
Redes neurais para classificação de testes de eletroforese de imunofixação

Alexandre de Cássio Vilarinho Filho



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Alexandre de Cássio Vilarinho Filho

**Redes neurais para classificação de testes de
eletroforese de imunofixação**

Dissertação de mestrado apresentada ao Programa de Pós-graduação da Faculdade de Computação da Universidade Federal de Uberlândia como parte dos requisitos para a obtenção do título de Mestre em Ciência da Computação.

Área de concentração: Ciência da Computação

Orientador: Prof. Dr. Leandro Nogueira Couto

Coorientador: Prof. Dr. Jefferson Rodrigo de Souza

Uberlândia

2025

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU
com dados informados pelo(a) próprio(a) autor(a).

V697 Vilarinho Filho, Alexandre de Cássio, 1998-
2025 Redes neurais para classificação de testes de eletroforese de
imunofixação [recurso eletrônico] / Alexandre de Cássio Vilarinho
Filho. - 2025.

Orientador: Leandro Nogueira Couto.

Coorientador: Jefferson Rodrigo de Souza.

Dissertação (Mestrado) - Universidade Federal de Uberlândia,
Pós-graduação em Ciência da Computação.

Modo de acesso: Internet.

DOI <http://doi.org/10.14393/ufu.di.2025.475>

Inclui bibliografia.

Inclui ilustrações.

1. Computação. I. Couto, Leandro Nogueira, 1986-, (Orient.). II.
Souza, Jefferson Rodrigo de, 1985-, (Coorient.). III. Universidade
Federal de Uberlândia. Pós-graduação em Ciência da Computação.
IV. Título.

CDU: 681.3

Bibliotecários responsáveis pela estrutura de acordo com o AACR2:

Gizele Cristine Nunes do Couto - CRB6/2091

Nelson Marcos Ferreira - CRB6/3074



ATA DE DEFESA - PÓS-GRADUAÇÃO

Programa de Pós-Graduação em:	Ciência da Computação				
Defesa de:	Dissertação, 17/2025, PPGCO				
Data:	30 de Julho de 2025	Hora de início:	09:00	Hora de encerramento:	10:10
Matrícula do Discente:	12122CCP001				
Nome do Discente:	Alexandre de Cássio Vilarinho Filho				
Título do Trabalho:	Redes neurais para classificação de testes de eletroforese de imunofixação				
Área de concentração:	Ciência da Computação				
Linha de pesquisa:	Ciência de Dados				
Projeto de Pesquisa de vinculação:	-----				

Reuniu-se por videoconferência, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Ciência da Computação, assim composta: Professores Doutores: Jefferson Rodrigo de Souza - FACOM/UFU (Coorientador), Marcelo Zanchetta do Nascimento - FACOM/UFU, Thomas Kauê Dal'Maso Peron - ICMC/USP e Leandro Nogueira Couto - FACOM/UFU, orientador do candidato.

Os examinadores participaram desde as seguintes localidades: Thomas Kauê Dal'Maso Peron - São Carlos/SP. Os outros membros da banca e o aluno participaram da cidade de Uberlândia.

Iniciando os trabalhos o presidente da mesa, Prof. Dr. Leandro Nogueira Couto, apresentou a Comissão Examinadora e o candidato, agradeceu a presença do público, e concedeu ao Discente a palavra para a exposição do seu trabalho. A duração da apresentação da Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor presidente concedeu a palavra, pela ordem sucessivamente, aos examinadores, que passaram a arguir ao candidato. Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o candidato:

Aprovado

Esta defesa faz parte dos requisitos necessários à obtenção do título de Mestre.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

*Este trabalho é dedicado ao meu filho, Henrique,
uma das razões pelas quais sigo firme diante das adversidades da vida.*

Agradecimentos

Aos meus pais, cuja dedicação, trabalho árduo e renúncias foram fundamentais para que eu pudesse alcançar este marco inédito em nossa família: a conclusão do mestrado. Sua criação ímpar e seu apoio constante me proporcionaram a força necessária para seguir adiante. Aos meus avós, meu porto seguro, exemplo de vida e fonte inesgotável de amor, expresso minha eterna gratidão. Também agradeço ao meu irmão Fellipe, meu amigo de toda a vida, e ao meu irmão Henrique, que, mesmo não estando mais entre nós, continua vivo em meu coração e nas lembranças.

À minha esposa, Mirian, que foi e é meu alicerce em todos os momentos. Sua força diária, apoio às minhas ideias e paciência ao ouvir minhas reflexões e reclamações foram indispensáveis. Sem o seu suporte incondicional, este trabalho não teria sido possível.

Ao meu filho, Henrique, cuja simples existência me ensinou o verdadeiro significado do amor incondicional e que diariamente me inspira a ser uma pessoa melhor.

Aos amigos e demais familiares, que, de diferentes formas, contribuíram para esta etapa tão significativa da minha vida, deixo meu sincero reconhecimento.

Não poderia deixar de agradecer aos meus orientadores, que acreditaram neste trabalho desde o início. Sua confiança, orientações e suporte contínuo foram essenciais para o desenvolvimento e a concretização deste projeto.

Aos colegas de trabalho, professores e a todas as pessoas que cruzaram meu caminho durante esta jornada, agradeço pelas trocas de experiências e pelos aprendizados compartilhados. Cada vivência, direta ou indireta, contribuiu para o meu crescimento pessoal e acadêmico.

Por fim, meu agradecimento especial vai aos alunos que tive a oportunidade de ensinar. Vocês reacenderam em mim a paixão pela docência e me mostraram que este é o caminho que quero seguir. Foi por meio de vocês que recuperei a fé em um futuro melhor e na capacidade transformadora da educação.

*“E no fim, o amor que você recebe é igual ao amor que você dá”
(Paul McCartney)*

Resumo

Anualmente, um único laboratório pode realizar milhares de testes de imunofixação. A identificação da presença ou ausência de proteínas na amostra é geralmente feita visualmente, exigindo a expertise de um especialista. O diagnóstico correto é essencial, pois condições potencialmente prejudiciais, como a gamopatia monoclonal, podem ser diagnosticadas a partir do exame. Este estudo propõe dois métodos baseados em modelos de *Machine Learning* (Aprendizado de Máquina) (ML) para classificar exames de imunofixação. Os métodos utilizam uma Convolutional Neural Network (Rede Neural Convolucional) (CNN) e uma rede neural densa com um vetor de características desenvolvido manualmente. Além disso, é proposto um método de pré-processamento para dividir a imagem em tiras verticais, que atuam como regiões de interesse. Este método utiliza o histograma de intensidade de cor das imagens para dividir a imagem. Os modelos foram validados usando um novo banco de dados que contém imagens desafiadoras de imunofixação por eletroforese. Os resultados obtidos com a rede *Visual Geometry Group* (Grupo de Geometria Visual) (VGGNet) foram um F1-Score de 0,9088, precisão de 0,9123 e recall de 0,9060. Por outro lado, os resultados da rede neural densa foram um F1-Score de 0,9625, precisão de 0,9628 e recall de 0,9621. Ambos os modelos de ML demonstraram excelente desempenho na classificação de imagens de exames de imunofixação. Além disso, discutimos por que a rede neural densa alcançou métricas semelhantes ou melhores em comparação com o estado da arte disponível.

Palavras-chave: Eletroforese de imunofixação. Classificação de imagens médicas. Rede neural..

Abstract

Annually, a single laboratory can perform thousands of immunofixation tests. The identification of the presence or absence of proteins in the sample is usually done visually, requiring the expertise of a specialist. A correct diagnosis is essential, as potentially harmful conditions, such as monoclonal gammopathy, can be diagnosed through the test. This study proposes two methods based on Machine Learning (ML) models to classify immunofixation tests. The methods use a Convolutional Neural Network (CNN) and a regular dense network with a manually engineered feature vector. Additionally, a preprocessing method is proposed to divide the image into vertical strips, which act as regions of interest. This method uses the color intensity histogram of the images to segment the image. The models were validated using a new database containing challenging immunofixation electrophoresis images. The results obtained with the VGGNet network were an F1-Score of 0.9088, precision of 0.9123, and recall of 0.9060. On the other hand, the results of the regular neural network were an F1-Score of 0.9625, precision of 0.9628, and recall of 0.9621. Both ML models demonstrated excellent performance in classifying immunofixation test images. Furthermore, we discuss why the regular neural network achieved similar or better metrics compared to the available state-of-the-art methods.

Keywords: Immunofixation electrophoresis. Medical image classification. Neural network..

Lista de ilustrações

- Figura 1 – Arquitetura de uma Rede Neural Artificial (RNA) com camada de entrada (*input layer*), camadas ocultas (*hidden layers*) e camada de saída (*output layer*), ilustrando o fluxo de informação entre os neurônios. Fonte: www.deeplearningbook.com.br/a-arquitetura-das-redes-neurais/ 32
- Figura 2 – Diagrama ilustrando uma CNN com camadas de entrada, convolucionais, densas e de saída, usadas para extração de características e classificação em tarefas de visão computacional. Fonte: <https://www.sakurai.dev.br/cnn-mapreduce/> 33
- Figura 3 – Estrutura da VGG16: camadas convolucionais (verde), *pooling* (roxo), totalmente conectadas (verde claro) e *softmax* (vermelho). A arquitetura reduz a dimensionalidade espacial e extrai características para classificação. Fonte: <https://www.geeksforgeeks.org/computer-vision/vgg-16-cnn-model/> 34
- Figura 4 – Representação esquemática da arquitetura U-Net. À esquerda, o caminho de contração (*encoder*) captura o contexto da imagem, enquanto à direita, o caminho de expansão (*decoder*) recupera a informação espacial para a segmentação. As setas horizontais representam as conexões de atalho (*skip connections*), que são fundamentais para preservar padrões locais e sua localização. Fonte: <https://medium.com/data-science/u-net-explained-understanding-its-image-segmentation-architecture-56e4842e313a> 35
- Figura 5 – Gel de agarose apresentando o perfil eletroforético da amostra (ELP) à esquerda, seguido pelas trilhas de imunoprecipitação. A reação com antissoros específicos revelou uma banda de precipitação bem definida nas trilhas correspondentes à cadeia pesada G (IgG) e à cadeia leve K (*kappa*). 37

Figura 6 – O diagrama ilustra as etapas do pré-processamento das imagens de imunofixação, incluindo a imagem do exame, redimensionamento, detecção da marcação "ELP" e segmentação das bandas para posterior análise.	42
Figura 7 – Eletroforese de proteínas séricas demonstrando um padrão de gamopatia monoclonal. (A) O traçado densitométrico exibe um pico estreito e proeminente na região gama (indicado pelo asterisco), conhecido como componente monoclonal (ou Proteína M). (B) O gel de agarose correspondente mostra uma banda densa e restrita na mesma posição (asterisco), confirmando a presença da proteína monoclonal. Fonte: https://emedicine.medscape.com/article/2087113-overview	49
Figura 8 – Gráfico das curvas de intensidade dos pixels de cada coluna das imagens do exame de imunofixação. Essa visualização foi possibilitada pela técnica implementada para a extração de características de cada coluna de proteína do exame.	49
Figura 9 – O diagrama ilustra o processo de extração do vetor de características a partir das faixas verticais. As etapas incluem a aquisição da imagem, seleção da coluna central de pixels e obtenção dos valores extraídos, que são utilizados no treinamento da rede neural densa.	50
Figura 10 – Amostra de imagens do banco de imagens utilizado.	55
Figura 11 – Matriz de confusão dos resultados obtidos no conjunto de teste utilizando a rede neural VGG16.	56
Figura 12 – Matriz de confusão dos resultados obtidos no conjunto de teste utilizando a rede neural U-Net adaptada.	57
Figura 13 – Matriz de confusão dos resultados obtidos no conjunto de teste utilizando a rede neural densa sobre o vetor de características extraído.	59

Lista de tabelas

Tabela 1	–	Trabalhos relacionados.	40
Tabela 2	–	Quantidade de proteínas positivas e negativas.	54
Tabela 3	–	Resultados de classificação do modelo VGG16 com limiar otimizado (0.3) no conjunto de teste.	56
Tabela 4	–	Resultados de classificação do modelo U-Net com limiar padrão (0.5) no conjunto de teste.	57
Tabela 5	–	Resultados de classificação do modelo U-Net com limiar otimizado (0.45) no conjunto de teste.	58
Tabela 6	–	Métricas de desempenho detalhadas do modelo de rede neural densa no conjunto de teste.	59

Lista de siglas

AUCROC Área sob a Curva Característica de Operação do Receptor

CNN *Convolutional Neural Network* (Rede Neural Convolucional)

DNN *Deep Neural Network* (Rede Neural Densa)

GAP Global Average Pooling (Agrupamento da média global)

IFE Imunofixação Eletroforética ou Imunofixação

ILSVRC *ImageNet Large Scale Visual Recognition Challenge* (Desafio de Reconhecimento Visual em Grande Escala do ImageNet)

IMT Imunotipagem

LCR Líquido Cefalorraquidiano

ML *Machine Learning* (Aprendizado de Máquina)

MLP *Multilayer Perceptron* (Perceptron de Múltiplas Camadas)

ReLU *Rectified Linear Unit* (Unidade Linear Retificada)

RNA *Redes Neurais Artificiais*

RFC *Random Forest Classifier* (Classificador de Floresta Aleatório)

VGGNet *Visual Geometry Group* (Grupo de Geometria Visual)

Sumário

1	INTRODUÇÃO	23
1.1	Motivação	25
1.2	Objetivos e Desafios da Pesquisa	25
1.2.1	Desafios	25
1.2.2	Objetivo Geral	26
1.2.3	Objetivos Específicos	26
1.3	Hipótese	27
1.4	Contribuições	28
1.5	Organização da Tese	28
2	FUNDAMENTAÇÃO TEÓRICA	31
2.1	Redes Neurais Artificiais	31
2.1.1	Redes Neurais Convolucionais (CNNs)	32
2.1.2	VGG16	32
2.1.3	U-Net	34
2.2	Exame de Imunofixação	36
2.3	Classificação Automática de Exames de Imunofixação	37
3	METODOLOGIA	41
3.1	Método para a Avaliação	41
3.1.1	Pré-processamento de Imagem	41
3.1.2	VGG16	43
3.1.3	U-Net Adaptada	45
3.1.4	Rede Neural Densa (DNN)	47
4	EXPERIMENTOS	53
4.1	Descrição do Banco de Dados	53
4.2	Resultados: VGG16	54

4.3	Resultados: U-Net	56
4.4	Resultados: Rede Neural Densa	58
4.5	Avaliação dos Resultados	60
5	CONCLUSÃO	63
5.1	Principais Contribuições	64
5.2	Contribuições em Produção Bibliográfica	65
5.3	Trabalhos Futuros	65
REFERÊNCIAS		67

Introdução

A classificação automatizada de imagens médicas usando aprendizado de máquina, especialmente por meio de redes neurais, desempenha um papel crucial na modernização e aprimoramento dos processos diagnósticos. Ao automatizar a análise de imagens médicas, essa abordagem oferece uma série de vantagens significativas. Primeiramente, ela proporciona resultados mais consistentes, eliminando a variação inerente à interpretação humana, que pode ser influenciada por fatores subjetivos e variáveis como o nível de experiência do profissional ou condições específicas de cada exame. Além disso, a automatização reduz substancialmente o risco de erros humanos, que são uma preocupação constante em processos diagnósticos, onde a precisão é essencial para o tratamento eficaz de pacientes (ANWAR et al., 2018).

Os avanços computacionais recentes têm possibilitado o processamento automatizado de imagens por meio de técnicas de aprendizado de máquina. Imagens médicas, em particular, apresentam um rico conteúdo visual, com padrões complexos que exigem uma análise detalhada. Além disso, o grande volume de dados gerado nesses exames torna essencial o uso de abordagens computacionais para garantir eficiência e precisão na interpretação. A capacidade de processar grandes volumes de imagens de maneira eficiente também é uma vantagem notável, permitindo que grandes quantidades de dados sejam analisadas em um tempo relativamente curto, o que é especialmente importante em contextos clínicos, onde o tempo é um fator crítico (CAI; GAO; ZHAO, 2020)(CHLAP et al., 2021).

Neste contexto, este trabalho concentra-se especificamente na aplicação de aprendizado de máquina em exames de imunofixação por eletroforese, uma técnica laboratorial utilizada principalmente na detecção de proteínas anormais no sangue, que podem ser indicativas de várias condições patológicas, incluindo doenças hematológicas e imunológicas. A imunofixação combina duas metodologias essenciais: a eletroforese e a imunoprecipitação. A eletroforese é uma técnica que permite a separação de moléculas com base em seu tamanho e carga elétrica, facilitando a análise visual de proteínas ou a identificação de microrganismos em uma amostra. Esta técnica é amplamente utilizada em diagnósticos

laboratoriais devido à sua precisão na separação molecular. A imunoprecipitação, por sua vez, é uma técnica derivada da cromatografia de afinidade e envolve o uso de anticorpos específicos para isolar e identificar proteínas-alvo presentes na amostra. Juntas, essas técnicas fornecem uma base robusta para a detecção precisa e confiável de proteínas anormais (KEREN, 2003)(RODRIGUEZ et al., 2020)(ZACHARIOU, 2008).

Apesar do potencial dessas técnicas, a análise automatizada de imunofixação por eletroforese utilizando aprendizado de máquina é uma área relativamente nova. Somente nos últimos anos, essa tarefa começou a ser explorada com maior profundidade, resultando em alguns trabalhos promissores que demonstram desempenhos impressionantes, com precisão de classificação superior a 90% em muitos casos. Esses avanços têm sido alcançados principalmente através do uso de modelos baseados em Redes Neurais Convolucionais (LEE et al., 2022)(THIEMANN et al., 2022)(WEI et al., 2021), ou CNNs, que são projetados para capturar padrões complexos em dados visuais. Diferentes arquiteturas de CNN têm sido exploradas, cada uma oferecendo vantagens específicas em termos de desempenho e eficiência. Li et al. (LI et al., 2021), por exemplo, propuseram uma abordagem híbrida que combina vetores de características desenvolvidos manualmente com modelos de CNN, destacando a diversidade de técnicas que podem ser aplicadas nessa área. No entanto, apesar dos resultados encorajadores, ainda há espaço significativo para melhorias, particularmente na escolha e no tratamento de imagens, na seleção de vetores de características e no desenvolvimento de modelos de classificação mais robustos que possam atingir taxas de classificação correta ainda mais altas.

Neste estudo, são apresentados métodos inovadores para a análise automatizada de exames de imunofixação por eletroforese usando aprendizado de máquina. A proposta inclui uma estratégia eficiente para a extração de tiras verticais tingidas das imagens dos exames, um passo crucial para a análise subsequente. Essas tiras, que representam a expressão de proteínas na amostra, são então classificadas usando três abordagens distintas: uma rede neural VGG16, uma rede U-Net adaptada para classificação e uma abordagem que utiliza uma rede neural densa, com um vetor de características lineares extraídas das imagens pré-processadas. Essa diversidade de abordagens permite explorar diferentes perspectivas sobre o problema e avaliar qual estratégia oferece o melhor equilíbrio entre precisão e eficiência, já que não há consenso na literatura em relação ao tipo de rede neural mais adequado para a aplicação. Além disso, o estudo faz uso de um novo conjunto de dados que inclui imagens desafiadoras, com um número comparável ou superior de amostras positivas em relação à literatura existente, e uma distribuição balanceada de exemplos de tiras de imunofixação de eletroforese positivas e negativas. Isso garante que o modelo desenvolvido seja treinado e avaliado em um ambiente que reflete de maneira mais fiel os desafios encontrados na prática clínica, aumentando sua aplicabilidade e robustez.

1.1 Motivação

O diagnóstico de doenças a partir de exames de imunofixação é um desafio significativo em grandes laboratórios do Brasil e do mundo. A interpretação desses exames apresenta uma curva de aprendizado acentuada e exige um elevado nível de expertise, resultando frequentemente em resultados subjetivos que dependem da avaliação visual de especialistas. A diferença entre um exame positivo e um negativo pode ser extremamente sutil, tornando a análise ainda mais desafiadora. Essa subjetividade pode levar a inconsistências e erros no diagnóstico, além de necessitar de análises extras para confirmação, aumentando assim o tempo e os custos do processo.

Atualmente, a análise automatizada de imagens médicas utilizando aprendizado de máquina, especialmente Redes Neurais Convolucionais (CNNs), tem-se mostrado uma abordagem promissora para enfrentar esses desafios. No entanto, apesar dos avanços recentes, a aplicação de aprendizado de máquina para a análise de imunofixação por eletroforese ainda enfrenta problemas não resolvidos. Esses incluem a necessidade de métodos mais eficientes de segmentação de imagens, a criação de vetores de características mais representativos e com maior explicabilidade, e a busca por modelos de classificação que possam superar as limitações dos métodos atuais (ÇUBUKÇU; TOPCU; YENICE, 2024) (DODIG; ĆEPELAK; DODIG, 2025).

Neste contexto, o presente trabalho se propõe a desenvolver e avaliar comparativamente metodologias para a análise automatizada de exames de imunofixação utilizando técnicas de aprendizado de máquina. A próxima seção detalhará os objetivos específicos deste estudo, incluindo a implementação de novas estratégias de pré-processamento de imagens e a comparação de diferentes arquiteturas de redes neurais na tarefa de classificação de exames.

1.2 Objetivos e Desafios da Pesquisa

A classificação automatizada de exames de imunofixação por eletroforese apresenta uma série de desafios e objetivos que este estudo pretende abordar e alcançar.

1.2.1 Desafios

- ❑ Curva de Aprendizado e Subjetividade: A interpretação dos exames de imunofixação requer um alto nível de especialização devido à complexidade dos resultados, que são frequentemente subjetivos. Isso pode levar a inconsistências no diagnóstico.
- ❑ Necessidade de Métodos Automatizados e Eficientes: A automatização da análise de imagens de imunofixação precisa ser suficientemente robusta para lidar com a diversidade e complexidade das imagens, proporcionando resultados precisos e con-

sistentes. Esses métodos necessitam de alto grau de acurácia para uso clínico, e, se possível, explicabilidade para aumentar a confiança dos resultados.

- ❑ Segmentação e Pré-Processamento de Imagens: Imagens de exames de imunofixação não são necessariamente padronizadas, apresentando tiras verticais em diversas posições que precisam ser detectadas de forma precisa.
- ❑ Desenvolvimento de Vetores de Características: CNNs usam a imagem como entrada, mas outros classificadores funcionam melhor com vetores de características que capturem de forma eficaz e em um vetor de menos dimensões as informações relevantes das imagens, garantindo um alto potencial discriminativo para melhorar o desempenho dos modelos de classificação.
- ❑ Várias Arquiteturas de Redes Neurais Possíveis: Dada a diversidade de arquiteturas disponíveis, é importante avaliar e comparar o desempenho dos diferentes modelos de redes neurais na tarefa específica de classificação de exames de imunofixação.

1.2.2 Objetivo Geral

- ❑ Classificar exames de imunofixação por eletroforese utilizando técnicas de aprendizado de máquina, contribuindo para a melhoria do diagnóstico de doenças por meio da análise automatizada.

1.2.3 Objetivos Específicos

- ❑ Desenvolver uma nova estratégia de pré-processamento baseada na extração de tiras verticais tingidas das imagens dos exames.
- ❑ Comparar o desempenho de diferentes arquiteturas de redes neurais na classificação de exames de imunofixação.
- ❑ Elaborar e organizar um novo conjunto de dados de exames de imunofixação, contendo imagens desafiadoras e uma distribuição relativamente balanceada de exemplos positivos e negativos, para validar os modelos propostos.
- ❑ Avaliar e comparar a precisão, *recall*, *F1-Score* e outras métricas de desempenho dos modelos propostos, destacando as vantagens e limitações de cada abordagem.

Com esses objetivos, este estudo não apenas visa desenvolver um sistema ou explorar um método, mas também gerar novos conhecimentos que contribuam para a melhoria dos processos diagnósticos em laboratórios, reduzindo a subjetividade e aumentando a precisão e a consistência dos resultados.

1.3 Hipótese

Com base nos objetivos e desafios delineados, este trabalho parte das seguintes hipóteses, as quais serão testadas pelos experimentos descritos no Capítulo 3:

1. Hipótese 1 (H1):

A estratégia de pré-processamento proposta, que segmenta as imagens em tiras verticais de interesse, baseada no histograma de intensidade de cor e na detecção da tira de controle, é eficaz em extrair as regiões de interesse do exame, que são as tiras verticais.

- Pergunta associada (Q1): Quão eficaz é a extração de tiras verticais usando a estratégia proposta? E qual a redução aproximada de dimensionalidade da imagem de entrada quando apenas a tira vertical, que é a região de interesse, é usada?

2. Hipótese 2 (H2):

A adaptação da U-Net para tarefa de classificação, por meio de Global Average Pooling (Agrupamento da média global) (GAP) sobre seu *decoder* e posterior cabeça densa, é capaz de capturar características hierárquicas relevantes e atingir performance comparável à de CNNs especificamente desenvolvidas para classificação como VGG16.

- Pergunta associada (Q2): A U-Net adaptada para classificação alcança *F1-Score* semelhante ao modelo VGG16 quando avaliada no mesmo conjunto de tiras verticais?

3. Hipótese 3 (H3):

Um classificador baseado em Rede Neural Densa, alimentado por um vetor de características extraído manualmente a partir da coluna central de cada tira (intensidade de pixel ao longo do eixo vertical), poderá superar as CNNs em termos de *F1-Score*, uma vez que explora diretamente os picos de intensidade correlacionados à presença de proteínas monoclonais, enquanto oferece ganhos em termos de espaço e tempo de processamento.

- Pergunta associada (Q3): Qual a diferença de desempenho (precisão, *recall* e *F1-Score*) entre o classificador DNN com vetor manual e as CNNs (VGG16 e U-Net), utilizando o mesmo conjunto de dados segmentados?

Todas as hipóteses serão validadas quantitativamente no Capítulo 3, por meio de métricas de classificação (acurácia, precisão, *recall*, *F1-Score* e Área sob a Curva Característica de Operação do Receptor (AUCROC)) e análise comparativa entre modelos.

1.4 Contribuições

As principais contribuições deste trabalho podem ser resumidas da seguinte forma:

- ❑ Um método de pré-processamento para segmentar eficientemente imagens de exames de eletroforese em tiras verticais junto com seus dados.
- ❑ O uso de redes neurais em um novo conjunto de dados de exames de eletroforese mais balanceado em comparação com a literatura existente.
- ❑ Um vetor de características definido manualmente a partir de imagens de exames que, quando combinado com uma rede neural densa, se equipara às CNNs em precisão.
- ❑ Uma análise comparativa de redes neurais para a classificação de imunofixação por eletroforese.

1.5 Organização da Tese

Esta tese foca na pesquisa e aplicação de métodos de aprendizado profundo, especificamente redes neurais, para a classificação automática de imagens provenientes de Imunofixação Eletroforética ou Imunofixação (IFE). O objetivo principal é desenvolver e avaliar modelos capazes de identificar padrões relevantes nesses exames, como a presença de componentes monoclonais, auxiliando no diagnóstico de condições associadas.

A abordagem central deste trabalho reside na utilização de Redes Neurais Convolucionais (CNNs), arquiteturas de aprendizado profundo especialmente eficazes na análise de dados visuais. Investiga-se como essas redes podem ser treinadas para extrair características discriminativas das imagens de IFE e realizar a classificação de forma eficiente e precisa. Exploram-se diferentes configurações e técnicas para otimizar o desempenho dos modelos na distinção entre resultados normais e anormais.

A estrutura da dissertação organiza-se da seguinte forma: A próxima seção elenca os objetivos gerais e específicos deste trabalho. O Capítulo 2 apresenta uma revisão bibliográfica, abordando os conceitos fundamentais da eletroforese de imunofixação, os desafios inerentes à sua interpretação visual e os avanços recentes na aplicação de inteligência artificial e aprendizado de máquina para análise de imagens médicas, com foco em dados de eletroforese.

Em seguida, o Capítulo 3 detalha a metodologia empregada, incluindo a descrição do conjunto de dados utilizado, as etapas de pré-processamento das imagens, as arquiteturas de redes neurais exploradas ou propostas e os protocolos de treinamento e validação dos modelos desenvolvidos.

O Capítulo 4 é dedicado à apresentação e análise dos experimentos e seus resultados. Descreve-se a configuração experimental, as métricas de avaliação utilizadas e os

resultados obtidos pelos modelos de classificação. Realiza-se uma análise comparativa do desempenho alcançado, potencialmente contrastando com métodos anteriores ou *benchmarks* da área.

Por fim, o Capítulo 5 apresenta as considerações finais, sumariando as principais contribuições desta dissertação para a classificação automatizada de testes de IFE. Discutem-se as conclusões obtidas, as limitações do estudo e propõem-se direções para trabalhos futuros que possam expandir ou aprimorar as abordagens aqui desenvolvidas.

Fundamentação Teórica

2.1 Redes Neurais Artificiais

As Redes Neurais Artificiais (RNA) são modelos computacionais inspirados no funcionamento do cérebro humano, projetados para executar tarefas que envolvem aprendizado e tomada de decisão. Elas são compostas por neurônios artificiais organizados em camadas – geralmente divididas em entrada, ocultas e saída – interligadas por conexões ponderadas, como representado na Figura 1. Cada neurônio recebe múltiplas entradas, que são combinadas linearmente por meio de pesos sinápticos e um viés (*bias*), elementos fundamentais para ajustar a influência de cada entrada na ativação do neurônio. O resultado dessa combinação é então processado por uma função de ativação, como *Rectified Linear Unit* (Unidade Linear Retificada) (ReLU), *sigmoide* ou *softmax*, que introduz não linearidade ao modelo e permite a aprendizagem de relações complexas nos dados. Essa estrutura possibilita a aplicação das RNAs em áreas como reconhecimento de padrões, processamento de linguagem natural e análise de imagens (PRIDDY, 2005) (LECUN; BENGIO; HINTON, 2015)

O aprendizado das RNAs ocorre por meio de três etapas principais: propagação direta, cálculo do erro e retropropagação. Inicialmente, os dados percorrem a rede na propagação direta, sendo transformados progressivamente pelas conexões ponderadas até a obtenção da saída final. Em seguida, o erro entre a predição e o valor esperado é calculado por meio de uma função de perda, fornecendo uma métrica para avaliar o desempenho do modelo. Para minimizar esse erro, a retropropagação ajusta os pesos da rede utilizando o método do gradiente descendente e suas variantes, permitindo que a RNA refine suas conexões ao longo das iterações do treinamento. Esse processo iterativo capacita o modelo a identificar padrões nos dados e generalizar para novas entradas, tornando as RNAs altamente eficazes na resolução de problemas complexos (GOODFELLOW, 2016)

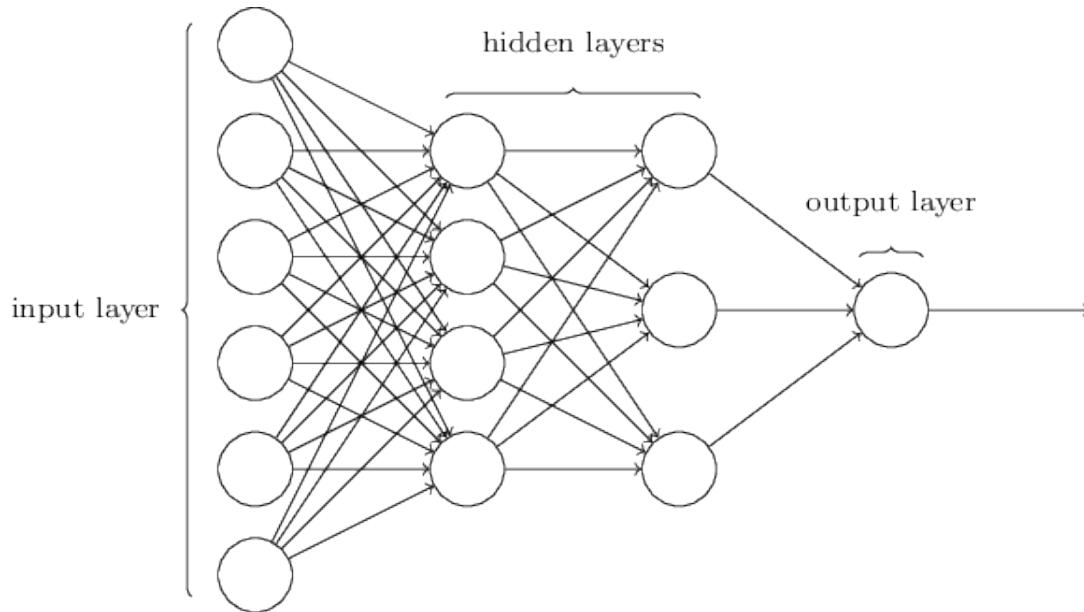


Figura 1 – Arquitetura de uma Rede Neural Artificial (RNA) com camada de entrada (*input layer*), camadas ocultas (*hidden layers*) e camada de saída (*output layer*), ilustrando o fluxo de informação entre os neurônios.

Fonte: www.deeplearningbook.com.br/a-arquitetura-das-redes-neurais/

2.1.1 Redes Neurais Convolucionais (CNNs)

As CNNs são uma classe de RNAs amplamente utilizadas para análise e processamento de dados visuais, como imagens e vídeos. Sua arquitetura difere das RNAs tradicionais ao incluir camadas convolucionais, que realizam operações de convolução para extrair características locais dos dados, conforme exemplificado na Figura 2. Essas redes também utilizam operações de *pooling*, que reduzem a dimensionalidade dos dados, preservando as informações mais relevantes (KRIZHEVSKY; SUTSKEVER; HINTON, 2017).

A principal vantagem das CNNs é sua habilidade de aprender hierarquias de características. Nas primeiras camadas, as redes identificam padrões básicos, como bordas e texturas. Em camadas mais profundas, as CNNs capturam informações mais abstratas, como formas e objetos completos. Esse comportamento torna as CNNs ideais para tarefas de classificação de imagens, detecção de objetos e segmentação semântica (RAWAT; WANG, 2017).

2.1.2 VGG16

A VGG Net, proposta por Simonyan e Zisserman (2014), é uma arquitetura de *Convolutional Neural Network* (Rede Neural Convolucional) (CNN) que se destaca pela simplicidade e alto desempenho em tarefas como o *ImageNet Large Scale Visual Recognition Challenge* (Desafio de Reconhecimento Visual em Grande Escala do ImageNet) (ILSVRC). Sua principal característica é o uso sistemático de pequenas janelas de convolução (3x3), empilhadas em profundidade, o que permite extrair informações detalhadas sem aumentar

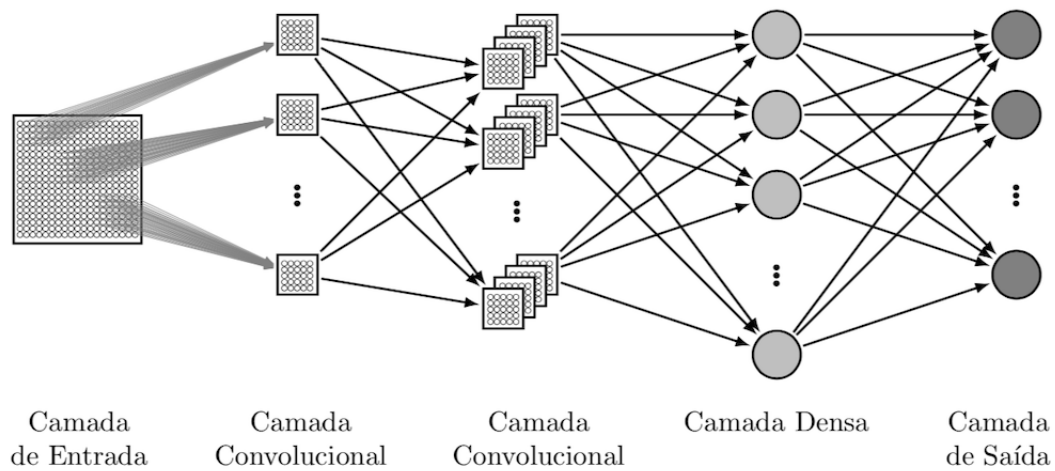


Figura 2 – Diagrama ilustrando uma CNN com camadas de entrada, convolucionais, densas e de saída, usadas para extração de características e classificação em tarefas de visão computacional.

Fonte: <https://www.sakurai.dev.br/cnn-mapreduce/>

significativamente o custo computacional.

Entre os modelos da família VGG, a VGG16 é uma das versões mais amplamente utilizadas. Seu nome refere-se às 16 camadas com pesos treináveis: 13 convolucionais e 3 totalmente conectadas (SIMONYAN; ZISSERMAN, 2014). É comumente empregada como modelo de base (*baseline*) para tarefas de classificação e como extrator de características em aplicações como detecção de objetos e segmentação de imagens.

A VGG16 processa imagens de entrada com dimensão fixa (tipicamente $224 \times 224 \times 3$) por meio de cinco blocos convolucionais. Cada bloco utiliza exclusivamente filtros de 3×3 com *stride* 1, seguidos por camadas de *max pooling* 2×2 com *stride* 2, que reduzem pela metade a resolução espacial, preservando as informações mais relevantes (SIMONYAN; ZISSERMAN, 2014).

A organização dos blocos é a seguinte:

- ❑ Blocos 1 e 2: Cada um possui duas camadas convolucionais seguidas de uma de *pooling*. Os filtros começam com 64 no primeiro bloco e aumentam para 128 no segundo.
- ❑ Blocos 3 a 5: Contêm três camadas convolucionais e uma de *pooling*. O número de filtros dobra progressivamente: 256 no terceiro bloco, e 512 nos dois últimos.

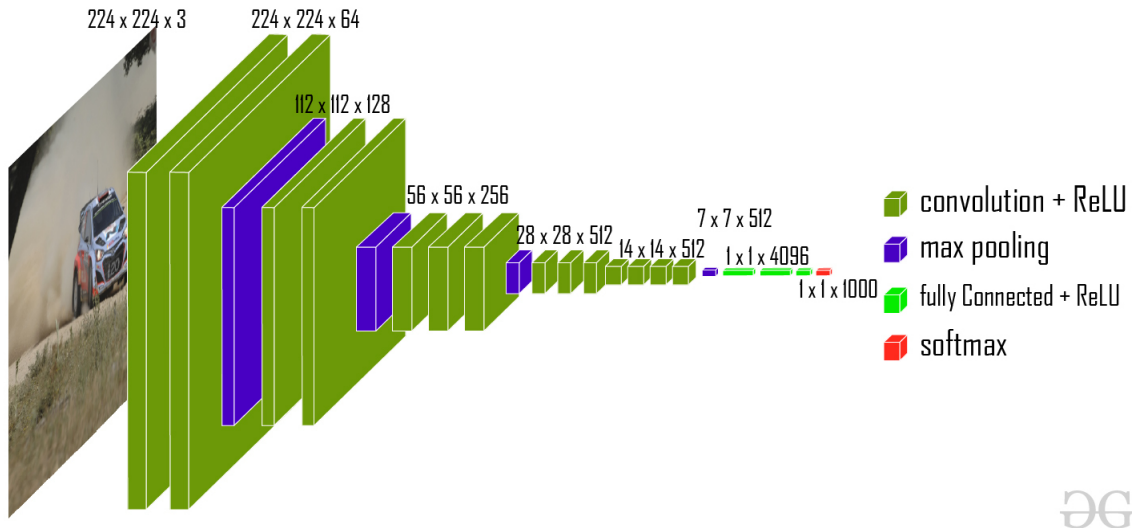


Figura 3 – Estrutura da VGG16: camadas convolucionais (verde), *pooling* (roxo), totalmente conectadas (verde claro) e *softmax* (vermelho). A arquitetura reduz a dimensionalidade espacial e extrai características para classificação.

Fonte: <https://www.geeksforgeeks.org/computer-vision/vgg-16-cnn-model/>

Após o último bloco, o mapa de características é achatado (*flattened*) em um vetor, que é processado por uma rede totalmente conectada composta por:

- ❑ Duas camadas ocultas com 4096 neurônios cada.
- ❑ Uma camada de saída com 1000 neurônios (correspondente às classes do *ImageNet*), seguida por uma ativação *softmax*.

Embora tenha um design relativamente simples, o elevado número de parâmetros, especialmente nas camadas totalmente conectadas, torna a VGG16 intensiva em termos computacionais e de memória. Ainda assim, sua base convolucional demonstrou ser altamente eficaz para extração de características, o que a consolidou como uma escolha popular em estratégias de *transfer learning*, aproveitando pesos pré-treinados em bases como a *ImageNet* (QASSIM; VERMA; FEINZIMER, 2018).

2.1.3 U-Net

A U-Net é uma arquitetura de rede neural convolucional (CNN) desenvolvida primordialmente para tarefas de segmentação de imagens biomédicas. Proposta originalmente por Ronneberger, Fischer e Brox (2015), sua principal característica é a capacidade de realizar uma segmentação precisa mesmo com um conjunto de dados de treinamento relativamente pequeno, uma situação comum em aplicações médicas. A arquitetura demonstrou um desempenho notável em diversas competições de segmentação de imagens, tornando-se um padrão de referência na área (RONNEBERGER; FISCHER; BROX, 2015).

Desde sua proposição, a U-Net tem sido amplamente adotada e aprimorada em estudos voltados à segmentação de imagens médicas, apresentando resultados de alta acurácia em diferentes modalidades de imagem, como ressonância magnética, biomédicas, tomografia computadorizada e imagens microscópicas (DU et al., 2020) (??). Por exemplo, Zhang et al. (2024) demonstram o uso eficaz da U-Net com aprimoramentos baseados em atenção para segmentação de tumores cerebrais, alcançando desempenho superior a métodos clássicos (ZHANG; LIU; WANG, 2024). De forma semelhante, o trabalho de Akinyelu e Adewumi (2023) também destaca o sucesso da U-Net na segmentação de lesões pulmonares em imagens de tomografia, reforçando sua versatilidade e robustez nesse domínio (AKINYELU; ADEWUMI, 2023).

A estrutura da U-Net, que visualmente se assemelha à letra U, conforme observado na Figura 4, é o que lhe confere o nome e grande parte de sua eficácia. Ela consiste em dois caminhos principais: um caminho de contração (*encoder*) e um caminho de expansão (*decoder*).

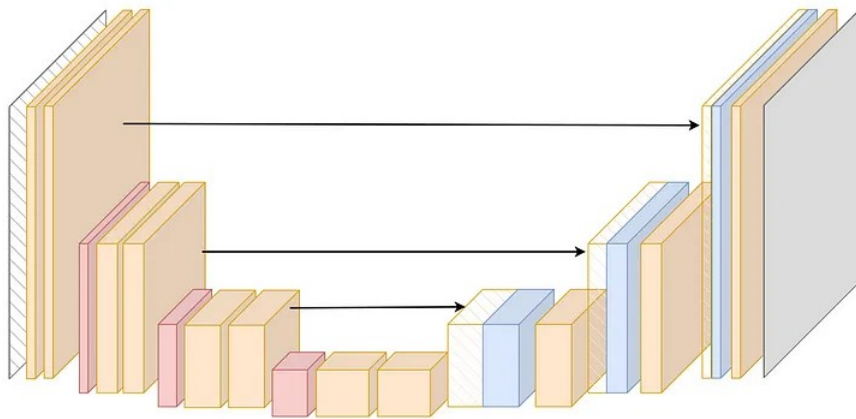


Figura 4 – Representação esquemática da arquitetura U-Net. À esquerda, o caminho de contração (*encoder*) captura o contexto da imagem, enquanto à direita, o caminho de expansão (*decoder*) recupera a informação espacial para a segmentação. As setas horizontais representam as conexões de atalho (*skip connections*), que são fundamentais para preservar padrões locais e sua localização.

Fonte: <https://medium.com/data-science/u-net-explained-understanding-its-image-segmentation-architecture-56e4842e313a>

Caminho de Contração (*Encoder*): O *encoder* segue a arquitetura típica de uma rede convolucional. Ele é composto por uma sequência de blocos, onde cada bloco consiste em duas convoluções de 3x3 seguidas por uma função de ativação *ReLU* e uma operação de *max pooling* de 2x2. O objetivo do *encoder* é capturar o contexto da imagem. A cada operação de *max pooling*, a resolução espacial da imagem é reduzida pela metade, enquanto o número de canais de características é duplicado. Isso permite que a rede aprenda características hierárquicas em diferentes escalas (SIDDIQUE et al., 2021).

Caminho de Expansão (*Decoder*): O *decoder* tem como objetivo reconstruir a imagem segmentada com a resolução original, utilizando as características aprendidas pelo

encoder. Para isso, a cada passo do *decoder*, o mapa de características é submetido a uma "convolução para cima" (*up-convolution*) de 2x2, que aumenta a resolução espacial, seguida pela concatenação com o mapa de características correspondente do caminho de contração. Essas conexões, conhecidas como *skip connections* (conexões de atalho), são a inovação fundamental da U-Net. Elas permitem que o *decoder* recupere informações espaciais de alta resolução que foram perdidas durante a contração, combinando o "o quê" (características semânticas do *decoder*) com o "onde" (características de localização do *encoder*). Após a concatenação, seguem-se duas convoluções de 3x3 e uma ativação *ReLU* (RONNEBERGER; FISCHER; BROX, 2015).

A camada final da rede utiliza uma convolução de 1x1 para mapear o vetor de características de múltiplos canais para o número desejado de classes de segmentação.

Desde sua concepção, a U-Net inspirou uma vasta gama de variantes que buscam aprimorar sua performance para diferentes aplicações. Pesquisadores têm explorado modificações na estrutura do *encoder-decoder* e nas *skip connections* para melhorar ainda mais a precisão da segmentação.

Em resumo, a U-Net representa uma família de arquiteturas poderosas e flexíveis para segmentação de imagens. Sua estrutura fundamental de *encoder-decoder* com *skip connections* estabeleceu um novo paradigma, e as inovações contínuas, como blocos residuais, conexões densas, mecanismos de atenção e a integração com *Transformers*, continuam a expandir as fronteiras da precisão e da eficiência na segmentação semântica, especialmente no domínio médico (HATAMIZADEH et al., 2022).

2.2 Exame de Imunofixação

O exame de imunofixação (IFE) é uma técnica laboratorial que combina a eletroforese com a imunofixação para identificar proteínas específicas, especialmente imunoglobulinas, em amostras como soro, urina e Líquido Cefalorraquidiano (LCR). A eletroforese separa as proteínas com base em sua carga e tamanho, enquanto a imunofixação utiliza anticorpos específicos para detectar as proteínas de interesse após a separação. Este método é fundamental para identificar imunoglobulinas monoclonais, que são proteínas anormais produzidas por um único clone de plasmócitos, auxiliando no diagnóstico de diversas condições médicas (CAILLON et al., 2013) (GILES; KARUNANITHI, 2024).

O IFE possui diversas aplicações clínicas, sendo crucial no diagnóstico e monitoramento de gamopatias monoclonais, como o mieloma múltiplo, a macroglobulinemia de Waldenström e a amiloidose. É considerado o "padrão ouro" para confirmar a presença de paraproteínas (proteínas monoclonais) no sangue ou na urina, demonstrando maior sensibilidade na detecção dessas proteínas em comparação com a eletroforese de proteínas séricas (EPS). No LCR, o IFE é utilizado para detectar bandas oligoclonais, um importante marcador de produção intratecal de imunoglobulinas associado a condições como a

esclerose múltipla (O'CONNELL; HORITA; KASRAVI, 2005).

O procedimento do IFE envolve inicialmente a separação das proteínas da amostra por eletroforese em gel de agarose. Após a separação, antissoros contendo anticorpos específicos para diferentes classes de imunoglobulinas (IgG, IgA, IgM) e cadeias leves (*kappa* e *lambda*) são aplicados no gel. Se a proteína correspondente estiver presente, ocorrerá a formação de um complexo antígeno-anticorpo insolúvel, que é então visualizado como uma banda após coloração, permitindo a identificação da proteína monoclonal específica (O'CONNELL; HORITA; KASRAVI, 2005) (CAWLEY et al., 1976).

Como exemplo do procedimento descrito, a Figura 5 demonstra um resultado de imunofixação. A formação de um precipitado visível, que forma uma banda nítida nas trilhas onde foram aplicados os antissoros anti-IgG e anti-*kappa*, evidencia a detecção e a caracterização de uma proteína monoclonal específica, do isotipo IgG *kappa*.

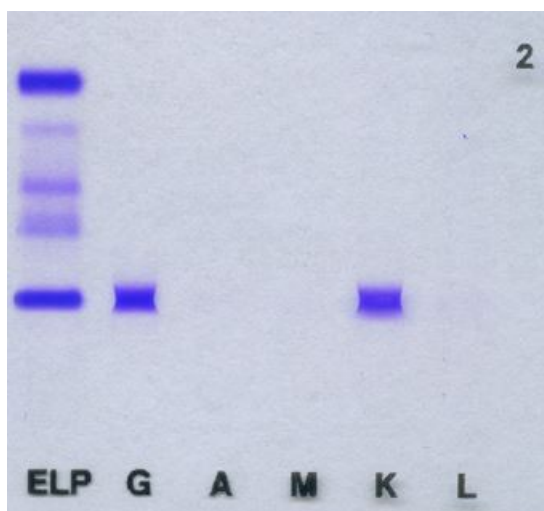


Figura 5 – Gel de agarose apresentando o perfil eletroforético da amostra (ELP) à esquerda, seguido pelas trilhas de imunoprecipitação. A reação com antissoros específicos revelou uma banda de precipitação bem definida nas trilhas correspondentes à cadeia pesada G (IgG) e à cadeia leve K (*kappa*).

2.3 Classificação Automática de Exames de Imunofixação

A análise automatizada de exames de imunofixação por eletroforese utilizando aprendizado de máquina baseado em imagens tem sido um tema de crescente interesse e pesquisa nos últimos anos. Essa linha de investigação surge da necessidade de aprimorar métodos diagnósticos tradicionais, combinando técnicas avançadas de processamento de imagens com algoritmos de aprendizado de máquina. O objetivo principal é aumentar a precisão e a eficiência das análises, minimizando a interferência de fatores subjetivos e reduzindo a possibilidade de erros humanos. A implementação de técnicas automatizadas nessa

área também é motivada pela crescente demanda por diagnósticos rápidos e precisos em ambientes clínicos, onde o tempo e a acurácia são cruciais para o tratamento adequado dos pacientes. Diante desse cenário, diversas abordagens têm sido propostas e testadas, buscando otimizar a análise das imagens e, consequentemente, os resultados diagnósticos.

Wei et al. (WEI et al., 2021) obtiveram resultados significativos ao empregar uma abordagem de *deep learning* (aprendizado profundo) que integra uma máscara de atenção guiada, estratégia conhecida como "*coaching* de pré-processamento". Essa técnica visa aprimorar o processo de pré-processamento das imagens, direcionando a atenção do modelo para as áreas mais relevantes das imagens, o que potencialmente melhora a precisão da classificação. Além disso, os autores adotaram uma abordagem conservadora na construção do conjunto de dados, excluindo exames de baixa qualidade que poderiam introduzir ruído no modelo e comprometer os resultados. O conjunto de dados utilizado por eles, composto por 4.352 imagens, foi analisado utilizando arquiteturas de redes neurais convolucionais avançadas, como ResNet, VGGNet e Inception-v3. Essas arquiteturas são amplamente reconhecidas por sua capacidade de capturar características complexas em imagens, o que foi fundamental para o sucesso do modelo proposto pelos autores.

Li et al. (LI et al., 2021), por sua vez, seguiram uma abordagem diferente ao desenvolver manualmente um vetor de características baseado em Modelo de Mistura de Gaussianas e detecção de picos, visando à classificação eficaz das imagens. Essa abordagem, embora menos dependente de arquiteturas profundas de redes neurais, mostra que técnicas clássicas de extração de características ainda têm um papel importante na análise de imagens médicas, especialmente quando combinadas com métodos de aprendizado de máquina. O conjunto de dados utilizado neste estudo foi consideravelmente menor, com 191 imagens, mas os resultados obtidos foram suficientes para demonstrar a viabilidade da abordagem proposta, destacando-se como uma alternativa relevante em contextos onde a disponibilidade de dados é limitada.

Mais recentemente, Hu et al. (HU et al., 2023) alcançaram taxas de classificação extremamente altas ao empregar um conjunto de três redes neurais profundas, aplicadas a um vasto conjunto de dados contendo 12.703 imagens. Essa abordagem de ensemble, que combina os resultados de múltiplas redes neurais, é particularmente eficaz em melhorar a robustez e a precisão do modelo, pois permite que diferentes modelos capturem aspectos complementares das imagens analisadas. A utilização de um grande conjunto de dados também contribuiu para o sucesso do modelo, proporcionando uma base diversificada e abrangente para o treinamento das redes neurais.

Thiemann et al. (THIEMANN et al., 2022) exploraram a arquitetura ResNet18 de CNN para avaliação automatizada das imagens, com um enfoque especial na aplicação de um limite de confiança para classificar os resultados de baixa confiança como inconclusivos. Essa estratégia é importante em contextos clínicos, onde a incerteza na classificação pode ter implicações significativas para o diagnóstico e tratamento do paciente. O conjunto de

dados utilizado por esses autores consistiu em 14.053 imagens, com uma predominância de casos não patológicos (9.160 imagens) ou inconclusivos (1.915 imagens). Essa distribuição de dados reforça a necessidade de desenvolver modelos que sejam capazes de lidar adequadamente com a variedade e complexidade das amostras encontradas em cenários reais.

Elfert et al. (ELFERT et al., 2024) investigaram a aplicação de cinco algoritmos de *machine learning*, incluindo um *Random Forest Classifier* (Classificador de Floresta Aleatório) (RFC) e uma CNN, para a detecção automatizada de Proteínas-M a partir de gráficos de eletroforese de proteínas séricas. O estudo destacou a superioridade do RFC, que não apenas superou o desempenho médio de especialistas de laboratório, mas também demonstrou a capacidade de ter seu limiar de predição ajustado para otimizar a sensibilidade e especificidade conforme o contexto clínico. O treinamento e a validação dos modelos foram realizados em um robusto conjunto de dados de 69.722 amostras de pacientes, sendo 4.273 positivas para Proteína-M e 65.449 negativas, com a classificação confirmada por Imunotipagem (IMT) como padrão-ouro.

Por fim, Lee et al. (LEE et al., 2022) empregaram uma arquitetura DenseNet CNN para classificar 2.578 imagens. O que diferencia o trabalho desses autores é a categorização dos resultados em diagnósticos específicos de doenças, em vez de simplesmente identificar a presença ou ausência de proteínas alvo. Essa abordagem agrega um nível adicional de precisão e utilidade clínica, pois permite que o modelo não apenas detecte anomalias, mas também sugira possíveis diagnósticos, auxiliando diretamente os especialistas no processo de tomada de decisão.

Apesar dos impressionantes índices de desempenho relatados, uma análise crítica revela limitações importantes que devem ser consideradas. Uma questão central é a generalização dos modelos. Isso introduz um risco de que os modelos aprendam características específicas dos equipamentos, reagentes e protocolos locais, não se mostrando igualmente eficazes quando aplicados a dados de outros laboratórios, uma limitação conhecida como falta de validação externa. Em contrapartida, abordagens como a de Li et al. (LI et al., 2021), que utilizam bancos de dados consideravelmente menores, enfrentam um risco aumentado de sobreajuste, onde o modelo se adapta excessivamente aos dados de treino e perde a capacidade de generalizar para novos exemplos não vistos, comprometendo a robustez de seus resultados em um cenário clínico mais amplo e diversificado.

Outro ponto crítico reside na metodologia de validação e na curadoria dos dados. Por exemplo, a decisão de Wei et al. (WEI et al., 2021) de excluir exames de baixa qualidade, embora compreensível para otimizar a performance, pode resultar em um modelo com desempenho superestimado, cuja eficácia em um fluxo de trabalho clínico real, que inevitavelmente inclui amostras desafiadoras ou com artefatos, permanece incerta. A abordagem de Thiemann et al. (THIEMANN et al., 2022) de classificar resultados de baixa confiança como inconclusivos representa uma solução mais pragmática para este

problema, mas reforça que a automação completa ainda não é uma realidade. Além disso, a variação nas métricas de avaliação reportadas (ora precisão, ora *recall*) dificulta uma comparação direta e justa entre os diferentes estudos, evidenciando a falta de um protocolo de avaliação padronizado na área. A ausência de validação externa na maioria desses trabalhos é, talvez, a limitação mais significativa, pois a comprovação do desempenho em múltiplos centros clínicos é um passo indispensável para a translação de qualquer ferramenta de diagnóstico baseada em IA para a prática clínica.

Tabela 1 – Trabalhos relacionados.

Trabalho	Número de imagens	Desempenho
Wei(WEI et al., 2021)	4352	72.94% (<i>recall</i>)
Li(LI et al., 2021)	191	86.20% (<i>recall</i>)
Hu(HU et al., 2023)	12,703	93.17% (<i>recall</i>)
Thiemann(THIEMANN et al., 2022)	14,053	98.70% (precisão)
Lee(LEE et al., 2022)	2,578	82.81% (precisão)
Elfert (ELFERT et al., 2024)	69,722	89.9% (<i>recall</i>)

Todos os trabalhos mencionados conseguiram taxas de classificação corretas superiores a 80% ou 90%, demonstrando o potencial significativo das abordagens baseadas em aprendizado de máquina para a análise automatizada de exames de imunofixação por eletroforese. Os dois últimos trabalhos, em particular, enfatizaram a utilidade de suas abordagens como ferramentas de auxílio diagnóstico para especialistas, ressaltando, no entanto, que a automação completa pode exigir refinamentos adicionais. Essa observação sublinha a importância de pesquisas contínuas nesta área, uma vez que o aprimoramento desses sistemas pode levar a uma maior acurácia diagnóstica e, conseqüentemente, a melhores resultados clínicos. Os resultados resumidos desses trabalhos são apresentados na Tabela 1, que oferece uma visão comparativa das metodologias e performances alcançadas.

Da análise dos trabalhos acima, é importante realçar que as CNNs são a ênfase principal na exploração de soluções para classificação automática de exames de eletroforese de imunofixação, enquanto vetores de características extraídas manualmente são menos prevalentes na literatura.

Metodologia

Neste capítulo, é descrito o método utilizado para a avaliação das imagens de exames de eletroforese por imunofixação. A abordagem abrange o pré-processamento das imagens, que inclui redimensionamento e segmentação, e também a aplicação de redes neurais para a classificação dos resultados. Três arquiteturas de rede foram testadas: VGG16, U-Net e uma rede neural densa tradicional *Multilayer Perceptron* (Perceptron de Múltiplas Camadas) (MLP). A escolha dos métodos e a comparação entre eles foi intencional, como forma de explorar experimentalmente três frentes distintas de experimentos: a rede VGG16 é uma CNN especializada em classificação; a rede U-Net é uma rede tipicamente usada para segmentação que valoriza padrões locais e sua localização, aqui adaptada para classificação; e a rede MLP é uma rede que tradicionalmente tem desempenho melhor em vetores com menos dimensões, motivo pelo qual a entrada desta rede foi usada com um vetor de características gerado manualmente, em oposição às outras duas arquiteturas em que a entrada das redes foi a imagem. As três abordagens enfatizam características específicas e distintas das imagens, de forma que a análise poderá oferecer resultados informativos em relação às particularidades da aplicação e à importância de aspectos como o posicionamento das tiras de controle e a necessidade de ignorar variações horizontais irrelevantes para obter resultados de alta acurácia na classificação.

3.1 Método para a Avaliação

3.1.1 Pré-processamento de Imagem

As imagens dos exames foram pré-processadas para garantir a consistência e a eficácia na análise subsequente. O primeiro passo desse processo envolveu o redimensionamento das imagens para um tamanho padronizado de 250x250 pixels. Esse ajuste foi necessário para uniformizar todas as imagens, facilitando a comparação e análise automática das características das tiras. O redimensionamento assegura que todas as imagens de exames sejam tratadas da mesma forma, reduzindo variações que poderiam impactar negativa-

mente o desempenho dos algoritmos de análise.

Após o redimensionamento, foi realizado o passo crucial de identificar e localizar a tira de controle dentro da imagem, rotulada como "ELP". Essa tira de controle é fundamental no contexto dos exames, pois a sua posição ajuda a identificar as outras tiras na imagem. A razão para isso é que a posição das letras que indicam as proteínas pode variar; elas podem aparecer na parte superior ou inferior da imagem. Portanto, localizar a tira de controle permite que se faça uma correção de posicionamento e se obtenha uma leitura precisa das tiras restantes, assegurando que as informações obtidas sejam corretas e confiáveis.

A segmentação das colunas das tiras foi realizada com base na identificação da posição das letras "ELP" na tira de controle. Para isso, foi aplicada uma máscara binária à imagem, um procedimento que facilita a detecção das letras e números presentes no exame. A aplicação dessa máscara permite isolar e identificar os elementos relevantes da imagem. Em seguida, uma operação de dilatação foi realizada para expandir as áreas detectadas, tornando-as mais visíveis e fáceis de manipular. Posteriormente, o contorno da coluna de controle encontrado na imagem foi selecionado; a coluna de controle é distinta por conter três letras, ao contrário das outras colunas que possuem apenas uma. Esse procedimento é essencial para garantir que a coluna de controle seja identificada corretamente e que sua informação seja utilizada para a segmentação das demais colunas.

A Fig. 6 ilustra de forma detalhada o processo de pré-processamento das imagens. Esta figura mostra o fluxo de trabalho desde a captura inicial da imagem da lâmina de exame até a sua divisão em tiras verticais, com cada tira sendo tratada individualmente para análise posterior. O fluxograma ajuda a visualizar cada etapa do processo, permitindo uma compreensão clara das operações realizadas para preparar as imagens para análise.

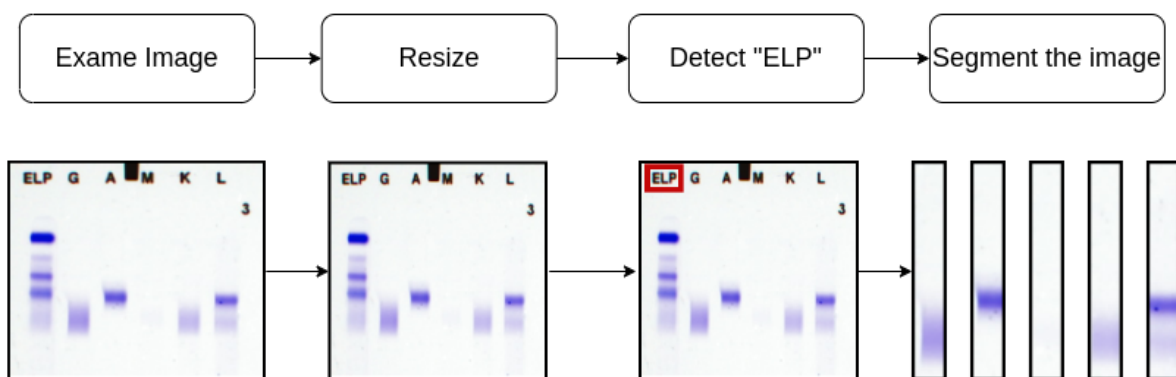


Figura 6 – O diagrama ilustra as etapas do pré-processamento das imagens de imunofixação, incluindo a imagem do exame, redimensionamento, detecção da marcação "ELP" e segmentação das bandas para posterior análise.

Após a segmentação, as etiquetas de identificação das proteínas foram removidas de cada coluna. Essa etapa foi essencial para garantir que o modelo de aprendizado de máquina não utilizasse a identidade textual da proteína como um fator de classificação, o que introduziria um viés na análise. Com a remoção desses rótulos, que se localizavam

nas extremidades, a altura efetiva de cada imagem de coluna foi reduzida de 250 para 150 pixels, deixando exclusivamente a área de interesse. A média de área das regiões de interesse é 27,1% da área da imagem original completa do exame, uma redução significativa do tamanho da entrada.

Essas colunas resultantes, agora livres de etiquetas e com altura de 150 pixels, formaram a base para a extração de características. Para a rede neural densa, a coluna central de cada imagem foi extraída e convertida em um vetor unidimensional de dimensão $n=150$, servindo como o vetor de características. As colunas inteiras foram então utilizadas como entrada para três arquiteturas distintas: a VGG16, escolhida por ser um modelo consagrado de extração de características; a U-Net, que, apesar de ser voltada para segmentação, foi adaptada para classificação; e, por fim, a já mencionada rede neural densa para uma abordagem de classificação mais direta. A combinação dessas três arquiteturas permitiu uma análise multifacetada, capitalizando sobre as diferentes formas de processamento de cada modelo.

3.1.2 VGG16

Para a implementação e avaliação do modelo baseado na arquitetura VGG16, foi adotada uma metodologia criteriosa, detalhada nas subseções seguintes, com foco na preparação dos dados, otimização do treinamento e na mitigação de *overfitting*.

3.1.2.1 Preparação e Estruturação do Conjunto de Dados

O conjunto de dados inicial, composto por 4.067 imagens da classe “negativo” e 2.728 da classe “positivo”, revelou, em testes preliminares, uma falha na avaliação de desempenho devido a uma divisão de dados não representativa. Para garantir uma avaliação fidedigna do modelo VGG16, foi realizado um particionamento manual e estratificado do *dataset*.

1. **Divisão Estratificada:** Utilizou-se a função `train_test_split` da biblioteca *Scikit-learn* para segmentar o conjunto de dados, alocando 25% das amostras para o conjunto de teste. Foi empregado o parâmetro `stratify` para assegurar que a proporção original das classes fosse mantida tanto no conjunto de treino quanto no de teste. Esta etapa é fundamental para evitar vieses de avaliação e garantir que o desempenho do modelo seja medido em uma amostra estatisticamente representativa da distribuição real dos dados.
2. **Criação de Estrutura de Diretórios:** Com base na divisão estratificada, foi gerada uma nova estrutura de diretórios contendo subpastas distintas para `train/` e `test/`, cada uma com suas respectivas subpastas `positivo/` e `negativo/`. Os

arquivos de imagem foram subsequentemente copiados para seus locais designados, estabelecendo uma fonte de dados íntegra e bem definida.

3.1.2.2 Arquitetura do Modelo e Aprendizado por Transferência

A arquitetura VGG16 foi selecionada com base na estratégia de Aprendizado por Transferência (*Transfer Learning*). Esta abordagem aproveita os pesos da rede pré-treinada na base de dados *ImageNet*, o que é particularmente eficaz para *datasets* de tamanho limitado, pois permite que o modelo utilize características de baixo e médio nível já consolidadas (e.g., bordas, texturas) sem a necessidade de aprendê-las do zero.

A construção do modelo consistiu em:

- ❑ **Base Convolucional:** A base da VGG16 foi instanciada sem suas camadas de topo (`include_top=False`). Inicialmente, todas as camadas da base foram “congeladas” (`trainable=False`) para preservar os pesos pré-treinados durante a fase inicial de treinamento do novo classificador.
- ❑ **Classificador Personalizado:** Um novo classificador foi adicionado ao topo da base congelada, composto por uma camada `GlobalAveragePooling2D` (para redução de parâmetros e regularização), uma camada densa com 512 neurônios e ativação *ReLU*, uma camada `Dropout` com taxa de 0.5, e uma camada de saída densa com um neurônio e ativação *sigmoid* para a classificação binária.

3.1.2.3 Estratégia de Treinamento e Otimização

O treinamento foi conduzido em duas fases para otimizar a performance:

1. **Fase 1 - Treinamento do Classificador:** Nesta fase, com a base convolucional congelada, apenas as camadas do classificador personalizado foram treinadas. Isso permitiu que o novo classificador aprendesse a interpretar as características extraídas pela VGG16 de forma estável.
2. **Fase 2 - Ajuste Fino (Fine-Tuning):** Subsequentemente, as camadas do último bloco convolucional da VGG16 (a partir de `block5_conv1`) foram “descongeladas”. O modelo foi então re-treinado com uma taxa de aprendizado (*learning rate*) muito baixa ($1e-5$) para adaptar sutilmente as características mais especializadas às imagens do estudo, buscando um ganho de performance incremental.

Para a compilação, utilizou-se o otimizador `Adam` e a função de perda `binary_crossentropy`.

3.1.2.4 Técnicas para Prevenção de Overfitting

Uma preocupação central foi a mitigação de *overfitting*. As seguintes técnicas de regularização foram empregadas:

- ❑ **Aumento de Dados (*Data Augmentation*):** Foram aplicadas transformações aleatórias (rotações, zooms, deslocamentos e inversões) em tempo real nas imagens de treino para aumentar artificialmente a diversidade do *dataset*, promovendo a generalização.
- ❑ **Dropout:** Uma camada de Dropout com taxa de 50% foi inserida no classificador para desativar aleatoriamente neurônios durante o treinamento, forçando a rede a aprender representações mais robustas.
- ❑ **Parada Antecipada (*Early Stopping*):** Foi utilizada uma *callback* para monitorar a perda no conjunto de validação (*val_loss*). O treinamento foi configurado para ser interrompido automaticamente caso esta métrica não apresentasse melhora após 10 épocas consecutivas, evitando o sobreajuste aos dados de treino.

3.1.3 U-Net Adaptada

Para a terceira abordagem de classificação dos fragmentos de exames de imunofixação, foi selecionada a arquitetura de rede neural convolucional U-Net. Embora tenha sido originalmente proposta para tarefas de segmentação semântica de imagens biomédicas (RONNEBERGER; FISCHER; BROX, 2015), sua estrutura de codificador-decodificador com conexões de salto (*skip connections*) constitui um poderoso mecanismo para a extração de características hierárquicas, e essa capacidade de aprender representações ricas em diferentes escalas espaciais enquanto mantém a informação posicional das características da imagem pode ser eficaz para a tarefa de classificação binária proposta.

3.1.3.1 Preparação dos Dados e Estratégias de Mitigação de Sobreajuste

O conjunto de dados, composto por imagens pré-rotuladas como "positivo" ou "negativo", foi inicialmente processado para adequação ao modelo. Todas as imagens foram redimensionadas para uma resolução padrão de 128×128 pixels. O conjunto total foi então particionado, com 75% das amostras destinadas ao treinamento e validação, e os 25% restantes reservados como um conjunto de teste, utilizado apenas para a avaliação final do modelo generalizado. Do subconjunto de treinamento e validação, 20% foram separados para formar o conjunto de validação, essencial para monitorar o desempenho do modelo durante o treinamento e para a aplicação de técnicas de regularização.

Uma preocupação central no treinamento de redes neurais profundas, especialmente com conjuntos de dados de tamanho limitado, é o risco de sobreajuste (*overfitting*). Para mitigar este efeito e promover a capacidade de generalização do modelo, foi implementado um conjunto de técnicas de regularização:

- ❑ **Aumento de Dados (*Data Augmentation*):** O conjunto de amostras positivas do split de treino foi expandido artificialmente no processo de treinamento por

meio da aplicação de transformações geométricas e de intensidade aleatórias a cada imagem. As transformações incluíram rotações, inversões horizontais e verticais, escala, e alterações de contraste. Essa técnica de aumento artificial da quantidade de amostras de treino é comumente usada a fim de garantir que o conjunto de treino seja balanceado, uma propriedade importante para a qualidade de muitos classificadores.

- ❑ **Normalização em Lote (*Batch Normalization*):** Camadas de normalização em lote foram inseridas após as operações de convolução. Esta técnica normaliza as ativações de cada camada, estabilizando a distribuição dos dados de entrada para as camadas subsequentes. O resultado é uma convergência mais rápida e estável durante o treinamento, além de prover um efeito regularizador sutil.
- ❑ ***Dropout*:** Uma camada de *dropout*, com uma taxa de 0.5, foi aplicada na cabeça de classificação do modelo. Durante o treinamento, esta camada zera aleatoriamente 50% das ativações dos neurônios, impedindo que unidades vizinhas se tornem co-dependentes. Isso força a rede a aprender representações mais robustas e menos dependentes de neurônios específicos.
- ❑ **Parada Antecipada (*Early Stopping*):** O treinamento foi monitorado por um mecanismo de parada antecipada. A métrica de perda no conjunto de validação (*validation loss*) foi acompanhada a cada época, e o treinamento era interrompido caso não houvesse melhora nesta métrica por um número pré-definido de épocas consecutivas (paciência de 20). O modelo com o melhor desempenho no conjunto de validação foi então restaurado, garantindo que a versão final do modelo não correspondesse a um estado de sobreajuste.

3.1.3.2 Arquitetura e Adaptação para Classificação

A arquitetura U-Net canônica é composta por um caminho contrativo (codificador) e um caminho expansivo (decodificador). O codificador consiste em uma sequência de blocos convolucionais e operações de *max-pooling*, que reduzem progressivamente a resolução espacial da imagem enquanto aumentam a profundidade das características, capturando o contexto da imagem. O decodificador, por sua vez, utiliza convoluções transpostas para aumentar a resolução espacial, recuperando a localização dos detalhes. As conexões de salto ligam diretamente as saídas do codificador às entradas correspondentes do decodificador, permitindo que o modelo combine características semânticas de alto nível com informações espaciais de baixo nível.

O uso de uma arquitetura baseada em U-Net para extração de características, seguido por uma etapa de classificação, é uma técnica consolidada na área de imagens médicas. Para adaptar esta arquitetura a uma tarefa de classificação binária, a abordagem utilizada

neste trabalho se assemelha a outras aplicações encontradas na literatura para o mesmo fim (SUDHAN et al., 2022) (ALAKWAA; NASSEF; BADR, 2017). A principal modificação foi na saída final do decodificador que, em um problema de segmentação, seria um mapa de probabilidades. Para que a saída fosse uma única classificação binária, aplicou-se uma camada de *Global Average Pooling* GAP sobre o último mapa de características. Esta camada calcula a média de cada mapa, transformando-o em um único vetor que resume a informação espacial. Este vetor foi então processado por um classificador MLP denso, composto por uma camada *fully-connected* com ativação *ReLU* e uma camada de saída final com um único neurônio e ativação sigmoide, que produz a probabilidade de a imagem pertencer à classe "positivo".

3.1.3.3 Processo de Treinamento e Otimização de Hiperparâmetros

O modelo foi treinado utilizando a função de perda de Entropia Cruzada Binária (*Binary Cross-Entropy*), que é o padrão para problemas de classificação com duas classes. Devido a um desequilíbrio observado na distribuição de classes do conjunto de treinamento, foram calculados e aplicados pesos de classe inversamente proporcionais à frequência de cada classe. Esta abordagem penaliza mais fortemente os erros na classe minoritária, incentivando o modelo a não desenvolver um viés em favor da classe majoritária.

Como otimizador, foi selecionado o AdamW, uma variante do otimizador Adam que implementa de forma mais eficaz a técnica de decaimento de peso (*weight decay*), um método de regularização L2 que penaliza pesos de grande magnitude. A taxa de aprendizado foi gerenciada por um agendador do tipo *Cosine Decay*, que ajusta suavemente a taxa de aprendizado de um valor inicial de 1×10^{-4} até zero ao longo do treinamento, auxiliando o modelo a convergir para um mínimo mais estável e robusto da função de perda.

O desempenho do modelo foi avaliado com base em um conjunto de métricas: acurácia, precisão, *recall*, *F1-Score* e a Área Sob a Curva ROC (AUC). Após a conclusão do treinamento, uma etapa final de otimização foi realizada. O limiar (*threshold*) de classificação padrão de 0.5 foi ajustado. Para isso, as probabilidades de saída do modelo no conjunto de validação foram utilizadas para determinar um novo limiar, especificamente aquele que maximizava o *F1-Score* da classe "positivo". Este limiar otimizado foi então utilizado para gerar o relatório de classificação final no conjunto de teste, garantindo que a avaliação final refletisse o ponto de operação ótimo do modelo para a métrica de maior interesse.

3.1.4 Rede Neural Densa (DNN)

Uma *Deep Neural Network* (Rede Neural Densa) (DNN) foi proposta em conjunto com a CNN para a tarefa de classificação das imagens, explorando a combinação de aprendizado profundo e redes convolucionais para aprimorar a precisão do sistema. O

desenvolvimento desse modelo exigiu uma análise minuciosa dos dados, a fim de compreender o comportamento das imagens categorizadas como positivas e negativas e identificar padrões que poderiam influenciar a decisão do classificador. Além disso, essa etapa inicial permitiu uma investigação detalhada da relação entre as faixas de proteínas e a faixa de controle, aspecto essencial para garantir que o modelo fosse capaz de distinguir corretamente as classes de interesse. Essa abordagem cuidadosa visou ajustar os parâmetros da rede neural às características específicas das imagens processadas, assegurando que a extração e a interpretação das características relevantes fossem conduzidas de maneira eficiente.

Estudos prévios e a experiência de especialistas indicaram que o eixo vertical das imagens das faixas de exame contém informações essenciais para a correta interpretação dos resultados. Em contrapartida, variações na intensidade ao longo do eixo horizontal foram classificadas como ruído, conforme evidenciado por Keren et al. (KEREN, 2003). A análise dos padrões de intensidade dos pixels ao longo do eixo vertical permitiu identificar regiões de interesse associadas à presença de proteínas específicas, como as proteínas G e L. Esses picos de intensidade de cor, correspondentes a resultados positivos para determinadas proteínas (exemplificado na Figura 7, apresentaram uma forte correlação com os picos observados na faixa de controle, indicando uma possível relação direta entre suas características espectrais. A posição, largura e intensidade desses picos desempenham um papel fundamental na análise das imagens, uma vez que fornecem informações relevantes para a identificação e quantificação das proteínas presentes nas amostras analisadas. Dado o alto volume de dados gerados e a complexidade das relações entre as intensidades das faixas, a utilização de um modelo baseado em redes neurais profundas se mostrou apropriada, pois permite uma extração mais robusta de características e uma classificação mais precisa dos padrões identificados.

A Figura 8 apresenta um gráfico das curvas de intensidade dos pixels para cada coluna das imagens obtidas no exame de imunofixação. Essa visualização foi viabilizada por meio da técnica implementada para a extração de características de cada coluna da imagem, permitindo uma análise detalhada dos padrões de intensidade que correspondem às diferentes proteínas. Esse processo é fundamental para a segmentação e categorização das bandas proteicas, fornecendo subsídios para a avaliação automática das amostras e aprimorando a acurácia do diagnóstico computacional.

Diferentemente da abordagem baseada em CNN, que realiza a extração automática de características por meio de camadas convolucionais, o treinamento da rede neural densa seguiu uma estratégia distinta, fundamentada na extração explícita de um vetor de características descritivo a partir das imagens das faixas verticais. Esse processo teve início com a conversão das imagens em uma representação linear, permitindo que a rede capturasse padrões relevantes a partir de uma estrutura de dados simplificada.

Para garantir uma extração eficiente das características mais representativas das fai-

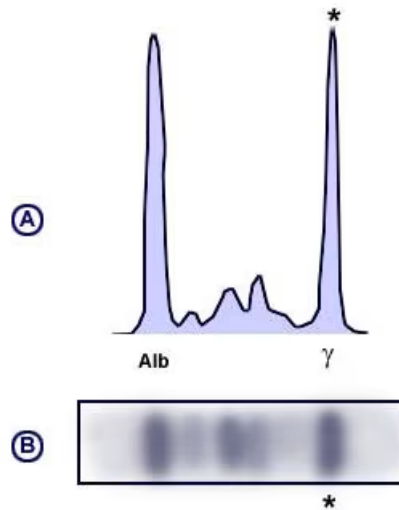


Figura 7 – Eletroforese de proteínas séricas demonstrando um padrão de gamopatia monoclonal. (A) O traçado densitométrico exibe um pico estreito e proeminente na região gama (indicado pelo asterisco), conhecido como componente monoclonal (ou Proteína M). (B) O gel de agarose correspondente mostra uma banda densa e restrita na mesma posição (asterisco), confirmando a presença da proteína monoclonal.

Fonte: <https://emedicine.medscape.com/article/2087113-overview>

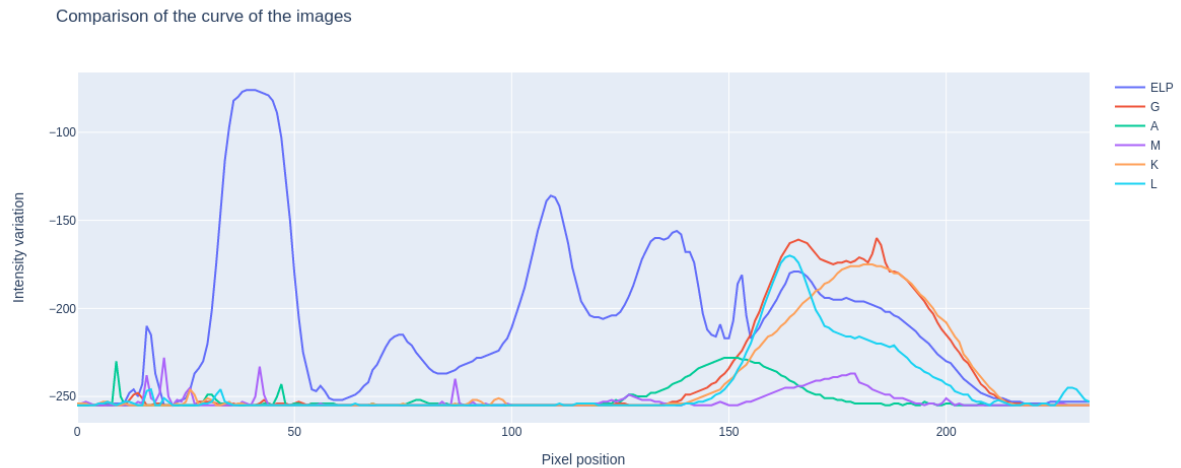


Figura 8 – Gráfico das curvas de intensidade dos pixels de cada coluna das imagens do exame de imunofixação. Essa visualização foi possibilitada pela técnica implementada para a extração de características de cada coluna de proteína do exame.

xas, optou-se por selecionar apenas a coluna central de cada tira como base para a formação do vetor de características. Essa coluna, que possui uma largura de um único pixel, foi convertida em um vetor unidimensional de dimensão $n = 150$, representado como $F = f_1, f_2, \dots, f_n$. Esse vetor preserva as informações essenciais sobre a intensidade dos pixels ao longo do eixo vertical, reduzindo a complexidade computacional sem compro-

meter a fidelidade dos dados extraídos. A Figura 9 ilustra esse fluxo de processamento, detalhando a sequência de operações empregadas para transformar as imagens em vetores numéricos interpretáveis pela rede neural.

Cada pixel pertencente à coluna central foi utilizado diretamente como uma entrada para a rede neural densa, que, por sua vez, foi projetada para conter 150 neurônios na camada de entrada, correspondendo exatamente à quantidade de elementos do vetor extraído. Esse design permitiu que o modelo realizasse uma análise detalhada e direta das características das faixas verticais, capturando padrões de intensidade e variações sutis que poderiam estar associadas à presença ou ausência de determinadas proteínas. Essa abordagem, ao aliar a extração manual de características à capacidade de aprendizado da rede neural densa, proporcionou um equilíbrio entre interpretabilidade e desempenho computacional, otimizando o processo de classificação das amostras analisadas.

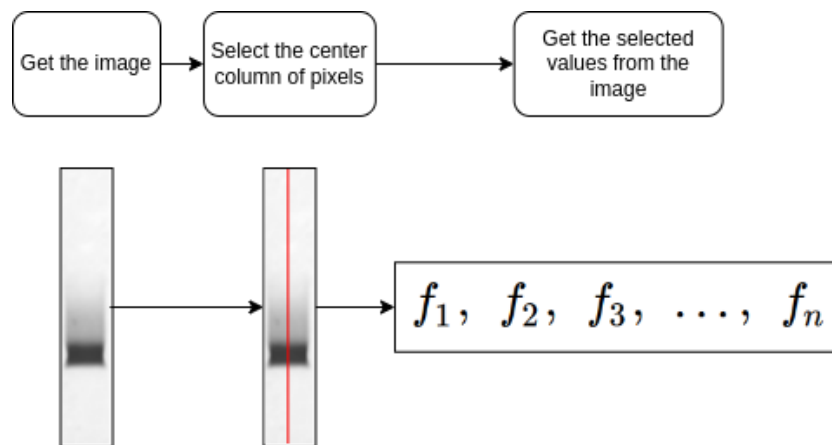


Figura 9 – O diagrama ilustra o processo de extração do vetor de características a partir das faixas verticais. As etapas incluem a aquisição da imagem, seleção da coluna central de pixels e obtenção dos valores extraídos, que são utilizados no treinamento da rede neural densa.

O conjunto de dados foi dividido em conjuntos de treinamento, validação e teste, mantendo as mesmas proporções de amostras, selecionadas aleatoriamente para garantir a representatividade. Durante o treinamento da rede neural densa, a otimização de hiperparâmetros envolveu múltiplos testes para definir a arquitetura e os hiperparâmetros ideais. Para mitigar o risco de sobreajuste (*overfitting*), foi empregada a técnica de Parada Antecipada (*Early Stopping*). Essa estratégia monitorava a perda no conjunto de validação e interrompia o treinamento caso não houvesse melhora após um número pré-definido de épocas (paciência), garantindo que o modelo salvasse apenas os pesos correspondentes ao seu melhor desempenho. O otimizador Adam foi utilizado, e a combinação de Parada Antecipada com um experimento de ajuste em menor escala ajudou a identificar as configurações ideais para o modelo final.

Após a conclusão dessa etapa de otimização, os parâmetros principais da rede neural foram definidos: a rede, contendo 4 camadas ocultas e um total de 1216 neurônios, foi

treinada com um tamanho de lote de 750 amostras. A função de perda utilizada foi a entropia cruzada binária e o inicializador de *kernel* padrão foi empregado. O treinamento foi configurado para um máximo de 2.500 épocas; no entanto, a parada antecipada atuou como o critério efetivo de conclusão, assegurando que o treinamento fosse finalizado no momento ótimo para evitar o sobreajuste e promover a capacidade de generalização do modelo para dados não vistos.

Considerações Finais sobre as Abordagens

As três abordagens implementadas, VGG16, U-Net Adaptada e Rede Neural Densa, foram selecionadas e desenvolvidas com o objetivo de explorar diferentes paradigmas de aprendizado profundo na tarefa de classificação das imagens de imunofixação. Cada uma das arquiteturas apresenta vantagens específicas em termos de capacidade de generalização, robustez na extração de características e complexidade computacional. Essa diversidade metodológica fornece uma base sólida para a comparação experimental entre os modelos, cujos resultados serão discutidos no próximo capítulo.

Experimentos

Neste capítulo, é realizada a apresentação do banco de dados utilizado no estudo, que contém imagens de exames de eletroforese por imunofixação com proteínas positivas e negativas. A partir dessas imagens, foram treinados e avaliados modelos de redes neurais, incluindo a VGG16, U-Net e uma rede neural densa, para classificar as amostras com base nas características das proteínas. Além disso, são discutidos os resultados obtidos por cada modelo, incluindo métricas como *F1-Score*, precisão e *recall*, além de analisar as matrizes de confusão e o desempenho dos modelos. Também é explorada a relação entre os resultados e as peculiaridades das imagens de eletroforese, que influenciam o desempenho das redes neurais. O capítulo segue com a avaliação dos resultados, comparando o desempenho das diferentes redes, com destaque para as características específicas do banco de dados e a escolha do vetor de características.

4.1 Descrição do Banco de Dados

O estudo utilizou um conjunto de 1359 imagens provenientes de um laboratório clínico, sendo 87 de exames completamente negativos e 1272 de exames com pelo menos uma proteína positiva. Essa distribuição mais balanceada permitiu uma média de aproximadamente duas amostras positivas e três negativas por exame, considerando a separação das faixas de proteína. O banco de dados resultante compreendia 6795 colunas de proteínas segmentadas, conforme descrito na Tabela 2, refletindo uma diversidade nas amostras utilizadas.

As imagens apresentadas na Figura 10 representam uma amostra do conjunto de imagens utilizadas no estudo, coletadas entre junho e agosto de 2022. Observa-se variação significativa no posicionamento das letras indicativas das proteínas, nos níveis de contraste e nos equipamentos utilizados durante os exames. Essas diferenças impactam diretamente na qualidade das imagens, tornando a análise mais complexa. Além disso, foram identificados diversos tipos de ruído, como artefatos gerados pelo gel de agarose e anotações manuais feitas com caneta, que também influenciam a interpretação visual dos resultados.

Tabela 2 – Quantidade de proteínas positivas e negativas.

Proteína e Rótulo	Quantidade de imagens	Percentual
IgG - Positivo	923	13,5%
IgA - Positivo	250	3,7%
IgM - Positivo	107	1,5%
κ - Positivo	812	11,9%
λ - Positivo	636	9,4%
IgG - Negativo	436	6,4%
IgA - Negativo	1109	16,4%
IgM - Negativo	1252	18,5%
κ - Negativo	547	8,0%
λ - Negativo	723	10,7%
TOTAL	6795	100,0%

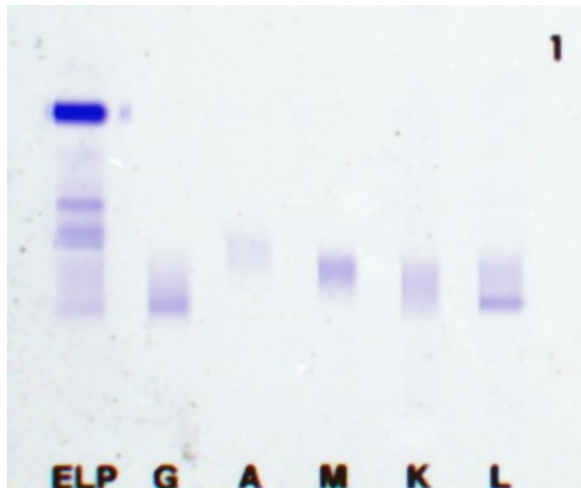
Para validar o modelo, foram realizadas avaliações especializadas em cada faixa de exame, com o objetivo de garantir a precisão e a confiabilidade dos resultados. Após ajustes nos parâmetros do modelo, foi conduzida uma validação cruzada com 10 iterações, utilizando diferentes conjuntos de dados para teste e treinamento. Esse processo de validação foi fundamental para fornecer uma avaliação abrangente e rigorosa do desempenho do modelo, permitindo uma análise precisa e confiável das métricas obtidas.

4.2 Resultados: VGG16

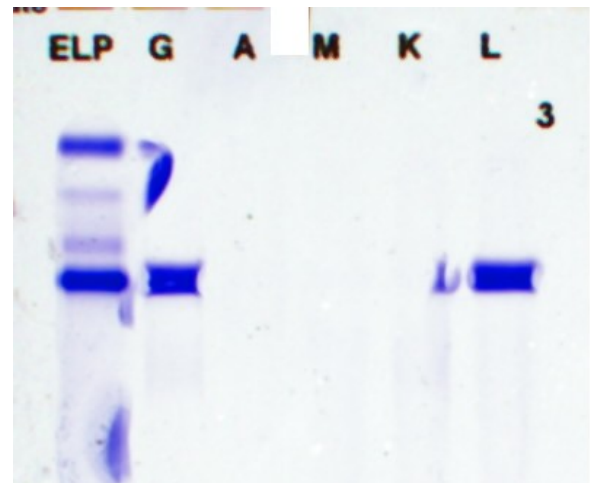
Após o treinamento e ajuste fino (*fine-tuning*) do modelo VGG16, foi realizada uma análise do limiar de decisão (*decision threshold*) com o objetivo de otimizar o desempenho para a tarefa de classificação, com foco particular no aumento da sensibilidade (*recall*) para a classe "positivo". A classificação binária padrão utiliza um limiar de 0.5 sobre as probabilidades geradas pela função de ativação *sigmoid*. Contudo, este valor nem sempre representa o ponto de operação ideal, especialmente em contextos onde o custo de Falsos Negativos é elevado.

Para determinar um ponto de corte otimizado, o modelo foi avaliado sistematicamente em um intervalo de limiares de 0.1 a 0.9. Os resultados demonstraram o clássico *trade-off* entre precisão e *recall*. Conforme o limiar era reduzido, o *recall* da classe "positivo" aumentava, ao custo de uma diminuição em sua precisão e no *recall* da classe "negativo". No limiar de 0.1, o modelo atingiu um *recall* máximo de 0.95 para a classe "positivo", identificando com sucesso 95% de todos os casos positivos. No entanto, este ganho foi acompanhado por uma queda acentuada na precisão para 0.72 e na acurácia geral do modelo para 0.83.

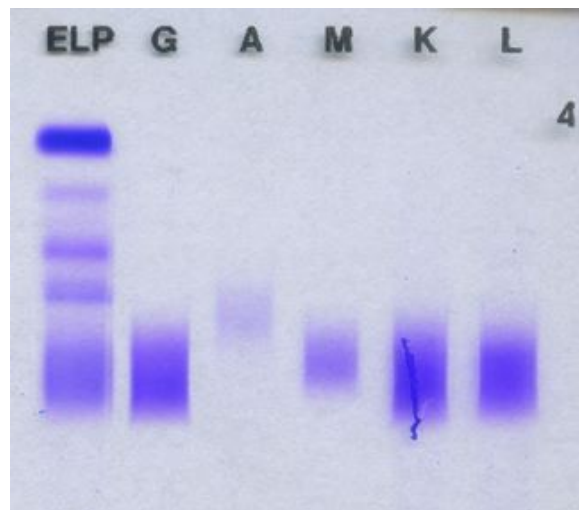
Considerando a necessidade de equilibrar a capacidade de detecção da classe de interesse com a robustez geral do modelo, o limiar de 0.3 foi selecionado como o ponto de



(a) Amostra positiva, nas proteínas G e L, com baixo contraste e bandas menos nítidas.



(b) Amostra positiva, nas proteínas G e L, com forte contraste e bandas bem definidas.



(c) Amostra negativa com alto nível de ruído e coloração mais uniforme.

Figura 10 – Amostra de imagens do banco de imagens utilizado.

operação ótimo. Esta escolha é justificada pelos seguintes fatores:

- ❑ Aumento Significativo do *Recall*: O limiar de 0.3 elevou o *recall* da classe "positivo" para 0.89, um aumento relevante em relação ao valor de 0.86 obtido com o limiar padrão de 0.5, cumprindo o objetivo primário de reduzir o número de Falsos Negativos.
- ❑ Manutenção de um Desempenho Equilibrado: Diferente dos limiares mais baixos, o ponto de corte de 0.3 manteve a precisão da classe "positivo" em um nível elevado (0.87) e preservou um excelente desempenho para a classe "negativo" (precisão de 0.92 e *recall* de 0.91).
- ❑ Melhor *F1-Score* Balanceado: O modelo alcançou um *F1-Score* médio (*macro* e

weighted average) de 0.90 com este limiar, indicando o melhor equilíbrio harmônico entre a precisão e o *recall* para ambas as classes.

Portanto, ao adotar o limiar de 0.3, o modelo final apresenta não apenas uma alta acurácia geral (0.90), mas também uma sensibilidade aprimorada para a classe "positivo", sem comprometer significativamente sua capacidade de classificar corretamente os casos negativos, conforme observado na Tabela 3. Esta otimização direcionada torna o modelo mais alinhado aos objetivos práticos da pesquisa.

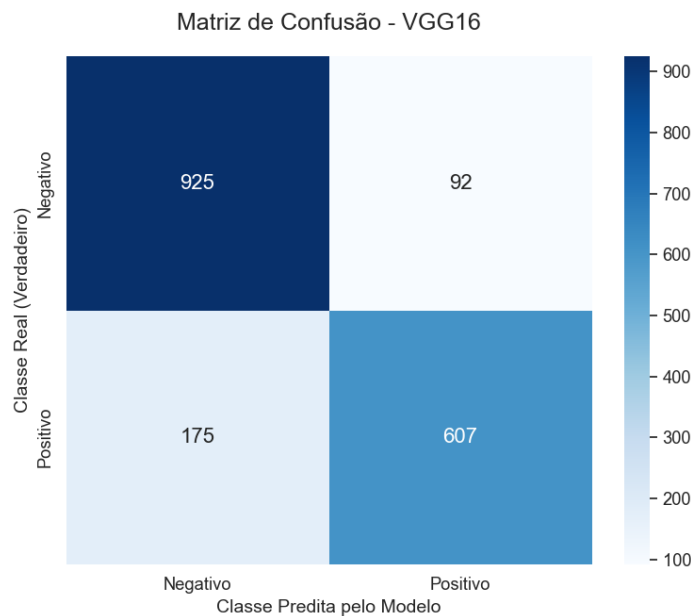


Figura 11 – Matriz de confusão dos resultados obtidos no conjunto de teste utilizando a rede neural VGG16.

Tabela 3 – Resultados de classificação do modelo VGG16 com limiar otimizado (0.3) no conjunto de teste.

Classe	Precisão	Recall	F1-score	Suporte
Negativo	0.92	0.91	0.92	1017
Positivo	0.87	0.89	0.88	682
Acurácia	0.90			1699

4.3 Resultados: U-Net

A avaliação do modelo U-Net treinado foi conduzida no conjunto de teste, que não foi utilizado em nenhuma etapa anterior do treinamento ou validação. A análise foi realizada em duas fases: primeiramente, utilizando o limiar de decisão padrão de 0.5, e subsequentemente, após um processo de otimização desse limiar.

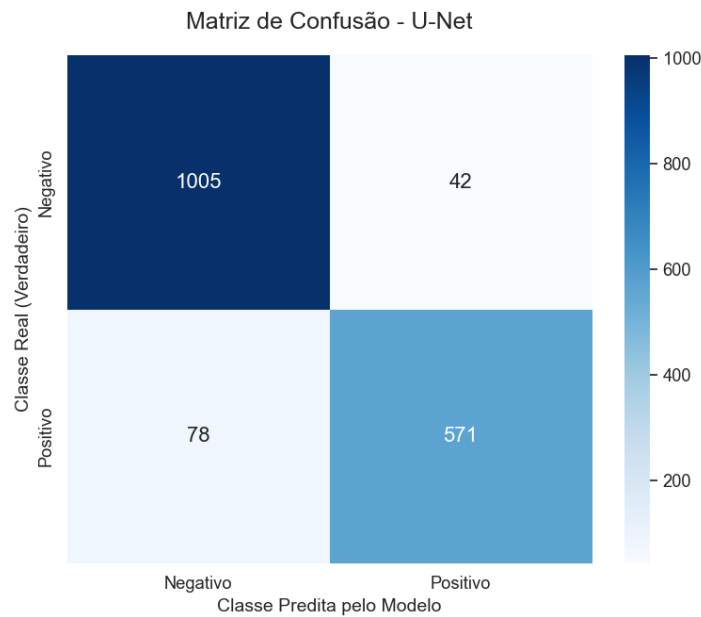


Figura 12 – Matriz de confusão dos resultados obtidos no conjunto de teste utilizando a rede neural U-Net adaptada.

Avaliação com Limiar de Classificação Padrão

Com o limiar de classificação padrão de 0.5, o modelo demonstrou um desempenho geral robusto, alcançando uma acurácia de 93% no conjunto de teste. A Tabela 4 detalha as métricas de precisão, *recall* e *F1-score* para cada classe. Enquanto a Figura 12 demonstra a matriz de confusão.

Tabela 4 – Resultados de classificação do modelo U-Net com limiar padrão (0.5) no conjunto de teste.

Classe	Precisão	Recall	F1-score	Suporte
Negativo	0.93	0.96	0.94	1047
Positivo	0.93	0.88	0.90	649
Acurácia		0.93		1696

A análise da Tabela 4 evidencia que o modelo foi altamente eficaz na identificação da classe ‘Negativo’, com um *F1-Score* de 0.94. Para a classe ‘Positivo’, o modelo obteve um *F1-Score* de 0.90, resultante de uma alta precisão (0.93) e um *recall* moderadamente alto (0.88). A precisão de 0.93 indica que, quando o modelo prevê uma amostra como positiva, ele está correto em 93% dos casos. O *recall* de 0.88 indica que o modelo identificou corretamente 88% de todas as amostras que eram de fato positivas.

Otimização do Limiar e Avaliação Final

Embora o *F1-Score* de 0.90 para a classe ‘Positivo’ seja um resultado sólido, buscou-se investigar se um ajuste no ponto de operação do classificador poderia otimizar o equilíbrio entre precisão e *recall* para esta classe. Conforme descrito na metodologia, um novo limiar de classificação foi determinado utilizando o conjunto de validação, com o objetivo de maximizar o *F1-Score* da classe ‘Positivo’. O valor ótimo encontrado foi de 0.45.

O modelo foi então reavaliado no conjunto de teste utilizando este novo limiar. Os resultados finais são apresentados na Tabela 5.

Tabela 5 – Resultados de classificação do modelo U-Net com limiar otimizado (0.45) no conjunto de teste.

Classe	Precisão	Recall	F1-score	Suporte
Negativo	0.93	0.95	0.94	1047
Positivo	0.92	0.89	0.91	649
Acurácia		0.9287		1696

A aplicação do limiar otimizado resultou em uma melhora na métrica de interesse para a classe ‘Positivo’. O *F1-Score* aumentou de 0.90 para 0.91. Este ganho foi obtido através de um aumento no *recall* (de 0.88 para 0.89), com uma leve e aceitável redução na precisão (de 0.93 para 0.92). Isto indica que a redução do limiar permitiu ao modelo identificar corretamente mais casos positivos que antes eram perdidos (reduzindo os falsos negativos), ao custo de introduzir um número marginalmente maior de falsos positivos.

Notavelmente, o desempenho para a classe ‘Negativo’ permaneceu estável com um *F1-Score* de 0.94, e a acurácia geral do modelo sofreu uma variação mínima, mantendo-se em 92.87%. Conclui-se que o ajuste de limiar foi uma etapa de pós-processamento bem-sucedida, que refinou o ponto de equilíbrio do modelo e otimizou sua performance para a detecção da classe ‘Positivo’ sem comprometer significativamente seu desempenho geral.

4.4 Resultados: Rede Neural Densa

A abordagem final consistiu na avaliação de uma rede neural densa (Perceptron de Múltiplas Camadas), que, diferentemente dos modelos convolucionais, opera sobre um vetor de características extraído das imagens, e não sobre os dados brutos de pixels. A performance deste modelo foi avaliada no conjunto de teste, e seus resultados são detalhados a seguir.

Análise de Desempenho de Classificação

A matriz de confusão para o modelo de rede neural densa é apresentada na Figura 13. Ela fornece uma visualização direta da performance de classificação do modelo.

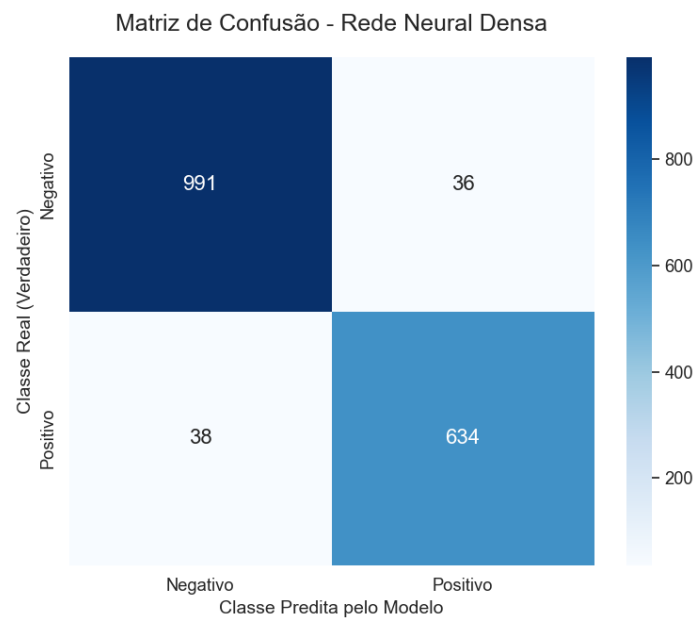


Figura 13 – Matriz de confusão dos resultados obtidos no conjunto de teste utilizando a rede neural densa sobre o vetor de características extraído.

Tabela 6 – Métricas de desempenho detalhadas do modelo de rede neural densa no conjunto de teste.

Classe	Precisão	Recall	F1-score	Suporte
Negativo	0.96	0.96	0.96	1027
Positivo	0.95	0.94	0.94	672
Acurácia	0.9564			1699

A análise da matriz revela um desempenho de classificação notavelmente alto. Das 1027 amostras pertencentes à classe ‘Negativo’, 991 foram corretamente identificadas, o que corresponde a uma Taxa de Verdadeiros Negativos (Especificidade) de 96,49%. Para a classe ‘Positivo’, de um total de 672 amostras, 634 foram classificadas corretamente, resultando em uma Taxa de Verdadeiros Positivos (*Recall* ou Sensibilidade) de 94,35%.

As métricas de desempenho por classe, calculadas a partir da matriz de confusão, estão sumarizadas na Tabela 6.

Os resultados indicam um desempenho excepcional e equilibrado. Ambas as classes apresentaram valores de *F1-Score* extremamente altos, 0.96 para a classe ‘Negativo’ e 0.94 para a classe ‘Positivo’, demonstrando que o modelo não apenas é preciso, mas também robusto na detecção de ambas as condições. A acurácia geral do modelo foi de 95,64%.

Análise de Eficiência e Padrões de Erro

Todos os experimentos, incluindo o treinamento e a avaliação dos modelos, foram conduzidos em um notebook equipado com um processador *Intel Core i5* de 12ª geração, 16

GB de *RAM* e uma *GPU NVIDIA GeForce GTX 1650*. Em termos de eficiência computacional, a rede neural densa apresentou um perfil de treinamento distinto em relação às CNNs. Devido ao menor tamanho dos dados de entrada (vetores de características em vez de imagens completas) e à ausência de operações convolucionais, o tempo de treinamento por época foi significativamente menor, aproximadamente 0,96 segundos. Em contrapartida, a CNN VGGNet demandou cerca de 8 segundos por época. Embora a rede neural densa tenha sido treinada por um número maior de épocas (2500 vs. 300 da CNN), o tempo total de treinamento para ambos os modelos foi similar, em torno de 40 minutos. Isso demonstra que houve equilíbrio entre a complexidade do modelo, a duração do treinamento por etapa, e o número de etapas de treinamento.

A inspeção visual e qualitativa dos erros de classificação revelou padrões consistentes. Os erros de falso positivo (classificar uma amostra como "Positivo" incorretamente) ocorreram com frequência em imagens que apresentavam manchas fortes e nítidas, mas cujas faixas não possuíam bordas claramente definidas. Por outro lado, os erros de falso negativo (não detectar uma amostra "Positivo") foram mais comuns em cenários onde uma mancha positiva bem definida se sobrepunha a uma mancha de fundo suave e com bordas difusas, sugerindo que a sobreposição de padrões pode confundir o classificador.

4.5 Avaliação dos Resultados

Uma análise comparativa dos resultados revela o desempenho notavelmente superior da MLP em relação às arquiteturas convolucionais (VGG16, U-Net). Enquanto o modelo U-Net otimizado alcançou uma acurácia de 92,87%, a rede neural densa atingiu uma acurácia de 95,64%, com *F1-Scores* equilibrados e elevados de 0.96 para a classe negativa e 0.94 para a positiva. A principal hipótese para essa discrepância reside na natureza do problema. As CNNs são projetadas com base no princípio de invariância à translação, uma propriedade vantajosa para diversas tarefas de classificação em que a característica de interesse pode aparecer em qualquer lugar da imagem. No entanto, no contexto da eletroforese por imunofixação, essa propriedade pode se tornar uma limitação, pois a posição espacial das bandas de proteína é determinante e semanticamente crucial para o diagnóstico. A importância do posicionamento das bandas nos exames torna a robustez das CNNs a variações de perspectiva e localização um atributo potencialmente prejudicial, o que pode explicar por que um modelo que não assume essa invariância, como a MLP, obteve melhor performance.

Adicionalmente, um fator determinante para a superioridade da rede neural densa foi a etapa de extração de características, que antecedeu a classificação. As arquiteturas de CNN operam diretamente sobre os dados brutos de pixels e, portanto, são expostas a todas as variações e ruídos presentes nas imagens, como artefatos do gel de agarose, bordas de faixas e anotações manuais, conforme descrito na Seção 2. O modelo convolucional é

forçado a aprender a ignorar essas informações irrelevantes, o que pode complexificar o treinamento e levar à aprendizagem de correlações espúrias. Em contrapartida, a abordagem com a rede neural densa utilizou um vetor de características que, por sua concepção, funciona como um filtro inteligente, isolando o perfil de intensidade da proteína e descartando ruídos e dados espacialmente irrelevantes. Essa representação mais enxuta dos dados de entrada se mostrou adequada para a tarefa de classificação na aplicação de exames de eletroforese de imunofixação, permitindo o treinamento de um modelo MLP concentrado nos aspectos mais relevantes dos dados, o que foi refletido em suas métricas de desempenho.

Conclusão

Neste capítulo, apresentamos as principais contribuições do trabalho desenvolvido, focando nas inovações e nos resultados obtidos na classificação de imagens de exames de imunofixação por eletroforese. A introdução estabeleceu objetivos gerais e específicos, como o desenvolvimento de um método eficaz para a classificação de imagens e a validação de abordagens alternativas que possam superar as redes neurais convolucionais tradicionais.

O desenvolvimento da pesquisa permitiu alcançar cada um desses objetivos de maneira clara e eficaz. O vetor de características simples, proposto e testado durante o estudo, não apenas superou a VGG16 e U-Net em termos de precisão, *recall* e *F1-Score*, mas também proporcionou uma redução significativa no tempo de processamento, essencial em contextos clínicos. Esses resultados confirmam que métodos mais simples e rápidos podem, em certos casos, ser superiores a abordagens complexas, atingindo os objetivos específicos traçados na introdução.

Assim, este capítulo não apenas resume as conquistas do trabalho, mas também evidencia como o processo de pesquisa contribuiu para atingir plenamente os objetivos estabelecidos, reforçando a relevância das contribuições apresentadas para a área de análise automatizada em diagnósticos médicos.

Antes de detalhar as contribuições e direções futuras, é pertinente revisitar as hipóteses formuladas no início desta pesquisa, confrontando-as com os resultados experimentais obtidos.

A primeira hipótese (H1) postulava que a estratégia de pré-processamento, baseada na segmentação de tiras verticais a partir do histograma de intensidade, seria eficaz para isolar as regiões de interesse. Os resultados confirmaram esta hipótese, uma vez que o método não apenas permitiu a correta separação das seis faixas de análise em todas as imagens do conjunto de dados, mas também resultou em uma redução dimensional significativa, com a área de interesse representando, em média, apenas 27,1% da imagem original. Essa eficácia foi fundamental para o sucesso das etapas subsequentes de classificação.

A segunda hipótese (H2) previa que uma arquitetura U-Net, originalmente concebida para segmentação e aqui adaptada para classificação, poderia atingir um desempenho comparável ao de uma CNN canônica para esta tarefa, como a VGG16. Esta hipótese foi igualmente validada. A rede U-Net adaptada obteve um $F1$ -Score de 0,91 para a classe positiva, superando ligeiramente o desempenho da VGG16, que alcançou 0,88 na mesma métrica. Tal resultado sugere que a capacidade da U-Net de preservar informações espaciais por meio de suas *skip connections* confere uma vantagem na análise de padrões posicionais fixos, característicos dos exames de imunofixação.

Finalmente, a terceira e principal hipótese (H3) propunha que uma rede neural densa (DNN), alimentada por um vetor de características extraído manualmente, poderia superar as arquiteturas de CNN em acurácia e $F1$ -Score. Os experimentos corroboraram fortemente esta hipótese. O modelo DNN alcançou um $F1$ -Score de 0,94 para a classe positiva e uma acurácia geral de 95,64%, métricas substancialmente superiores às obtidas pelas redes VGG16 (0,88 e 90,0%) e U-Net (0,91 e 92,87%). Este resultado evidencia que a extração manual de características, ao focar exclusivamente no perfil de intensidade vertical e descartar ruídos espaciais irrelevantes, forneceu uma representação de dados mais limpa e discriminativa.

5.1 Principais Contribuições

Este estudo apresentou avanços significativos na classificação de imagens de exames de imunofixação por eletroforese, propondo e validando três métodos distintos: um modelo de CNN VGG16, uma rede neural densa baseada em um vetor de características simples e uma adaptação da rede U-Net, voltada para classificação. As contribuições principais incluem:

1. Desenvolvimento de um vetor de características eficiente: A proposta de um vetor de características simples, que superou o desempenho da VGG16 e U-Net em todas as métricas avaliadas, demonstra a eficácia de soluções mais leves e rápidas para a análise de imagens médicas. Esse vetor de características permitiu alcançar um $F1$ -Score de 0,9625, uma precisão de 0,9628 e um *recall* de 0,9621, superando os resultados obtidos pelas demais redes neurais analisadas.
2. Redução do tempo de processamento: A utilização da rede neural densa, aliada ao vetor de características proposto, resultou em um tempo de processamento significativamente menor em comparação com as demais redes. Isso é particularmente relevante em contextos clínicos, onde a agilidade no diagnóstico pode impactar diretamente o tratamento do paciente.
3. Validação da hipótese: Os experimentos realizados validaram a hipótese de que métodos baseados em características manuais podem, em alguns casos, superar redes

neurais convolucionais complexas, especialmente em cenários onde o controle sobre as variáveis do modelo e a rapidez de processamento são cruciais.

Essas contribuições destacam a importância de explorar abordagens alternativas e otimizadas para a classificação de imagens em diagnósticos médicos, mostrando que a escolha cuidadosa de técnicas pode resultar em avanços significativos na precisão e eficiência dos métodos propostos.

5.2 Contribuições em Produção Bibliográfica

Como resultado direto deste projeto, foi publicado o capítulo de conferência abaixo, que divulga os principais métodos desenvolvidos e os resultados obtidos na classificação automatizada de exames de imunofixação eletroforética:

- Vilarinho Filho, A. C., Couto, L. N. & Souza, J. R. (2024). *Neural Networks for Classification of Immunofixation Electrophoresis Tests*. In: X. Le & Z. Zhang (Eds.), *Advances in Neural Networks – ISNN 2024*, LNCS 14827, pp. 600–609. Springer, Singapore. DOI:10.1007/978-981-97-4399-5_57

5.3 Trabalhos Futuros

Embora o método proposto tenha apresentado resultados promissores, há várias áreas onde melhorias podem ser implementadas e novas direções de pesquisa exploradas:

1. Exploração de outros conjuntos de dados e modelos: Um próximo passo lógico seria testar o método proposto em diferentes conjuntos de dados para verificar a generalização do modelo. Além disso, seria interessante explorar outras arquiteturas de redes neurais, como ResNet e DenseNet, para avaliar se combinações ou variações podem oferecer ganhos adicionais em precisão e eficiência.
2. Calibração baseada na faixa de controle: Uma melhoria potencial envolve a calibração dos modelos com base na análise da faixa de controle antes de processar as demais faixas do exame. Como a nitidez e o contraste da faixa de controle muitas vezes indicam os resultados esperados, essa calibragem pode aumentar ainda mais a precisão da classificação.
3. Refinamento do vetor de características: Embora o vetor de características desenvolvido manualmente tenha se mostrado eficaz, futuras pesquisas poderiam focar na otimização desse vetor, incorporando novas características extraídas manualmente que possam capturar informações ainda mais detalhadas das imagens, como detecção de picos, análise de comportamento da linha usando derivadas, e cálculo de estatísticas do vetor de características.

4. Desenvolvimento de uma ferramenta clínica integrada: A partir do sucesso deste estudo, um projeto futuro poderia envolver a criação de uma ferramenta clínica integrada que combine a análise automatizada com *feedback* em tempo real para apoio aos profissionais de saúde, permitindo uma revisão rápida e eficiente dos resultados, ao mesmo tempo em que sugere possíveis diagnósticos baseados em análises anteriores.
5. Investigação de novas técnicas de aprendizado de máquina: Por fim, explorar técnicas emergentes de aprendizado de máquina, como aprendizado por reforço e redes neurais profundas auto-supervisionadas, pode abrir novas possibilidades para melhorar a robustez e a adaptabilidade do sistema em diferentes cenários clínicos.

Esses projetos futuros têm o potencial de não apenas resolver possíveis limitações identificadas no presente trabalho, mas também expandir o impacto e a aplicabilidade das soluções propostas, contribuindo para o avanço contínuo da análise automatizada em diagnósticos médicos.

Referências

AKINYELU, A.; ADEWUMI, A. Lung lesion segmentation in ct images using improved u-net architecture. **Global Clinical Engineering Journal**, v. 5, n. 1, p. 88–96, 2023. Disponível em: <<https://openurl.ebsco.com/EPDB%3Aagd%3A5%3A748069/detailv?sid=ebsco%3Aplink%3Ascholar&id=ebsco%3Aagd%3A143320555>>.

ALAKWAA, W.; NASSEF, M.; BADR, A. Lung cancer detection and classification with 3d convolutional neural network (3d-cnn). **International Journal of Advanced Computer Science and Applications**, Science and Information (SAI) Organization Limited, v. 8, n. 8, 2017. DOI: 10.14569/IJACSA.2017.080853.

ANWAR, S. M. et al. Medical image analysis using convolutional neural networks: a review. **Journal of medical systems**, Springer, v. 42, n. 11, p. 1–13, 2018. DOI: 10.1109/TPAMI.2024.3435571.

CAI, L.; GAO, J.; ZHAO, D. A review of the application of deep learning in medical image classification and segmentation. **Annals of translational medicine**, AME Publications, v. 8, n. 11, 2020. DOI: 10.14569/IJACSA.2017.080853.

CAILLON, H. et al. Difficulties in immunofixation analysis: a concordance study on the IFM 2007-02 trial. **Blood Cancer J.**, Springer Science and Business Media LLC, v. 3, n. 10, p. e154, out. 2013. DOI: 10.1038/bcj.2013.51.

CAWLEY, L. P. et al. Immunofixation electrophoretic techniques applied to identification of proteins in serum and cerebrospinal fluid. **Clinical chemistry**, Oxford University Press, v. 22, n. 8, p. 1262–1268, 1976. DOI: 10.1093/clinchem/22.8.1262.

CHLAP, P. et al. A review of medical image data augmentation techniques for deep learning applications. **Journal of Medical Imaging and Radiation Oncology**, Wiley Online Library, v. 65, n. 5, p. 545–563, 2021. DOI: 10.1111/1754-9485.13261.

ÇUBUKÇU, H. C.; TOPCU, D. İ.; YENICE, S. Machine learning-based clinical decision support using laboratory data. **Clinical Chemistry and Laboratory Medicine (CCLM)**, De Gruyter, v. 62, n. 5, p. 793–823, 2024. DOI: 10.1515/ccm-2023-1037.

DODIG, S.; ČEPELAK, I.; DODIG, M. Are we ready to integrate advanced artificial intelligence models in clinical laboratory? **Biochemia medica**, Hrvatsko društvo za medicinsku biokemiju i laboratorijsku medicinu, v. 35, n. 1, 2025. DOI: 10.11613/BM.2025.010501.

- DU, G. et al. Medical image segmentation based on u-net: A review. **Journal of Imaging Science & Technology**, v. 64, n. 2, 2020. DOI: 10.2352/J.ImagingSci.Technol.2020.64.2.020508.
- ELFERT, E. et al. Expert-level detection of m-proteins in serum protein electrophoresis using machine learning. **Clinical Chemistry and Laboratory Medicine (CCLM)**, De Gruyter, v. 62, n. 12, p. 2498–2506, 2024. DOI: 10.1515/cclm-2024-0222.
- GILES, H. V.; KARUNANITHI, K. Performance characteristics and limitations of the available assays for the detection and quantitation of monoclonal free light chains and new emerging methodologies. **Antibodies**, v. 13, n. 1, 2024. ISSN 2073-4468. DOI: 10.3390/antib13010019.
- GOODFELLOW, I. **Deep learning**. [S.l.]: MIT press, 2016.
- HATAMIZADEH, A. et al. Unetr: Transformers for 3d medical image segmentation. In: **Proceedings of the IEEE/CVF winter conference on applications of computer vision**. [S.l.: s.n.], 2022. p. 574–584. DOI: 10.1109/WACV51458.2022.00181.
- HU, H. et al. Expert-level immunofixation electrophoresis image recognition based on explainable and generalizable deep learning. **Clinical Chemistry**, Oxford University Press US, v. 69, n. 2, p. 130–139, 2023. DOI: 10.1093/clinchem/hvac190.
- KEREN, D. **Protein electrophoresis in clinical diagnosis**. [S.l.]: CRC Press, 2003. DOI: 10.1201/b13302.
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. **Communications of the ACM**, AcM New York, NY, USA, v. 60, n. 6, p. 84–90, 2017. DOI: 10.1145/3065386.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **nature**, Nature Publishing Group UK London, v. 521, n. 7553, p. 436–444, 2015. DOI: 10.1038/nature14539.
- LEE, N. et al. Development and validation of a deep learning-based protein electrophoresis classification algorithm. **PloS one**, Public Library of Science San Francisco, CA USA, v. 17, n. 8, p. e0273284, 2022. DOI: 10.1371/journal.pone.0273284.
- LI, H. et al. Learning to detect monoclonal protein in electrophoresis images. In: **IEEE. 2021 International Conference on Visual Communications and Image Processing (VCIP)**. [S.l.], 2021. p. 1–5. DOI: 10.1109/VCIP53242.2021.9675332.
- O’CONNELL, T. X.; HORITA, T. J.; KASRAVI, B. Understanding and interpreting serum protein electrophoresis. **American family physician**, v. 71, n. 1, p. 105–112, 2005.
- PRIDDY, K. **Artificial Neural Networks: an Introduction**. [S.l.]: Prentice-Hall, India, 2005. DOI: 10.1117/3.633187.
- QASSIM, H.; VERMA, A.; FEINZIMER, D. Compressed residual-vgg16 cnn model for big data places image recognition. In: **IEEE. 2018 IEEE 8th annual computing and communication workshop and conference (CCWC)**. [S.l.], 2018. p. 169–175. DOI: 10.1109/CCWC.2018.8301729.

- RAWAT, W.; WANG, Z. Deep convolutional neural networks for image classification: A comprehensive review. **Neural computation**, MIT Press, v. 29, n. 9, p. 2352–2449, 2017. DOI: 10.1162/neco_a00990.
- RODRIGUEZ, E. L. et al. Affinity chromatography: A review of trends and developments over the past 50 years. **Journal of Chromatography B**, Elsevier, v. 1157, p. 122332, 2020. DOI: 10.1016/j.jchromb.2020.122332.
- RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. **Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015**, Springer, p. 234–241, 2015. DOI: 10.1007/978-3-319-24574-4_28.
- SIDDIQUE, N. et al. U-net and its variants for medical image segmentation: A review of theory and applications. **IEEE access**, IEEE, v. 9, p. 82031–82057, 2021. DOI: 10.1109/ACCESS.2021.3086020.
- SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. **arXiv preprint arXiv:1409.1556**, 2014.
- SUDHAN, M. et al. Segmentation and classification of glaucoma using u-net with deep learning model. **Journal of Healthcare Engineering**, Wiley Online Library, v. 2022, n. 1, p. 1601354, 2022. DOI: 10.1155/2022/1601354.
- THIEMANN, C. et al. Automated assessment of immunofixations with deep neural networks. **LaboratoriumsMedizin**, De Gruyter, 2022. DOI: 10.1515/labmed-2022-0078.
- WEI, X.-Y. et al. Deep collocative learning for immunofixation electrophoresis image analysis. **IEEE Transactions on Medical Imaging**, IEEE, v. 40, n. 7, p. 1898–1910, 2021. DOI: 10.1109/TMI.2021.3068404.
- ZACHARIOU, M. **Affinity chromatography: methods and protocols**. [S.l.]: Springer, 2008. v. 421. DOI: 10.1007/978-1-59745-582-4.
- ZHANG, Y.; LIU, X.; WANG, J. Enhanced u-net with attention mechanism for brain tumor segmentation. **IEEE Access**, v. 12, p. 1–10, 2024.