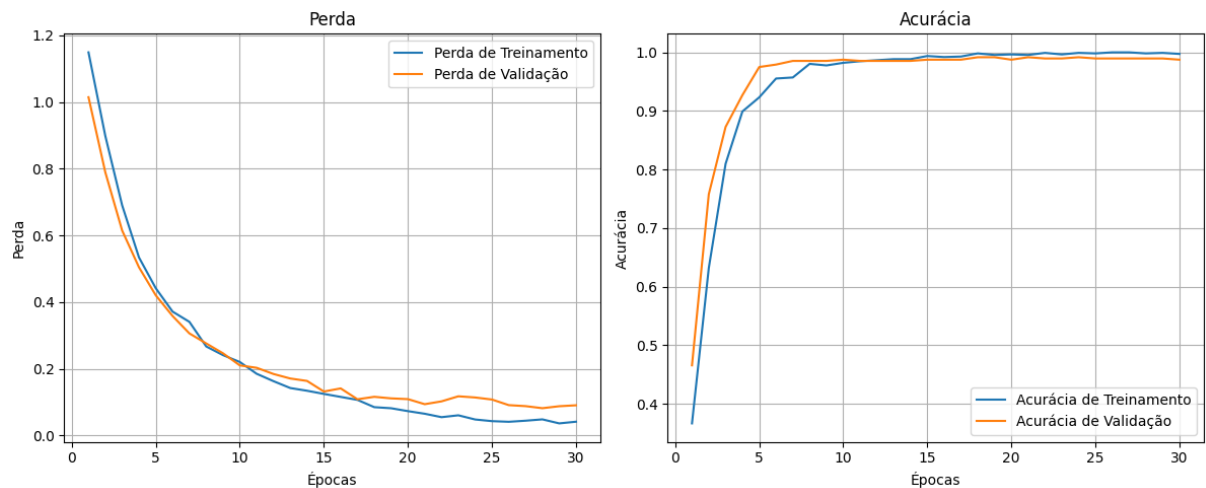


Fonte: Elaborada pelo autor (2025)

Os modelos para detectar estágios iniciais da doença apresentaram uma taxa de acurácia, no treinamento e na validação, semelhantes. A perda na validação de ambos os modelos estabiliza por volta da época da 15ª, após esta época passa a apresentar oscilações no modelo do estágio moderado e uma tendência crescente

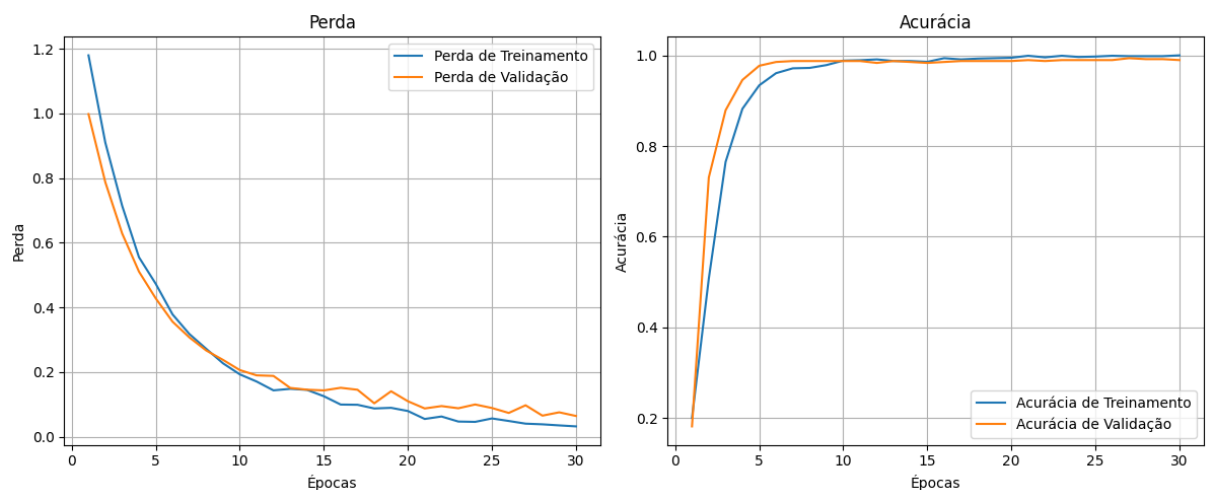
no modelo de estágio inicial. Tais padrões, indicam um início de sobreajuste de ambos os modelos.

Figura 7 - Gráficos do treinamento e validação do modelo de classificação entre olhos saudáveis e olhos no estágio severo



Fonte: Elaborada pelo autor (2025)

Figura 8 - Gráficos do treinamento e validação do modelo de classificação entre olhos saudáveis e olhos no estágio proliferativo



Fonte: Elaborada pelo autor (2025)

Os modelos de classificação do estágio severo e proliferativo conseguem ter uma diferença entre a perda de validação e treinamento melhor do que os treinamentos dos estágios leve e moderado. Esta diferença de desempenho se dá

pelo fato de olhos com os estados mais avançados da doença apresentarem mais características da mesma. Entretanto, a partir da 20ª época os modelos começam a apresentar oscilações nas perdas e uma leve queda na acurácia de validação.

Como se pode observar na Tabela 2, a especificidade dos modelos no conjunto de testes apresentam resultados semelhantes ou superiores a sensibilidade, ou seja, o modelo consegue prever melhor olhos saudáveis do que olhos doentes. Esta superioridade para detecção de olhos saudáveis se dá principalmente pela quantidade de amostras no treinamento.

Tabela 2 - Acurácia, sensibilidade e especificidade para cada modelo no conjunto de teste com limiar de 0.5

Modelo	Acurácia	Sensibilidade	Especificidade
Olho Saudável vs Olho Doente Leve	93.3%	79.7%	96.1%
Olho Saudável vs Olho Doente Moderado	97.1%	97.5%	97.0%
Olho Saudável vs Olho Doente Severo	98.2%	92.3%	98.8%
Olho Saudável vs Olho Doente Proliferativo	98.5%	92.3%	99.2%

Fonte: Elaborada pelo autor (2025)

Modelos que tiveram menos amostras de olhos doentes tiveram uma sensibilidade menor. Entretanto, mesmo os conjuntos de olhos doentes leve terem uma maior quantidade de imagens em relação ao estágio severo e proliferativo, os olhos doentes apresentam poucas características da doença, sendo mais difícil diferenciar de olhos saudáveis.

Ao final desses testes foi montado uma pipeline de execução, que foi aplicado o conjunto de teste do classificador de olho saudável e olho doente, em que cada imagem fosse testada em todos os modelos. Quando ao menos um modelo classifica se a imagem como doente, ela seria considerada positiva. A pipeline teve uma acurácia de 95.2%, sensibilidade de 94.6% e especificidade de 95.8%.

Ademais, no intuito de melhorar a sensibilidade do modelo, aumentado sua detecção de verdadeiros positivos, foi realizado o teste dos modelos com um limiar de classificação reduzido de 0.5 para 0.3.

Percebe-se, através da Tabela 3, que as acurácias de todos os modelos com o limiar reduzido, incluindo da *pipeline*, houve uma queda, junto com a especificidade. Todavia, houve um aumento na sensibilidade dos modelos, indicando uma maior taxa de acerto de verdadeiros positivos, menos no de Olho Saudável vs Olho Doente Leve que permaneceu igual, tendo apenas uma queda na especificidade.

Tabela 3 - Acurácia, sensibilidade e especificidade para cada modelo no conjunto de teste com limiar de 0.3

Modelo	Acurácia	Sensibilidade	Especificidade
Olho Saudável vs Olho Doente	95.8%	96.5%	95.0%
Olho Saudável vs Olho Doente Leve	92.9%	82.4%	95.0%
Olho Saudável vs Olho Doente Moderado	96.8%	98.0%	96.1%
Olho Saudável vs Olho Doente Severo	97.7%	97.4%	97.8%
Olho Saudável vs Olho Doente Proliferativo	98.2%	97.4%	98.3%
Pipeline Com os 4 Estágios	94.5%	95.7%	93.4%

Fonte: Elaborada pelo autor (2025)

5 CONSIDERAÇÕES FINAIS

Os modelos desenvolvidos neste estudo apresentaram ótimos resultados em cada uma das métricas selecionadas. Onde todos os modelos apresentaram resultados semelhantes aos modelos treinados por Khalifa *et al* (2019) e Mohanty *et al* (2023), que treinaram modelos em arquiteturas como *AlexNet*, *VGG16*, *ResNet18* e *DenseNet*.

Contudo, no treinamento do classificador binário de olho doente é possível identificar, em poucas épocas, um sobreajuste precoce do modelo, sendo isso algo que não é visível nos outros 4 modelos tão cedo no seus respectivos treinamentos.

A união de 4 modelos que identificam cada estágio da doença em uma *pipeline* ajuda a mitigar o sobreajuste presente no treinamento do modelo binário que identifica se há ou não sinais da doença. Pois, cada modelo aprende características de cada estágio da doença, e, para a pipeline classificar uma imagem com doença basta que apenas um dos modelos detecte a doença.

Este sobreajuste nos modelos provavelmente foi causado pelo tamanho do conjunto de dados utilizados para treinar os modelos, tal fato sendo a principal limitação deste trabalho. Portanto, em futuras pesquisas sugere-se aumentar a base de dados utilizada, com técnicas de aumento de dados ou com a adição de imagens de outros conjuntos.

REFERÊNCIAS

ALPAYDIN, Ethem. **Machine Learning**: revised and updated edition. Cambridge: The MIT Press, 2021.

ALYOUBI, Wejdan L.; SHALASH, Wafaa M.; ABULKHAIR, Maysoon F. **Diabetic retinopathy detection through deep learning techniques: A review**. *Informatics in Medicine Unlocked*, v. 20, p. 100377, 2020. DOI: 10.1016/j.imu.2020.100377. Acesso em: 27 nov. 2024.

BASTOS, Thaís Marino de Azeredo; OLIVEIRA, Ítalo Pena de; JORGE, Rodrigo. **Manifestações oculares de doenças sistêmicas II: retinopatia diabética e retinopatia hipertensiva**. *Medicina (Ribeirão Preto)*, Ribeirão Preto, Brasil, v. 55, n. 2, p. e-178543, 2022. DOI: 10.11606/issn.2176-7262.rmrp.2022.178543. Acesso em: 18 out. 2024.

BEZERRA, Eduardo. **Introdução à Aprendizagem Profunda**. Berlim: ResearchGate, 2016. Disponível em: https://www.researchgate.net/publication/309321510_Introducao_a_Aprendizagem_Profunda. Acesso em: 2 dez. 2024.

BORTOLI, Julia Quadri; SILBER, Paulo Caldas; PICETTI, Egídio; SILVA, Cássio Freitas da Silva; Pakter, Helena Messinger. **Retinografia como forma de rastreio de retinopatia diabética em hospital terciário do Sistema Único de Saúde**. *Revista Brasileira de Oftalmologia*, v. 81, p. e0057, 2022. DOI: 10.37039/1982.8551.20220057. Acesso em: 19 out. 2024.

BRASIL. Ministério da Saúde. **Portaria Conjunta nº 17, de 01 de outubro de 2021**. Aprova o Protocolo Clínico e Diretrizes Terapêuticas da Retinopatia Diabética. Brasília, DF: Ministério da Saúde, 2021. Disponível em: https://www.gov.br/conitec/pt-br/midias/relatorios/portaria/2021/20211217_portaria_conjunta_17.pdf/view. Acesso em: 18 out. 2024.

BRASIL. Ministério da Saúde. **Saúde de A a Z**. 2024. Disponível em: <https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/d/diabetes>. Acesso em: 3 dez. 2024.

CAMPOS, Rafael Saraiva. **Desmistificando a inteligência artificial: uma breve introdução conceitual ao aprendizado de máquina**. *Aoristo - International Journal of Phenomenology, Hermeneutics and Metaphysics*, v. 3, n. 1, p. 106-123, 2020. DOI: 10.48075/aoristo.v3i1.24880. Acesso em: 5 nov. 2024.

CASARIN, Daniele Escudeiro; DONADEL, Guilherme; DALMAGRO, Mariana; OLIVEIRA, Priscila Cogo de; CERANTO, Daniela de Cássia Faglioni Boleta; ZARDETO, Giuliana. **Diabetes mellitus: causas, tratamento e prevenção**.

Brazilian Journal of Development, v. 8, n. 2, p. 10062-10075, 2022. DOI: 10.34117/bjdv8n2-107. Acesso em: 21 out. 2024.

DATA SCIENCE ACADEMY. **Deep Learning Book**, 2022. Disponível em: <https://www.deeplearningbook.com.br/>. Acesso em: 15 nov. 2024.

FRANCO, Elisa Marques; SILVA, Leonardo Neves; ROCHA, Leonardo Dias Flor; BATISTA, Marina Andrade Aguiar; SATHLER, Yasmin Gomes; PAIVA, Lovily Duarte Toledo; FRANCO, Cláudia Marques Andrade; DURANTE, Brenda Castanha.

Revisão bibliográfica: retinopatia diabética. *Brazilian Journal of Development*, [S. l.], v. 8, n. 5, p. 35257–35264, 2022. DOI: 10.34117/bjdv8n5-168. Acesso em: 19 out. 2024.

GÉRON, Aurelien. **Mãos à Obra Aprendizado de Máquina com Scikit-Learn & TensorFlow**. Sebastopol: O'Reilly, 2019.

HE, Kaiming; ZHANG, Xiangyu; REN, Shaoqing; SUN, Jian. **Deep residual learning for image recognition**. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), p. 770–778, 2016. DOI: 10.1109/CVPR.2016.90. Acesso em: 02 jul. 2025.

KANADANI, Tereza Cristina Moreira; BRAGA, Carolina Serpa; IBRAHIM, Larissa Fouad; SILVA, Senice Alvarenga Rodrigues; KANADANI, Fábio Nishimura. **Comparação da eficácia da retinografia colorida e da angiografia fluorescente no diagnóstico da retinopatia diabética.** *Revista Brasileira de Oftalmologia*, v. 80, n. 6, p. 48, 2021. DOI: 10.37039/1982.8551.20210048. Acesso em: 21 out. 2024.

KHALIFA, Nour Eldeen M.; LOEY, Mohamed; TAHA, Hamed N.; MOHAMED, Hamed Nasr Eldin T. **Deep Transfer Learning Models for Medical Diabetic Retinopathy Detection**. *Acta Informatica Medica*, v. 27, n. 12, p. 327. DOI: 10.5455/aim.2019.27.327-332. Acesso em: 30 jun. 2025

LUDERMIR, Teresa Bernarda. **Inteligência Artificial e Aprendizado de Máquina: estado atual e tendências**. *Estudos Avançados*, v. 35, n. 101, p. 85–94, jan. 2021. DOI: 10.1590/s0103-4014.2021.35101.007. Acesso em: 13 nov. 2024.

MELO, Laura Gomes Nunes. **Prevalência da retinopatia diabética e seus fatores de risco em pacientes com diabetes tipo 1 no Brasil**. 2019. 102 f. Tese (Doutorado em Fisiopatologia Clínica e Experimental) - Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2019. Disponível em: <https://www.bdt.d.uerj.br:8443/handle/1/12589>. Acesso em: 22 out. 2024.

MILANO, Danilo de; HONORATO, Luciano Barrozo. **Visão computacional. Faculdade de Tecnologia, Universidade Estadual de Campinas**, 2010.

MITCHELL, Tom M.. **Machine learning**. Vol. 1. No. 9. New York: McGraw-hill, 1997.

MOHANTY, Cheena; MAHAPATRA, Sakuntala; ACHARYA, Biswaranjan; KOKKORAS, Fotios; GEROGIANNIS, Vassilis; KARAMITSOS, Ioannis; KANAVOS, Andreas. **Using Deep Learning Architectures for Detection and Classification of**

Diabetic Retinopathy. Sensors, v. 23, n. 12, p. 5726, 2023. DOI: 10.3390/s23125726. Acesso em: 6 jul. 2025.

MONARD, Maria Carolina; Baranauskas, José Augusto. **Conceitos Sobre Aprendizado de Máquina**, Sistemas Inteligentes, Fundamentos e Aplicações, v. 1, n. 1. São Paulo: Manole Ltda, 2003.

O'SHEA, Keiron; NASH, Ryan. **An Introduction to Convolutional Neural Networks.** New York: Cornell University, 2015. Disponível em: <https://arxiv.org/abs/1511.08458>. Acesso em: 17 nov. 2024

PAIXÃO, Gabriela Miana de Mattos; SANTOS, Bruno Campos; ARAUJO, Rodrigo Martins de; RIBEIRO, Manoel Horta; MORAES, Jermana Lopes de; RIBEIRO, Antonio L.. **Machine Learning na Medicina: Revisão e Aplicabilidade.** Arquivos Brasileiros de Cardiologia, v. 118, n. 1, p. 95-102, jan. 2022

ROCHA, Rafael; SILVA, Cleison; GOMES, Ana; FERREIRA, Bruno; CARVALHO, Eduardo; SRAVENHA, Ana; CARVALHO, Schubert. **Image Inspection of Railcar Structural Components: An approach through Deep Learning and Discrete Fourier Transform.** Symposium on Knowledge Discovery, Mining and Learning, KDMILE, 2019. DOI: 10.5753/kdmile.2019.8786. Acesso em: 02 jun. 2025.

RODRIGUES, William Costa. **Metodologia científica.** Faetec/IST. Paracambi, v. 2, 2007.

SANCHEZ, C. I.; HORNERO, R.; LOPEZ, M. I.; POZA, J. **Retinal image analysis to detect and quantify lesions associated with diabetic retinopathy.** In: *IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY. Annual International Conference*, 26., 2004, San Francisco, CA, USA. Anais... New York: IEEE, 2004. p. 1624-1627. DOI: 10.1109/IEMBS.2004.1403492. Acesso em: 29 nov. 2024.

SOCIEDADE BRASILEIRA DE DIABETES (SBD). **Brasil já tem cerca de 20 milhões de pessoas com diabetes.** 2024. Disponível em: <https://diabetes.org.br/brasil-ja-tem-cerca-de-20-milhoes-de-pessoas-com-diabetes/>. Acesso em: 18 out. 2024.

SILVA, S. N. da; COSTA, A. M. S. da; SOARES, J. de O.; SILVA, L. da; PONTES, A. N. **Retinopatia em pacientes diabéticos: fatores de risco importantes.** Revista JRG de Estudos Acadêmicos, Brasil, São Paulo, v. 6, n. 13, p. 134–149, 2023. DOI: 10.5281/zenodo.7950770. Acesso em: 19 out. 2024.

TSIKNAKIS, Nikos; THEODOROPOULOS, Dimitris; MANIKIS, Georgios; KTISTAKIS, Emmanouil; BOUTSORA, Ourania; BERTO, Alexa; SCARPA, Fabio; SCARPA, Alberto; FOTIADIS, Dimitrios I.; MARIAS, Kostas. **Deep learning for diabetic retinopathy detection and classification based on fundus images: A review.** *Computers in Biology and Medicine*, v. 135, p. 104599, 2021. DOI: 10.1016/j.combiomed.2021.104599. Acesso em: 18 out. 2024.

WAN, Shaohua; LIANG, Yan; ZHANG, Yin. **Deep convolutional neural networks for diabetic retinopathy detection by image classification.** *Computers & Electrical*

Engineering, v. 72, p. 274-282, 2018. DOI: 10.1016/j.compeleceng.2018.07.042. Acesso em: 25 out. 2024.

ZANOTTO, Bruna Stella, ETGES, Ana Paula Beck da Silva ; SIQUEIRA, Ana Célia. **Avaliação Econômica de um Serviço de Telemedicina para ampliação da Atenção Primária à Saúde no Rio Grande do Sul: o microcusteio do Projeto TeleOftalmo**. *Ciência & Saúde Coletiva*, v. 25, n. 4, p. 1349–1360, 2020. DOI: <https://doi.org/10.1590/1413-81232020254.28992019>.

ZHOU, Zhi-Hua. **Machine learning**. Berlim: Springer nature, 2021.