



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

MAXMILIANO RODRIGUES BARBOSA DE SOUSA

**DEEP LEARNING APLICADO AO DIAGNÓSTICO DA DEGENERAÇÃO
MACULAR RELACIONADA À IDADE**

FLORIANO - PI

2025

MAXMILIANO RODRIGUES BARBOSA DE SOUSA

DEEP LEARNING APLICADO AO DIAGNÓSTICO DA DEGENERAÇÃO MACULAR
RELACIONADA À IDADE

Trabalho de Conclusão de Curso (artigo científico) apresentado como exigência parcial para obtenção do diploma do Curso de Tecnologia em Análise e desenvolvimento de sistemas do Instituto Federal do Piauí, Campus Floriano.

Orientador: Prof. Dr. Rafael Angelo Santos Leite

Coorientador: Prof. Dr. José Joaquim de Moura Ramos

DEEP LEARNING APLICADO AO DIAGNÓSTICO DA DEGENERAÇÃO MACULAR RELACIONADA À IDADE

Maxmiliano Rodrigues Barbosa de Sousa¹
Rafael Angelo Santos Leite²

RESUMO

A Degeneração Macular Relacionada à Idade representa uma das principais causas de cegueira irreversível em idosos, e seu diagnóstico precoce no Brasil enfrenta desafios significativos devido à desigualdade na distribuição de médicos especialistas. Este trabalho teve como objetivo principal desenvolver e avaliar um sistema de inteligência artificial para atuar como ferramenta de triagem da doença, utilizando imagens de retinografia, que é um exame mais acessível. A metodologia adotou uma abordagem quantitativa e experimental, comparando o desempenho de duas arquiteturas de redes neurais convolucionais, a DenseNet-121 e a EfficientNet-B1. Os modelos foram treinados com técnicas de aprendizado por transferência, aumento de dados e validados por meio de validação cruzada. Os resultados indicaram que o modelo EfficientNet-B1, apesar de uma acurácia geral ligeiramente menor, foi superior na métrica de sensibilidade (recall), alcançando 72,58%. Essa métrica é clinicamente mais relevante, pois demonstra maior capacidade de identificar corretamente os casos positivos, minimizando o risco de diagnósticos perdidos. Conclui-se que o sistema desenvolvido é uma ferramenta de triagem pragmática e eficaz, com alto potencial para otimizar o fluxo de trabalho dos especialistas e democratizar o acesso à saúde ocular, especialmente em regiões remotas.

Palavras-chave: aprendizado profundo; degeneração macular; diagnóstico por imagem; retinografia.

ABSTRACT

Age-Related Macular Degeneration represents one of the main causes of irreversible blindness in the elderly, and its early diagnosis in Brazil faces significant challenges due to the unequal distribution of medical specialists. This work's main objective was to develop and evaluate an artificial intelligence system to act as a screening tool for the disease, using retinography images, which is a more accessible exam. The methodology adopted a quantitative and experimental approach, comparing the performance of two convolutional neural network architectures, DenseNet-121 and EfficientNet-B1. The models were trained with transfer learning and data augmentation techniques and validated through cross-validation. The results indicated that the EfficientNet-B1 model, despite a slightly lower overall accuracy, was superior in the sensitivity (recall) metric, achieving 72.58%. This metric is clinically more relevant as it demonstrates a greater ability to correctly identify

¹Graduando em análise e Desenvolvimento de Sistemas. IFPI Campus Floriano.
sousamax57@gmail.com.

²Doutor em Ciência da Propriedade Intelectual pela UFS e docente do eixo Gestão e Negócios.
IFPI Campus Floriano. rafaelangelo@ifpi.edu.br

positive cases, minimizing the risk of missed diagnoses. It is concluded that the developed system is a pragmatic and effective screening tool, with high potential to optimize the workflow of specialists and democratize access to eye health, especially in remote regions.

Keywords: deep learning; macular degeneration; diagnostic imaging; retinography.

Data de aprovação: 09/07/2025

1 INTRODUÇÃO

A degeneração macular relacionada à idade (DMRI) é uma doença que acomete a parte central da retina, a mácula, causando a sua deterioração funcional e apresentando sintomas como o escurecimento, a visão distorcida em sua parte central e por fim a perda progressiva da visão. Pode ser classificada em dois tipos: a seca (atrófica) e a úmida (neovascular). A forma seca é caracterizada pelo aparecimento de manchas amarelas, causadas pelo acúmulo dos bastonetes e cones que ficam depositados na parte central da retina. Já a forma úmida se dá pela criação de neovasos na retina, deixando assim, vazar sangue e outros fluidos na mesma (Mehta, 2024).

A DMRI é a causa mais comum de cegueira nos países industrializados e, conforme a pesquisa realizada pelo Conselho brasileiro de oftalmologia (CBO), afeta cerca de 30% da população com mais de 75 anos, tornando-se uma grande ameaça para a saúde ocular, tendo em vista o aumento da expectativa de vida da população, a porcentagem de pessoas com a doença deve aumentar ao passar dos anos (Umbelino; Ávila, 2023)

No Brasil, um dos principais desafios para o diagnóstico precoce e o acompanhamento da DMRI é a falta de profissionais capacitados, em especial nas regiões mais afastadas. Conforme o censo de 2021 realizado pelo CBO, cerca de 101 municípios brasileiros com mais de 40 mil habitantes, não tem profissionais para um atendimento regular (Beniz Neto; Umbelino, 2021)

Atualmente o acompanhamento e o diagnóstico da DMRI é feito por meio de exames de imagem, em especial a retinografia, que se dá pela análise de imagens detalhadas da retina, possibilitando o profissional a identificar a presença de drusas e neovascularização. No entanto, a interpretação manual das imagens, exige um

alto grau de especialização, que torna o processo difícil e demorado, principalmente em locais onde a presença de profissionais é escassa.

Nesse cenário, surge a necessidade de novos métodos para otimizar o diagnóstico e acompanhamento da DMRI, surgindo assim, as tecnologias de aprendizado profundo (*deep learning*), que podem ser adequadas para o processamento de imagens, permitindo a classificação, de forma automática e precisa, dos diferentes tipos da DMRI. Aliviando a pressão dos especialistas e tornando mais acessível e rápido o diagnóstico nas cidades sem oftalmologistas qualificados.

Embora o *Deep Learning* seja amplamente usado para o diagnóstico da DMRI (Samagaio *et al.*, 2018; Santos, 2019), ainda há uma carência em estudos que abordem o diagnóstico através da retinografia, por isso, esse estudo busca preencher essa lacuna, já que o diagnóstico por meio da retinografia representa uma acessibilidade maior, em relação aos outros métodos.

Este estudo apresenta como principal contribuição a validação de um modelo pragmático e eficaz para a triagem automatizada da DMRI, utilizando a retinografia, um exame de custo menor, em relação ao OCT, e amplamente acessível. A pesquisa também contribui com uma análise comparativa de duas arquiteturas, DenseNet-121 e EfficientNet-B1, indicando qual obteve melhor desempenho e fornecendo uma direção para futuros trabalhos na área.

2 REFERENCIAL TEÓRICO

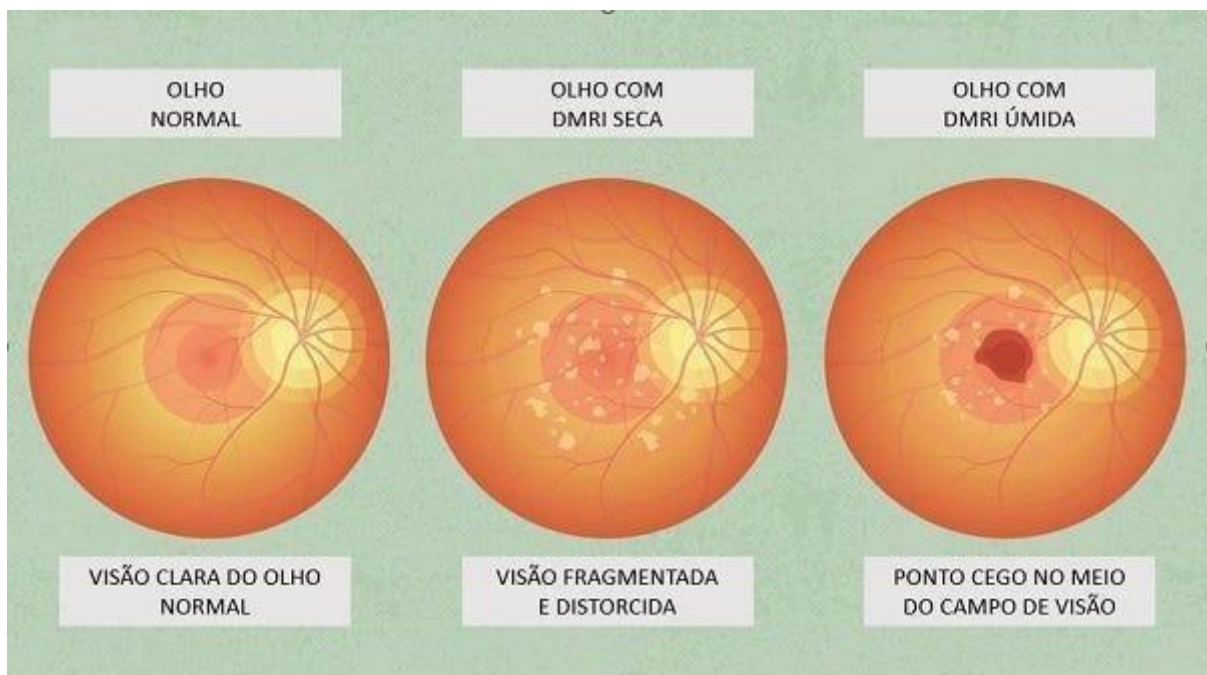
2.1 DEGENERAÇÃO MACULAR RELACIONADA À IDADE

2.1.1 Definição

A DMRI é uma doença degenerativa, progressiva e crônica, afetando a mácula, região central do olho responsável pela visão central, percepção das cores e atividades rotineiras, como a leitura e o reconhecimento de faces. Sendo a principal causa de cegueira em idosos, ela é um assunto que vem preocupando cada vez mais os estudiosos da área, já que com o envelhecimento e o aumento da expectativa de vida, a quantidade de casos da DMRI cresce exponencialmente ao redor do mundo (Torres *et al.*, 2008).

A DMRI apresenta duas formas clínicas: A Forma seca (atrófica) é a forma mais leve da doença e representa cerca de 90% dos casos, caracterizando-se pela destruição dos cones, formando pequenos depósitos de proteína e gordura no centro da retina, as drusas, podendo evoluir para um estágio de atrofia, causando a cegueira (Mehta, 2024). Já a Forma úmida (neovascular), embora menos comum, representa cerca de 10% dos casos e é a forma neovascular responsável pela maioria dos casos de cegueira, sendo caracterizada pela formação de novos vasos sanguíneos sob a mácula, que acabam inchando e se rompendo, gerando vazamento de fluidos, após a cicatrização dos vasos, o tecido cicatricial danifica permanentemente a visão dificultando a passagem de luz. Por causa disso a DMRI úmida é a forma mais agressiva da doença, podendo causar a cegueira em semanas ou meses (Mehta, 2024).

Figura 1 - Ilustração de Olho sem patologia, DMRI seca e DMRI úmida.



Fonte: (Mehta, 2024)

A figura 1 mostra claramente a diferença entre um olho saudável, com DMRI seca e úmida, possibilitando a melhor compreensão das características distintas de cada forma da doença e como isso reflete nos danos progressivos à visão.

2.1.2 Epidemiologia

Durante o último século, com o aumento exponencial da expectativa de vida, surgem novos problemas de saúde pública relacionados à idade, dentre elas se destaca a DMRI que segundo (Fabre *et al.*, 2022) cerca de 50% da população mundial com mais de 80 anos apresenta algum estágio da doença.

No Brasil a situação é igualmente alarmante, de acordo com o censo publicado pelo Conselho Brasileiro de Oftalmologia (CBO) (Umbelino; Ávila, 2023) cerca de 30% dos idosos com mais de 75 anos são diagnosticados com DMRI, configurando um grande problema para o país, sendo impulsionada pela rápida mudança na expectativa de vida. Como demonstrado no estudo publicado no site da câmara dos deputados do Brasil, estima-se que até 2050 a população idosa do Brasil triplicará, tornando-o a sexta população mais idosa do planeta, aumentando significativamente os casos de DMRI, fazendo com que o Brasil suba na lista dos países com maior incidência da doença (Camara, 2018).

Além de comprometer a visão, a doença afeta a qualidade de vida, impossibilitando o portador a desempenhar tarefas cotidianas como reconhecer rostos, dirigir, ler ou preparar refeições. Essa limitação funcional resulta na perda da autonomia, impactando negativamente a saúde mental dos portadores. Conforme demonstrado no estudo realizado por (Brody *et al.*, 2001), pacientes portadores da DMRI apresentavam sinais claros de depressão, muitas vezes causado pela dificuldade de interagir com o ambiente e com outras pessoas.

Além dos desafios psicológicos, também há os desafios para os cuidadores e familiares, que prestam suporte ao portador, aumentando a carga emocional e financeira. Neste cenário, o diagnóstico precoce se faz necessário para minimizar os impactos da doença, não só na saúde ocular, mas também no bem-estar do paciente (Brody *et al.*, 2001).

2.1.3 Métodos tradicionais de diagnóstico

Por ser uma doença crônica e sem cura, o diagnóstico precoce é essencial para controle e tratamento da patologia, atualmente os métodos mais utilizados para diagnóstico da DMRI incluem a retinografia, tomografia de coerência óptica (OCT) e angiografia fluorescente.

A retinografia é um exame que captura imagens da retina, possibilitando a visualização de sinais característicos da DMRI, como a presença de sangue na mácula, neovascularização e a presença de drusas. Por sua vez, a tomografia de coerência óptica (OCT) utiliza luz para criar imagens de alta definição das camadas da retina. Já a angiografia fluorescente é um método em que um líquido fluorescente é injetado na corrente sanguínea, servindo para o mapeamento e visualização de possíveis vazamentos ou neovascularização.

Embora eficazes, os métodos citados encontram vários empecilhos para um diagnóstico rápido e preciso, dentre eles podemos destacar o alto nível de capacitação exigido dos oftalmologistas, o alto custo dos exames e por fim a escassez dos profissionais da saúde ocular no Brasil. Como mostrado pelo Jornal da USP, a má distribuição dos oftalmologistas pelo Brasil vem dificultando o diagnóstico precoce da DMRI e de outras doenças (Fogaça, 2023). Essa desigualdade regional se intensifica em locais mais remotos do Brasil onde o acesso a um serviço especializado é limitado ou inexistente. Já que como demonstrado pelo censo do CBO, cerca de 101 municípios brasileiros não têm a presença de um profissional da saúde ocular (Beniz Neto; Umbelino, 2021).

Nesse contexto, soluções tecnológicas, como sistemas de inteligência artificial (IA) e aprendizado profundo (*deep learning*), têm o potencial de mitigar os empecilhos no diagnóstico e no tratamento.

Assim, a implementação de sistemas automatizados, na saúde, pode não só facilitar o diagnóstico da DMRI no Brasil, mas também pode democratizar o acesso à saúde ocular de qualidade no país.

2.2 APRENDIZADO DE MÁQUINA E *DEEP LEARNING*

2.2.1 Aprendizado de máquina

O aprendizado de máquina ou *machine learning* é um campo da inteligência artificial (IA) voltado para o desenvolvimento de algoritmos capazes de aprender com os dados que lhe são fornecidos. funcionando de maneira distinta da programação tradicional, em que o sistema só desempenha ações com base nas instruções programados, os sistemas de machine learning permitem com que a máquina aprenda e melhore o seu desempenho à medida que recebe mais informações.

Algoritmos de *machine learning* são desenvolvidos em 3 partes, sendo elas o pré-processamento, o treinamento e a avaliação. Na primeira fase temos o pré-processamento, é essa parte onde fica o processo de escolha do banco de dados, limpeza, definição dos objetivos do algoritmo e a segmentação dos dados definindo uma parte para o treinamento e outra para os testes (Paixão *et al.*, 2022).

Após isso, tem a fase de treinamento, que pode ocorrer de duas formas. No treinamento não supervisionado, utiliza-se um conjunto de dados não rotulados, ficando a cargo do software reconhecer padrões e agrupar os dados. Já no aprendizado supervisionado precisa que os dados estejam rotulados, em que cada dado é gravado com uma saída esperada, ficando a cargo do sistema definir qual as entradas e parâmetros possíveis para aquela saída (Ludermir, 2021).

2.2.2 Deep Learning

O *deep learning*, ou aprendizado profundo, é uma sub-área do *machine learning* que utiliza redes neurais artificiais com múltiplas camadas. Essas redes são programadas para funcionar de forma similar ao cérebro humano, fornecendo ao programa a capacidade de aprender padrões complexos e tomar decisões com base em um grande conjunto de dados, gerando assim um programa que não fica preso a linhas de códigos pré definidas por um programador (Avelar, 2025).

O *deep learning* diferencia-se das outras formas de aprendizado de máquina por sua forma de aprender diretamente de dados complexos, como imagens, vídeos e áudios, sem a necessidade de pré-definir as características específicas. Embora siga etapas similares ao *machine learning*, como pré-processamento, treinamento e

validação dos resultados, o *Deep Learning* se destaca por adotar abordagens distintas (Pereira Junior et al., 2024).

Em vez de depender da seleção manual das características de interesse, o sistema utiliza os próprios dados brutos como entrada e por meio de métodos, como *backpropagation* (retropropagação do erro), ajusta os pesos e parâmetros internos das redes neurais para alcançar os resultados esperados. Esse processo permite que o modelo aprenda diretamente com os dados, tornando-o extremamente eficaz em tarefas que envolvam padrões complexos e dados não lineares (Souza *et al.*, 2019).

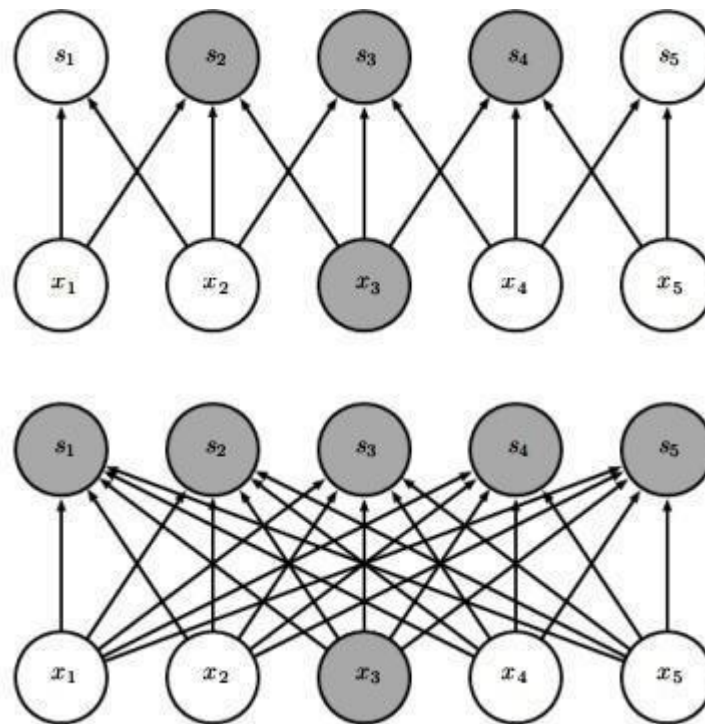
2.2.3 Redes Neurais Convolucionais

As redes neurais são a base dos sistemas de aprendizado profundo, com diversas configurações e adaptações para resolver problemas específicos. Dentre as redes neurais temos em destaque as convolucionais (CNNs - *Convolutional Neural Networks*), sendo as mais eficazes no processamento e análise de imagens (Korkmaz et al., 2024).

As CNNs são um tipo de rede neural especializada em processar dados em um tipo de topologia de grade. O seu funcionamento se dá por meio de operações matemáticas, chamadas de convolução, que exploram três ideias fundamentais para otimizar o aprendizado: interações esparsas, compartilhamento de parâmetros e representações equivariantes (Goodfellow; Bengio; Courville, 2016).

Quando falamos sobre interações esparsas estamos nos referindo a capacidade das CNNs de reduzir a complexidade computacional limitando as conexões entre as unidades de entrada e saída. Em vez de ligar cada unidade de entrada a todas unidades de saída, como ocorre nas demais redes neurais, cada unidade de entrada só está ligada a uma unidade de saída, utilizando filtros para restringir as conexões.

Figura 2 - representação gráfica das interações esparsas em CNNs



Fonte: (Goodfellow; Bengio; Courville, 2016)

Na figura 2, é ilustrada a representação gráfica de como se comportam as CNNs no que se trata a sua característica de interações esparsas. Na primeira rede neural, temos conectividade esparsas de filtro 3x3. Nesse caso, ao acionar o elemento central somente 3 elementos adjacentes são acionados, limitando as interações aos dados localizados dentro do filtro. Enquanto isso, na segunda rede neural, sem conectividade esparsa, todos os elementos são acionados, mesmo que o parâmetro de entrada seja o mesmo da rede anterior.

Isso demonstra a importância da conectividade esparsa nas CNNs, pois ao restringir as interações a um subconjunto da entrada, reduzimos significativamente os custos computacionais e de armazenamento, tornando o processamento mais eficiente. Além disso, essa abordagem melhora a capacidade do modelo de focar em padrões locais relevantes, sem desperdiçar recursos computacionais em interações desnecessárias.

Por sua vez, o compartilhamento de parâmetros é o uso do mesmo parâmetro em mais de uma função do modelo, permitindo que o mesmo conjunto de pesos seja aplicado em diferentes entradas. Essa característica reduz significativamente o

número de parâmetros armazenados, mais uma vez reduzindo os custos computacionais.

Além da redução de custos, o compartilhamento de parâmetros reduz a propensão de *Overfitting*, que é quando o modelo se ajusta tão bem aos dados de treinamento que acaba aprendendo características específicas deles, como ruídos na imagens, oq resulta em um desempenho excelente no treinamento mas pode gerar inconsistências em dados reais (Ponti; Costa, 2018).

A equivariança refere-se a uma propriedade matemática de certas funções, incluindo a convolução, que mantém consistência nas transformações da entrada e da saída. Em termos simples, uma função é considerada equivariante se, ao aplicar uma transformação na entrada, a saída também se transforma de maneira correspondente (Ponti; Costa, 2018).

2.3 PROCESSAMENTO DE IMAGENS E DMRI

2.3.1 Representação e pré-processamento de imagens

Conforme foi apresentado Chollet (2021), a representação de imagens em sistemas de *deep learning* baseia-se em tensores, estruturas de dados multidimensionais usadas para armazenar e manipular dados numéricos. Os tensores são apresentados em 4 formas, ou 'Ranks' diferentes, variando de 0 até 3, o rank 0 é a classificação de um tensor único, que armazena somente um número, o rank 1 armazena uma lista de números, o rank 2 é uma matriz de números, sendo formado por uma matriz bidimensional, e por fim, temos os tensores de rank 3, que é composto por uma matriz tridimensional (Chollet, 2021).

Por convenção, as imagens são representadas em tensores de rank 3, mesmo podendo ser representado em um rank 2, no caso de imagens em escala de cinza, já que o espaço destinado ao canal de cor sempre vai receber o valor '1'. Por motivos didáticos, consideremos uma imagem em escala de cinza com dimensões 3x3. Nesse caso, a imagem será representada como um tensor no formato channels-last, onde o índice da camada de cor é colocado na última dimensão.

Assim, a imagem será representada por um tensor com a forma (3, 3, 1), indicando 3 linhas, 3 colunas e 1 canal de cor (escala de cinza).

Partindo para o pré-processamento de imagens, etapa crucial na organização do banco para uso em *deep learning*, tem como objetivo melhorar a qualidade do banco de imagens, padronizando tamanhos, melhorando a resolução, removendo ruídos e facilitando a leitura de aspectos importantes, as principais formas de promover essas melhorias é através de 3 etapas, sendo elas, o redimensionamento, normalização dos pixels e remoção de ruídos.

O redimensionamento, uma das primeiras etapas no processo de organização do *dataset*, visa padronizar o formato e tamanho de cada imagem, buscando sempre um tamanho adequado que facilite o treinamento do modelo, mas sem perder a qualidade e preservando suas características essenciais (Rocha *et al.*, 2017).

Partindo para a normalização, nos deparamos com um método que busca ajustar a escala de cada pixel, de modo a evitar com que variáveis de escala maiores influenciam mais no treinamento. Seu objetivo de manter o desvio padrão das escalas de cores se mantenham em valores próximos de zero, promovendo uma distribuição adequada dos valores, de modo a impedir que cores em escalas maiores influenciam mais no modelo do que em escalas menores, permitindo um aprendizado mais eficiente (Rocha *et al.*, 2017).

A redução de ruído é uma das etapas mais importantes do pré-processamento de uma imagem. Ela tem como objetivo remover ou pelo menos reduzir interferências indesejadas, como qualquer tipo de distorção ou variação não que não pertence à imagem original. Esse ruído pode surgir de diferentes origens, dependendo de condições inadequadas de captura, sensores de baixa qualidade ou muita compressão (Rocha *et al.*, 2017).

2.3.2. Data augmentation

A *data augmentation*, ou técnica de aumento de dados é uma estratégia crucial para o desenvolvimento de redes convolucionais, principalmente aquelas com quantidade limitada de dados. No contexto do *deep learning* essa técnica

permite a criação de novas amostras através de modificações em imagens já existentes, aplicando procedimentos simples, como espelhamento, rotação, pequenas alterações no brilho e no contraste. Se tornando umas das principais etapas do pré-processamento de dados, pois através dela evitamos o overfitting, garantindo um sistema com maior capacidade de reconhecer padrões e robustez (Mumuni; Mumuni, 2022).

2.3.3 Uso de *deep learning* no diagnóstico da dmri

O diagnóstico da DMRI por meio do *deep learning* já vem sendo amplamente utilizado no mundo, há diversos estudos que comprovem a eficácia desses sistemas, como o estudo proposto por Santos (2019) que comprovou a eficácia das CNNs na análise das imagens provenientes de OCT para identificar sintomas da DMRI. Outro estudo relacionado foi realizado por Samagaio *et al.* (2018) que embora não esteja procurando especificamente sinais da DMRI, ainda faz análises da retina utilizando imagens vindas do OCT, e consegue identificar a presença de edema macular.

Conforme citado anteriormente, a falta de médicos capacitados e o crescente aumento da população vem criando uma enorme empecilho para o diagnóstico precoce da DMRI. Esse cenário exige soluções inovadoras e rápidas para um diagnóstico de alta qualidade, independente da localização geográfica, fatores financeiros e disponibilidade de profissionais. Nesse contexto, as tecnologias de *deep learning* emergem como uma solução promissora para enfrentar esses desafios.

Ao aplicar algoritmos de aprendizado profundo que utilizem as CNNs, é possível automatizar análises de retinografias, permitindo que o diagnóstico seja realizado de maneira rápida e precisa. O uso de retinografias para a realização do diagnóstico se justifica pelo fato de o exame ser mais barato e acessível que outros exames, como a OCT, que embora seja altamente eficaz, apresenta custos elevados e equipamentos especializados, não disponíveis em muitas regiões. Além disso, a retinografia é um exame de custo menor, em relação a OCT, e pode ser facilmente realizado em unidades de saúde de qualquer porte, tornando-se uma ferramenta

mais viável para o diagnóstico da DMRI, especialmente em locais com poucos recursos.

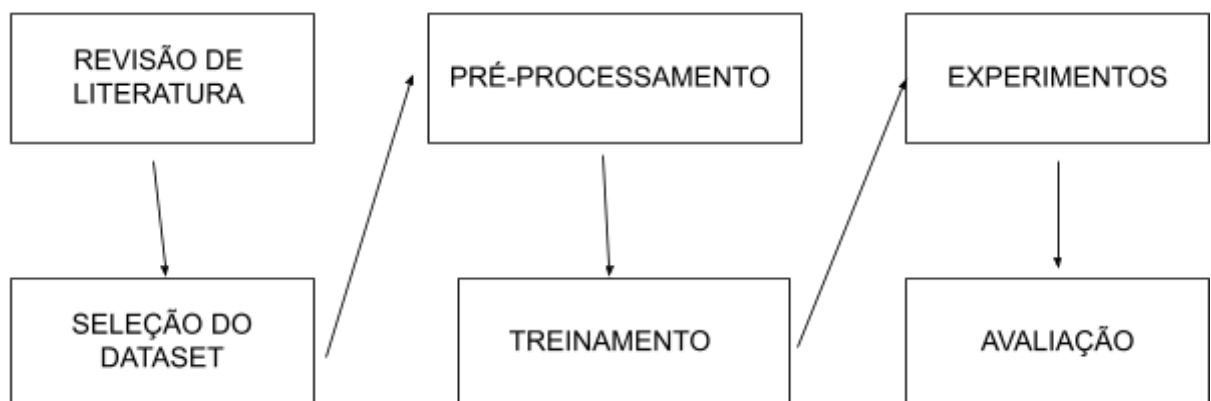
A relevância desse estudo se destaca ao considerar que, embora *deep learning* no diagnóstico da DMRI, seja um tema muito estudado, a maioria dos estudos se limita à análise de exames da OCT, enquanto os exames mais acessíveis acabam ficando de lado. Portanto o foco em retinografias, torna-se um diferencial para essa pesquisa, gerando um sistema mais acessível e aplicável em uma gama de cenários, incluindo aqueles com limitações financeiras e de infraestrutura, promovendo a democratização do diagnóstico da DMRI.

3 METODOLOGIA

A metodologia empregada neste estudo é caracterizada por uma abordagem predominantemente quantitativa e de natureza aplicada, com um delineamento experimental (Rosa et al., 2023). O estudo recebe essa classificação pois se baseia em uma análise quantitativa de uma solução prática, desenvolvida no mesmo estudo, através de uma metodologia experimental, testando hipóteses e avaliando os resultados.

Assim, a metodologia do estudo se dá em 6 etapas:

Figura 3 - Fluxograma da metodologia



Fonte: Autores (2025)

3.1 REVISÃO DA LITERATURA

A primeira etapa se dá pela realização de uma revisão sistemática com o objetivo de analisar os estudos mais importantes sobre o uso de *deep learning* no diagnóstico da DMRI em retinografias. A busca foi conduzida em bases de dados acadêmicas renomadas, como PubMed, IEEE Xplore, ScienceDirect e Scopus, utilizando palavras-chave como "DMRI", "*deep learning*", "retinografia", "classificação de doenças oculares", "redes neurais convolucionais (CNNs)", entre outras (Dong *et al.*, 2021).

A revisão teve como um de seus objetivos a identificação de limitações e desafios no âmbito do estudo proposto, a fim de construir uma base sólida para o desenvolvimento do modelo. As principais métricas usadas para avaliação do modelo também foram pesquisadas, com o objetivo de avaliar melhor o modelo que foi proposto.

3.1.1 *Deep learning* na oftalmologia

Com o constante envelhecimento da população global, tem se intensificado o impacto de doenças crônicas, com a DMRI se destacando como uma das principais causas de cegueira em idosos (Fabre *et al.*, 2022). Esta condição impõe uma sobrecarga no sistema de saúde, que se encontra defasado em diversas regiões do Brasil. A viabilidade de um diagnóstico preciso e precoce se torna mais difícil em áreas mais remotas no Brasil, onde o diagnóstico se torna um grande problema logístico e econômico (Beniz Neto; Umbelino, 2021).

Nesse a IA, em especial o *deep learning*, surgem como uma tecnologia inovadora com o objetivo de redefinir o paradigma oftalmológico. Modelos de *deep learning* demonstram uma notável capacidade de analisar imagens complexas, alcançando resultados que muitas vezes se equiparam especialistas humanos na detecção de patologias, como a DMRI (Nadeem *et al.*, 2022).

A escolha da retinografia como modalidade de imagem é uma escolha estratégica e fundamentada pela possibilidade de implementação em larga escala. Embora existam outros métodos tão eficazes quanto a retinografia, como a OCT, a retinografia se destaca pelo seu custo mais acessível. A aplicação do *deep learning*

na retinografia alinha-se com a necessidade de elaborar soluções que além de precisas sejam escaláveis e com a possibilidade de aplicar em diversos contextos clínicos, incluindo os que possuem infraestrutura limitada.

A análise de literatura revela um importante distinção entre testes feitos em laboratórios e a aplicação em ambientes clínicos. Por mais que alguns estudos apresentem resultados quase perfeitos, os algoritmos gerados ainda possuem grandes limitações. Como em toda IA, o maior gargalo para um desenvolvimento eficaz é a disponibilidade dos dados, na análise de exames médicos não seria diferente, Na análise de imagens médicas, duas limitações podem ser destacadas:

Primeiramente, a baixa variabilidade nas bases de dados de treino, a maioria das bases de dados disponíveis possuem uma diversidade de dados bem limitada, já que são construídas uma pequena pluralidade nos dispositivos de captura, tornando os algoritmos mais precisos quando recebem dados obtidos pelos dispositivos nos quais foram treinados (Dong *et al.*, 2021)

E em segundo lugar, a necessidade que as imagens obtidas, em uma aplicação prática, possua uma boa qualidade, já que variações na luz, borrões e contrastes influenciam na precisão do algoritmo.

Esses pontos reforçam que a IA como uma substituta para o trabalho de especialistas é uma ideia prematura. Em vez disso, uma abordagem mais pragmática é a de usar a IA como uma ferramenta de auxílio para o especialista. Um sistema automatizado pode analisar um grande número de imagens e sinalizar ao médico somente os pacientes que têm a patologia, assim diminuindo a carga de trabalho do especialista e garantindo que os pacientes com DMRI passem por uma segunda avaliação. Este modelo de colaboração entre homem-máquina tem potencial de democratizar o acesso diagnóstico, aliviar a sobrecarga dos especialistas e facilitar o diagnóstico precoce da DMRI.

No entanto, apesar de sua grande aplicabilidade, ainda há uma carência notável em estudos que abordem o diagnóstico da DMRI através da retinografia em cenários aplicados. Diante disso, este estudo busca preencher essa lacuna, focando no desenvolvimento e validação de um modelo para análise de retinografias.

3.1.2 Arquiteturas de redes convolucionais

As CNNs são a base para a análise de imagens médicas, pois permitem o aprendizado automático de características visuais, indo de padrões simples, como bordas e formas, até padrões mais complexos, como drusas em exames de DMRI (Cody, 2023).

A literatura demonstra uma intensa atividade de pesquisa na comparação e otimização das arquiteturas, para o diagnóstico de doenças da retina, incluindo a DMRI, a partir de retinografias. A escolha da arquitetura ideal não é trivial e envolve um compromisso entre precisão, eficiência computacional e complexidade.

A análise comparativa das arquiteturas mais adotadas revela que a escolha do modelo ideal não é algo absoluto, e sim uma decisão multifatorial, além da acurácia de teste, outros valores também entram na equação. A decisão final, da arquitetura a ser implementada baseia-se no balanço da acurácia de teste, o custo computacional e a dificuldade de implementação da arquitetura.

O estudo desenvolvido por Xia *et al.* (2024) traz uma análise das principais arquiteturas disponíveis, onde os pesquisadores alimentaram diversos algoritmos, com os dados gerados no mesmo estudo e avaliaram as principais métricas de cada modelo.

Embora algumas arquiteturas tenham se sobressaindo sobre as outras, como é o caso da DenseNet-121 (Recall 81,76%) e a InceptionV3 (Recall 81,91%) (Xia et al. 2024), a diferença entre as métricas de cada modelo não foi tão significativa, por esse motivo, a escolha das arquiteturas DenseNet-121 e EfficientNet-B1 foi baseada em seu reconhecido desempenho em tarefas de classificação de imagens médicas, conforme documentado na literatura.

A DenseNet-121 é notável por sua eficiência na reutilização de características, o que a torna robusta para *datasets* limitados, enquanto a EfficientNet-B1 oferece um excelente equilíbrio entre acurácia e eficiência computacional, um fator crucial para a implementação em larga escala.

3.1.3 Métricas de avaliação

A avaliação de um modelo de classificação não se prende ao acompanhamento de um só parâmetro, principalmente em casos que os diferentes tipos de erros têm pesos distintos. A base para toda análise profunda parte de uma matriz de confusão, que consiste em uma técnica que separa a performance em 4 classificações, os verdadeiros positivos(TP), os falsos positivos(FP), os verdadeiros negativos(TN) e os falsos negativos(FN) (Google, 2025).

A estrutura de uma matriz de confusão nos faz adotar um pensamento mais profundo do quão eficaz é o modelo. Por meio delas podemos ponderar sobre os custos que cada erro traz. A partir da matriz de confusão podemos analisar diferentes métricas, sendo elas:

A acurácia, que mede a proporção total de acertos, no geral, costuma ser o primeiro indicador a ser analisado devido a sua simplicidade. Contudo, ela também é a métrica mais enganosa em um *dataset* desbalanceado, como é comum em ambientes clínicos. Esse fenômeno pode levar um modelo a ter altíssima acurácia mas pouco valor clínico, por exemplo, ao classificar todas as imagens como normais em um ambiente com baixa prevalência da DMRI.

Uma das maneiras de contornar o problema do paradoxo da acurácia é utilizando em conjunto duas métricas, sendo elas:

A Precisão, que checa a confiabilidade das previsões positivas do modelo, essa métrica é de extrema importância em cenários onde o custo de um falso positivo é alto. E o recall, que mede a capacidade do modelo de encontrar todos os pacientes que têm algum tipo de patologia, essa métrica é crucial no nosso projeto, já que o custo de um falso negativo é catastrófico. Por esse motivo, essa métrica foi a que teve maior peso na classificação da melhor IA gerada.

3.2 IDENTIFICAÇÃO E CATALOGAÇÃO DOS *DATASETS*

A segunda etapa envolveu a identificação das principais bases de dados de retinografias utilizadas para o estudo da DMRI. A seleção do dataset se deu pela análise de 4 critérios principais e foi selecionado o que atendia melhor às exigências citadas abaixo :

1. Quantidade de imagens disponíveis: quanto mais imagens mais robusto será o modelo.
2. Qualidade das imagens: a qualidade e definição das imagens será analisada, já que a qualidade das imagens impacta na precisão do modelo.
3. Variabilidade dos dispositivos de captura: será registrado as especificações dos dispositivos de captura
4. Identificação dos tipos de DMRI: as imagens devem estar rotuladas, identificando se é o tipo seco ou úmido.

O *Dataset* utilizado no desenvolvimento do algoritmo foi disponibilizado por Xia *et al.* (2024) onde os autores capturaram um total de 28.877 imagens de retina, sendo elas distribuídas em 9 classificações diferentes. Desse conjunto de dados, usamos 15.000 imagens de olho normal e 7.095 imagens de olho com DMRI.

As imagens apresentam resoluções variadas, indo de 768 x 576 até 3264 x 2448. Embora o artigo não cita quais dispositivos foram usados nas capturas ele informa que foram utilizadas 5 câmeras diferentes, e assegura a qualidade e a variabilidade das imagens dispostas no *dataset* (Xia *et al.*, 2024).

Um ponto negativo no *dataset* é que a grande maioria das imagens contidas nele não classificam a DMRI entre seca e úmida, sendo somente 93 da DMRI seca e 92 da DMRI úmida. Devido a essa limitação, o algoritmo desenvolvido neste projeto seguiu de uma forma diferente, tendo a função de identificar a presença da patologia nos exames. Assim, os pacientes que são diagnosticados pelo algoritmo com DMRI são encaminhados para uma segunda avaliação clínica, fazendo com que o médico especialista analise somente os pacientes que de fato possuem DMRI.

Essa limitação do *dataset* exigiu uma adaptação estratégica no escopo do projeto. O foco foi redefinido para o desenvolvimento de um sistema de triagem, que cumpre um papel clinicamente valioso. Nessa abordagem, o algoritmo atua como uma ferramenta de primeira linha para separar os exames com suspeita de patologia dos normais.

3.3 DESENVOLVIMENTO DO MODELO DE *DEEP LEARNING*

Conforme a estratégia de pesquisa, nessa fase foram desenvolvidos e comparados dois modelos de rede neurais convolucionais (CNNS) para a classificação binária entre exames de retinas Normais (olhos sem nenhum tipo de patologia identificada) e com DMRI. O objetivo foi avaliar qual das arquiteturas DenseNet-121 e EfficientNet-B1, apresentaria melhor desempenho sob uma metodologia rigorosa de treinamento e validação. O treinamento e a avaliação dos modelos foram conduzidos na plataforma Google Colab, utilizando uma unidade de processamento gráfico (GPU) do tipo L4.

Para a construção dos modelos de classificação, foi empregada a técnica de aprendizado por transferências (Transfer Learning). Esta abordagem permite alavancar o conhecimento de arquiteturas de redes neurais convolucionais pré-treinadas no extenso *dataset* ImageNet.

3.3.1. Pré-processamento e preparação dos dados

Com o objetivo de padronizar os dados de entrada e otimizar o desempenho do modelo, um pipeline de pré-processamento foi implementado a todas as imagens. Inicialmente, foram redimensionadas para dimensões de entrada de cada arquitetura, sendo 224x224 para a DenseNet-121 e 240x240 para a EfficientNet-B1. Após isso, para balancear as classes, aumentar a diversidade do *Dataset* e a capacidade de generalização, um conjunto de transformações foi aplicado dinamicamente. Este processo incluiu inversões horizontais e verticais (RandomFlip), rotações em ângulos aleatórios (Randomrotation), aplicações de zoom (RandomZoom) e ajustes de contraste (RandomContrast).

Concluída essa etapa de manipulação, foi realizada a normalização dos pixels. Nesta fase, os valores de intensidade foram ajustados para o intervalo de [0,1]. Este processo é fundamental para garantir uma convergência mais rápida e estável durante o treinamento dos modelos.

3.3.2. Treinamento

Para permitir uma comparação justa e metodologicamente rigorosa entre as arquiteturas DenseNet-121 e EfficientNet-B1, uma estratégia de treinamento e validação idêntica foi aplicada a ambas. Os procedimentos foram desenhados para garantir a robustez da avaliação, lidar com o desbalanceamento dos dados e prevenir o overfitting.

Para assegurar a robustez estatística, foi empregada a validação cruzada k-fold utilizando uma semente aleatória, para garantir que ambas arquiteturas sejam treinadas com a mesma distribuição de dados.

Em ambas arquiteturas, os modelos foram iniciados com pesos pré-treinados da ImageNet e o treinamento focou no ajuste fino. Na DenseNet-121, foi descongelado as camadas a partir do bloco "conv5_block1_0_bn", enquanto na EfficientNet-B1 o treinamento de seu a partir do "block6a_expand_bn". Nessa fase foi utilizado uma taxa de aprendizado baixa " 5×10^{-6} " para ambos os modelos, as duas arquiteturas foram configuradas para rodar um máximo de 80 épocas. A fim mitigar o overfitting, forma empregada uma técnica de a callback, que monitorou a acurácia de validação e colocava uma condição de parada, caso a métrica não melhorasse nas últimas 10 épocas, salvando assim o modelo com maior acurácia.

Nas duas arquiteturas, foi utilizado a função de "BinaryCrossentropy", recebendo valores ponderados, com a DMRI recebendo peso 2 e o normal recebendo peso 1. Ao final, o desempenho de cada fold foi avaliado com um conjunto de métricas incluindo, acurácia, AUC, precisão e recall.

4 RESULTADOS E DISCUSSÃO

Esta seção apresenta uma análise detalhada e interpretativa dos resultados obtidos no treinamento e na avaliação dos modelos DenseNet-121 e EfficientNet-B1. O objetivo é dissecar o desempenho individual de cada arquitetura antes de realizar uma análise comparativa, contextualizando os achados no âmbito da tarefa de classificação de imagens oftálmicas para o diagnóstico de DMRI.

4.1. COMPARAÇÃO ENTRE ARQUITETURAS.

A avaliação das arquiteturas não se limitou a uma simples análise de acurácia e perda (loss). Após o treinamento, os 5 *folds* de cada arquitetura foram submetidos a um *dataset* de teste que compreende um total de 4.419 imagens. O resultado de cada fold foi analisado por meio de uma matriz de confusão e, por fim, o melhor *fold* de cada arquitetura foi selecionado. Essa seleção se deu pela avaliação criteriosa do perfil de erro de cada modelo, priorizando a métrica de maior impacto clínico para uma ferramenta de triagem.

Considerando o contexto de diagnóstico médico, um falso negativo (um paciente com DMRI classificado como saudável) representa um erro muito mais grave do que um falso positivo (um paciente saudável encaminhado para exames adicionais). Portanto, o critério de seleção para o "melhor fold" foi a maximização do recall (sensibilidade), que mede a capacidade do modelo de identificar corretamente os casos positivos, minimizando assim o risco de diagnósticos perdidos.

A tabela a seguir apresenta uma comparação direta das principais métricas de desempenho para o melhor fold de cada arquitetura, permitindo uma análise objetiva de seus pontos fortes e fracos.

Tabela 1 - Métricas de Avaliação

MÉTRICAS	DENSENET-121	EFFICIENTNET-B1
ACURÁCIA	76,06%±0,052%	80,06%±0,043%
PRECISÃO	61,62%±0,010%	67,67%±0,089%
RECALL	67,44%±0,010%	72,58%±0,013%

Fonte: Autores (2025)

A análise comparativa dos resultados, com os dados atualizados, revela um desempenho consistentemente superior da arquitetura EfficientNet-B1 em todas as métricas avaliadas. O modelo EfficientNet-B1 superou o DenseNet-121 em acurácia

(80,06% contra 76,06%), precisão (67,67% contra 61,62%) e, mais criticamente, no recall (72,58% contra 67,44%).

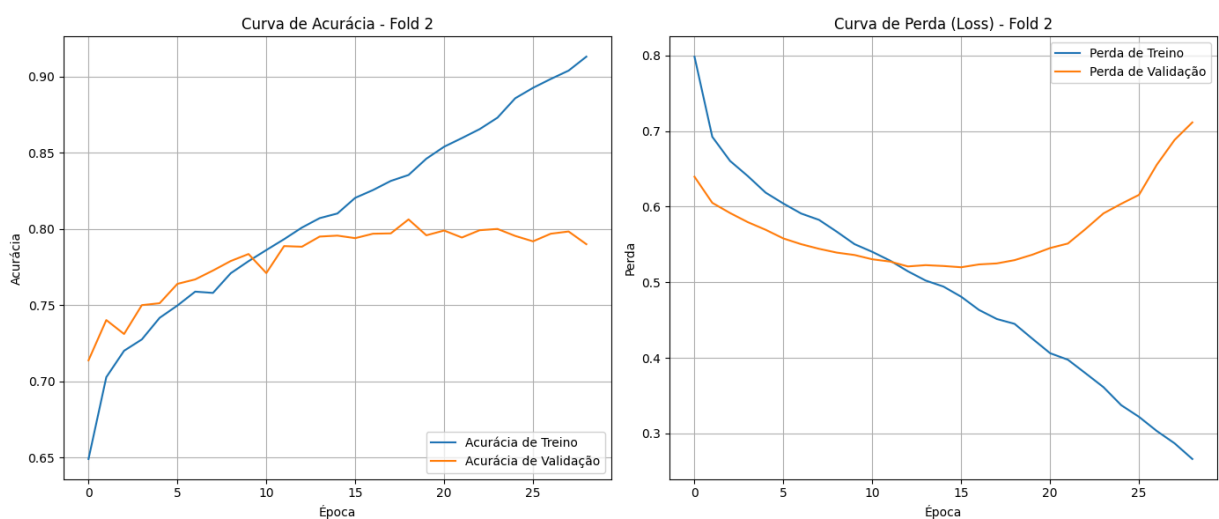
A métrica de maior importância para um sistema de triagem de doenças como a DMRI é o recall. Um recall mais alto indica que o modelo é mais eficaz em identificar corretamente os pacientes que de fato possuem a doença, minimizando assim a quantidade de falsos negativos.

A diferença de mais de 5 pontos percentuais no recall a favor da EfficientNet-B1 (72,58% vs. 67,44%) é clinicamente significativa e fundamental para a segurança do paciente.

Portanto, conclui-se que a EfficientNet-B1 se mostrou a melhor arquitetura para o propósito de uma ferramenta de triagem. Seu desempenho superior em todas as métricas, com destaque especial para o recall mais elevado, alinha-se diretamente com a principal exigência de um sistema de triagem clínica: priorizar a detecção de casos positivos e minimizar o risco de não diagnosticar um paciente doente.

4.2 AVALIAÇÃO DE DESEMPENHO

Figura 4 - Acurácia e Los da DenseNet-121



Fonte: Autores (2025)

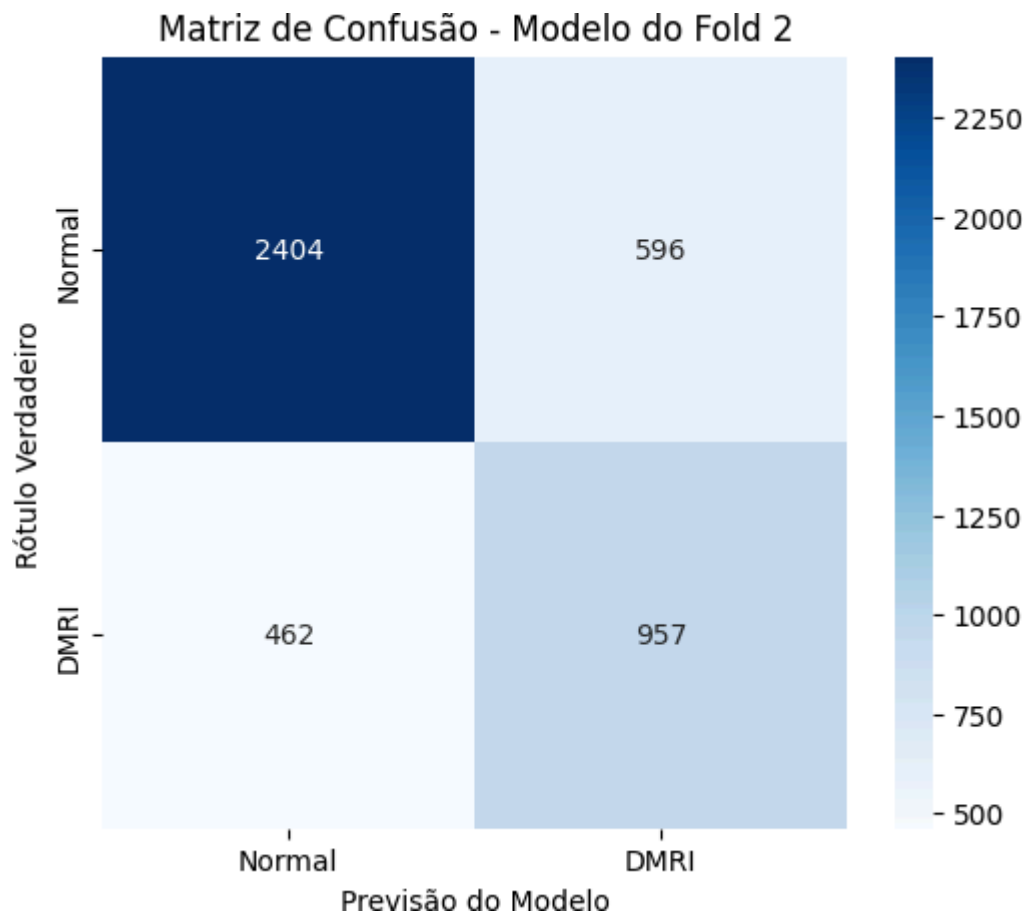
A análise da curva de aprendizado da DenseNet-121, apresentada na figura acima, oferece uma visão clara da sua curva de aprendizado ao longo de 29 épocas, a curva ascendente da acurácia treino mostra um aprendizado consistente do modelo, chegando a pouco mais de 90%. Em contraste, temos a curva de validação, que tem progresso até a época 19, onde atinge um platô, estabilizando-se em uma acurácia de 80%, não apresentando melhorias significativas após isso.

Uma observação mais crítica emerge da análise das curvas de perda. A perda de treino decresce de forma contínua e acentuada, atingindo um valor mínimo próximo a 0,25. Contudo, a perda de validação atinge seu valor mínimo, de aproximadamente 0,5, por volta da época 17 e, a partir daí, inverte sua trajetória, iniciando uma grande tendência de subida.

Esse comportamento demonstra um sinal claro de overfitting, o modelo continua a otimizar o seu desempenho no conjunto de dados de treino “memorizando” suas características, esse comportamento é prejudicial para o modelo, pois embora as taxas de acerto no treino continuem subindo, o modelo perde significativamente a sua capacidade de generalizar, perdendo assim a sua eficácia para imagens nunca vistas. ponto ótimo de treinamento, em termos de performance generalizável, ocorreu no momento em que a perda de validação foi mínima. A continuação do treinamento para além desse ponto mostrou-se contraproducente, pois o modelo começou a se ajustar ao ruído e às particularidades do conjunto de treino, perdendo robustez.

No entanto, este potencial problema foi efetivamente gerenciado pela utilização de uma callback de parada antecipada (early stopping). Ao fazer isso, a callback garantiu que a versão salva do modelo fosse aquela com a maior acurácia de validação, antes que os efeitos negativos do overfitting pudessem degradar sua performance. A análise da matriz de confusão (Figura 5) reflete, assim, o desempenho prático desta versão otimizada do DenseNet-121, e não de um modelo super ajustado.

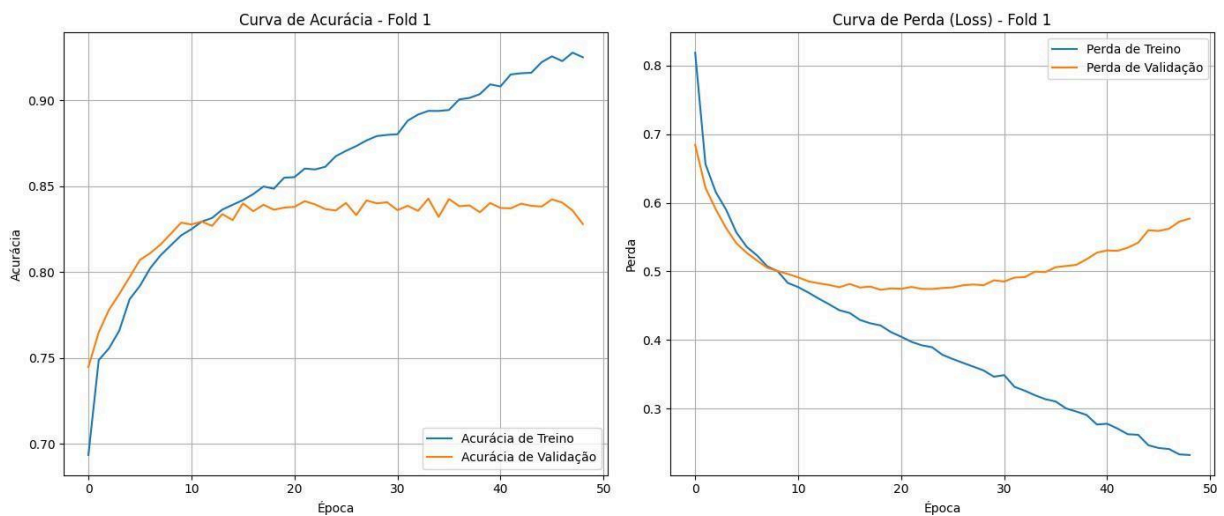
Figura 5 - Matriz de confusão da DenseNet-121



Fonte: Autores (2025)

A partir dos resultados apresentados na figura acima, as principais métricas de avaliação foram extraídas, a Acurácia geral do modelo ficou em 76,06%, a precisão foi de 61,62%, indicando que de todas as vezes que o modelo disse que era DMRI, ele estava certo em 61% das vezes, e por fim, a métrica de maior valor clínico, o recall, foi de 67,44%, o que significa que de todas as imagens que tinha de DMRI, o modelo conseguiu encontrar 67%.

Figura 6 - Acurácia e Los da EfficientNet-B1

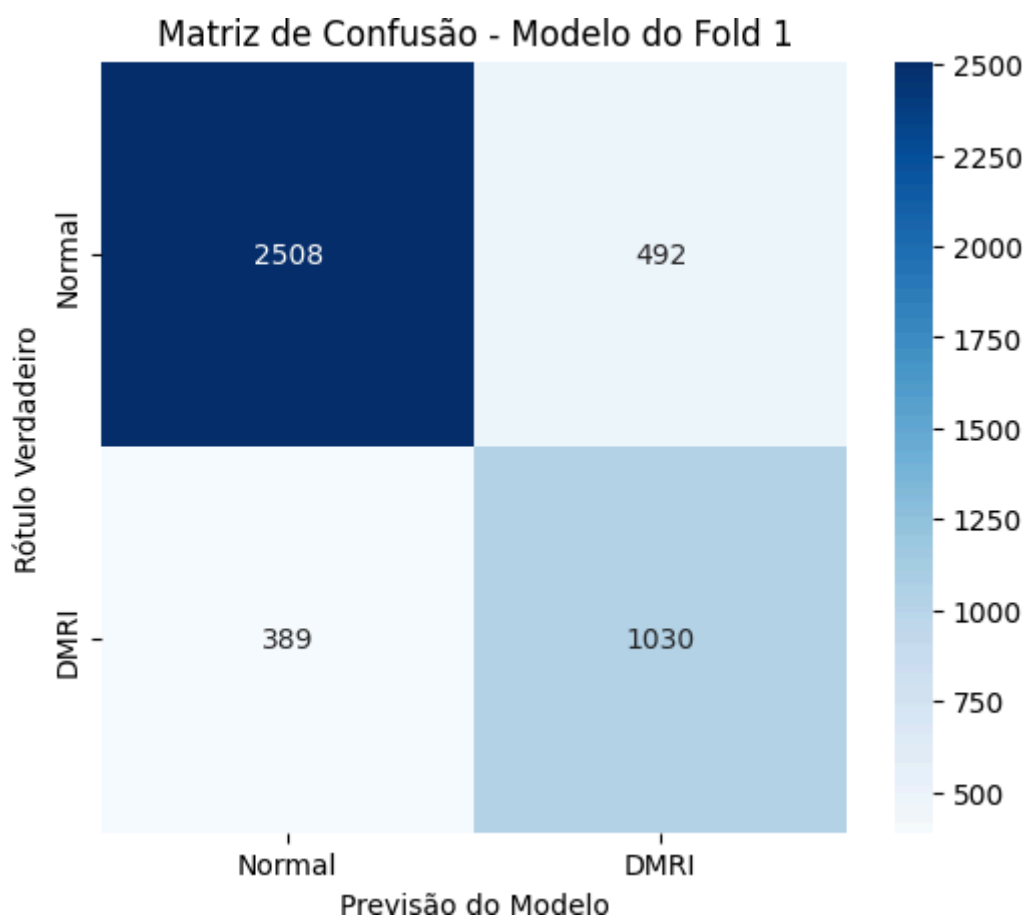


Fonte: Autores (2025)

Partindo para a EfficientNet-B1 (Figura 6), percebe-se que a curva da acurácia de teste tem uma ascensão mais agressiva, subindo rapidamente até a 10ª época, após isso a acurácia de teste tem uma estabilizada mas continua evoluindo até chegar na sua 34ª época, atingindo assim seu platô e estagnado com uma acurácia próxima a 85%.

Analisando o gráfico de perda (Loss), observa-se um comportamento que explica essa estagnação. A curva de perda de treino (em azul) apresenta uma queda acentuada e contínua ao longo de todo o processo, indicando que o modelo se ajusta cada vez mais aos dados de treinamento. Em contrapartida, a curva de perda de validação (em laranja) atinge seu valor mínimo por volta da 20ª época e, a partir desse ponto, inverte sua trajetória, começando a subir de forma clara. Essa divergência entre as curvas de perda de treino e validação é um forte indicativo de overfitting. Como as técnicas de mitigar o overfitting foram implementadas de maneira igual na Densenet e na EfficientNet-B1, o overfitting também não foi um problema nessa arquitetura.

Figura 7 - Matriz de confusão EfficientNet-B1



Fonte: Autores (2025)

A análise da matriz de confusão para o Fold 1 do modelo EfficientNet-B1, demonstrada na figura, permite uma avaliação detalhada do seu desempenho no conjunto de testes. O modelo classificou corretamente 1.030 casos como DMRI (Verdadeiros Positivos) e 2.508 como Normais (Verdadeiros Negativos). A acurácia do modelo ficou em 80,06%, a precisão foi de 67,67%. Isso indica que, quando o algoritmo identificou um exame como "Com DMRI", ele estava correto em aproximadamente 68% das vezes, e por fim o recall atingiu um patamar de 72,58%, isso significa que de todos os casos de DMRI o modelo conseguiu classificar corretamente 72,58%.

5 CONSIDERAÇÕES FINAIS

Este trabalho teve como objetivo central o desenvolvimento e a avaliação de um sistema de inteligência artificial, baseado em *deep learning*, para a detecção da DMRI a partir de imagens de retinografia. A motivação para este estudo emerge de um cenário de saúde pública desafiador no Brasil, caracterizado pelo envelhecimento populacional e por uma acentuada desigualdade na distribuição de oftalmologistas, que dificulta o diagnóstico precoce da DMRI, uma das principais causas de cegueira irreversível em idosos. O projeto propôs-se a explorar uma solução tecnológica que fosse não apenas precisa, mas também acessível e escalável, optando estrategicamente pela retinografia como modalidade de imagem.

A metodologia envolveu a comparação rigorosa de duas arquiteturas de Redes Neurais Convolucionais (CNNs) de ponta, DenseNet-121 e EfficientNet-B1, utilizando técnicas de aprendizado por transferência, aumento de dados e validação cruzada para garantir a robustez dos resultados. A análise comparativa revelou um desempenho superior da arquitetura EfficientNet-B1 em todas as métricas. Com um *recall* de 72.58%, a EfficientNet-B1 demonstrou uma maior capacidade de identificar corretamente os casos positivos de DMRI, em comparação com os 67,44% da DenseNet-121.

A principal contribuição deste estudo é a validação de um modelo pragmático e eficaz para um sistema de triagem automatizada da DMRI. Ao focar na retinografia, amplamente disponível no território brasileiro, a solução proposta apresenta um alto potencial de impacto social. O sistema pode atuar como uma poderosa ferramenta de primeira linha, especialmente em regiões remotas e carentes de especialistas, otimizando o tempo dos oftalmologistas ao direcionar sua atenção para os casos que realmente necessitam de uma avaliação aprofundada. Dessa forma, o trabalho contribui para a democratização do acesso à saúde ocular de qualidade e para a mitigação dos impactos da DMRI na população brasileira.

Apesar dos resultados promissores, é fundamental reconhecer as limitações deste estudo, que abrem caminhos claros para pesquisas futuras. A primeira limitação reside na incapacidade do modelo de diferenciar entre os subtipos de DMRI (seca e úmida), devido ao severo desbalanceamento de classes no *dataset* utilizado. Uma segunda limitação é a validação do modelo em um único conjunto de

dados, o que não garante sua generalização para imagens capturadas com diferentes equipamentos e em diversas populações.

Por fim, como a maioria dos modelos de *deep learning*, a solução opera como uma "caixa-preta", sem oferecer transparência sobre seu processo de decisão, o que representa uma barreira para a confiança e a adoção clínica.

Diante disso, um roteiro para trabalhos futuros pode ser traçado em três frentes principais. Primeiramente, é necessário o enriquecimento do conjunto de dados, buscando ou construindo uma base maior e mais equilibrada que possibilite o treinamento de um classificador multiclasse robusto.

Em segundo lugar, a validação multicêntrica do modelo em *datasets* de diferentes instituições de saúde brasileiras é essencial para testar e comprovar sua generalização e robustez no mundo real.

Finalmente, a etapa mais transformadora seria a integração de técnicas de Inteligência Artificial Explicável (XAI), como o Grad-CAM, para gerar mapas de ativação que visualizem as regiões da retina que mais influenciaram a predição do modelo. Essa transparência é crucial para construir a confiança dos profissionais de saúde e transformar o sistema de um classificador para uma verdadeira ferramenta de apoio à decisão clínica.

REFERÊNCIAS

AVELAR, E. A. AJUDANDO A “DESMITIFICAR” A INTELIGÊNCIA ARTIFICIAL EM EDUCAÇÃO: A ABORDAGEM DO APRENDIZADO DE MÁQUINA. **Educação em Revista**, v. 41, n. e58030, 2025.

BENIZ NETO, J.; UMBELINO, C. C. **CENSO CBO 2021**. [s.l.] CBO, 2021. Disponível em: https://cbo.net.br/admin/docs_upload/034327CensoCBO2021.pdf. Acesso em: 15 dez. 2024.

BRODY, B. L. et al. Depression, visual acuity, comorbidity, and disability associated with age-related macular degeneration. **Ophthalmology**, v. 108, n. 10, p. 1893–900; discussion 1900–1, out. 2001.

CÂMARA. **População idosa no Brasil pode triplicar até 2050 e se tornar a sexta maior do mundo** - Painel Eletrônico - Rádio Câmara. 2018. Disponível em: <https://www.camara.leg.br/radio/programas/541513-populacao-idosa-no-brasil-pode-triplicar-ate-2050-e-se-tornar-a-sexta-maior-do-mundo/>. Acesso em: 4 jul. 2025.

CHOLLET, F. **Deep learning with python, second edition**. New York, NY: Manning Publications, 2021.

CODY, E. **Redes neurais convolucionais na classificação de imagens médicas**. 2023. Disponível em:

<https://profissaobiotech.com.br/redes-neurais-convolucionais-classificacao-imagens-medicas/>. Acesso em: 5 jul. 2025.

DONG, L. et al. Artificial intelligence for the detection of age-related macular degeneration in color fundus photographs: A systematic review and meta-analysis. **EClinicalMedicine**, v. 35, n. 100875, p. 100875, maio 2021.

FABRE, M. et al. Recent advances in age-related macular degeneration therapies. **Molecules (Basel, Switzerland)**, v. 27, n. 16, p. 5089, 10 ago. 2022.

FOGAÇA, A. B. **Dificuldade no acesso a oftalmologistas é um empecilho para o uso de óculos no Brasil**. 2023. Disponível em: <https://jornal.usp.br/atualidades/dificuldade-no-acesso-a-oftalmologistas-e-um-empecilho-para-o-uso-de-oculos-no-brasil/>. Acesso em: 5 jul. 2025.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. London, England: MIT Press, 2016.

GOOGLE. **Classificação: precisão, recall, precisão e métricas relacionadas**. 2025. Disponível em: <https://developers.google.com/machine-learning/crash-course/classification/accuracy-precision-recall?hl=pt-br>. Acesso em: 5 jul. 2025.

KORKMAZ, M. et al. Redes neurais convolucionais no diagnóstico de mielopatia cervical. **Revista brasileira de ortopedia**, v. 59, n. 5, p. e689–e695, out. 2024.

LUDERMIR, T. B. Inteligência Artificial e Aprendizado de Máquina: estado atual e tendências. **Estudos Avançados**, v. 35, n. 101, p. 85–94, abr. 2021.

MEHTA, S. **Degeneração macular relacionada à idade (DMRI)**. 2024. Disponível em: <https://www.msmanuals.com/pt/casa/distúrbios-oftalmológicos/doenças-da-retina/degeneração-macular-relacionada-à-idade-dmri>. Acesso em: 4 jul. 2025.

MUMUNI, A.; MUMUNI, F. Data augmentation: A comprehensive survey of modern approaches. **Array (New York, N.Y.)**, v. 16, n. 100258, p. 100258, dez. 2022.

NADEEM, M. W. et al. Deep learning for diabetic retinopathy analysis: A review, research challenges, and future directions. **Sensors (Basel, Switzerland)**, v. 22, n. 18, p. 6780, 8 set. 2022.

PAIXÃO, G. M. DE M. et al. Machine Learning na Medicina: Revisão e Aplicabilidade. **Arquivos brasileiros de cardiologia**, v. 118, n. 1, p. 95–102, jan. 2022.

PEREIRA JUNIOR, W. M. et al. Avaliação da presença de fissuras em imagens de estruturas de concreto através do uso de redes neurais profundas. **Matéria (Rio de Janeiro)**, v. 29, n. 4, 2024.

PONTI, M. A.; COSTA, G. B. P. DA. **Como funciona o Deep Learning**. 20 jun. 2018. Disponível em: https://sites.icmc.usp.br/moacir/papers/Ponti_Costa_Como-funciona-o-Deep-Learning_2017.pdf. Acesso em: 8 jul. 2025.

ROCHA, R. L. et al. Avaliação de técnicas de Deep Learning aplicadas à identificação de peças defeituosas em vagões de trem. *In: CONFERENCE ON GRAPHICS*,

PATTERNS AND IMAGES (SIBGRAPI), 30., 2017, Niterói. **Anais do Workshop de Indústria (WIA)**. São José dos Campos: INPE, 2017. p. 1-8. Disponível em: <http://sibgrapi.sid.inpe.br/attachment.cgi/sid.inpe.br/sibgrapi/2017/09.26.14.23/doc/workshopsibgrapi2017.pdf>. Acesso em: 8 jul. 2025.

ROSA, A. C. F. DA et al. Uso de técnicas de aprendizado de máquina para classificação de fatores que influenciam a ocorrência de dermatites ocupacionais. **Revista Brasileira de Saúde Ocupacional**, v. 48, n. e4, 2023.

SAMAGAIO, G. et al. Automatic macular edema identification and characterization using OCT images. **Computer methods and programs in biomedicine**, v. 163, p. 47–63, set. 2018.

SANTOS, A. M. **Diagnóstico de Degeneração Macular Relacionada à Idade a partir de Tomografia de Coerência Óptica usando Geoestatística e Capsule Networks**. 2019. Tese (Doutorado em Engenharia de Eletricidade) - Programa de Pós-Graduação em Engenharia de Eletricidade, Universidade Federal do Maranhão, São Luís, 2019.

SOUZA, E. P. DE et al. Aplicações do Deep Learning para diagnóstico de doenças e identificação de insetos vetores. **Saúde em Debate**, v. 43, n. spe2, p. 147–154, nov. 2019.

TORRES, R. J. DE A. et al. Conceitos atuais e perspectivas na prevenção da degeneração macular relacionada à idade. **Revista brasileira de oftalmologia**, v. 67, n. 3, p. 142–155, jun. 2008.

UMBELINO, C. C.; ÁVILA, M. P. DE. **As condições de saúde ocular**. [s.l.] Conselho Brasileiro de Oftalmologia (CBO), 2023. Disponível em: https://www.cbo.net.br/admin/docs_upload/Condicoesdesaudeocularnobrasil.pdf. Acesso em: 7 jul. 2025.

XIA, X. et al. Benchmarking deep models on retinal fundus disease diagnosis and a large-scale dataset. **Signal processing. Image communication**, v. 127, n. 117151, p. 117151, set. 2024.