

# Classificação da Camada Lipídica do Filme Lacrimal usando *K-Means* e *Deep Learning*

Domingos A. Dias Junior<sup>1</sup>, Luana B. da Cruz<sup>1</sup> João O. B. Diniz<sup>2</sup>

<sup>1</sup>Laboratório de Inteligência Computacional Aplicada (LICA) —  
Universidade Federal do Cariri (UFCA)  
CEP 63.048-080 — Juazeiro do Norte — CE — Brasil

<sup>2</sup>Instituto Federal de Educação, Ciência e Tecnologia do Maranhão (IFMA)  
Caixa Postal 65.940-000 — Grajaú — MA — Brasil

domingos.adj@nca.ufma.br, luana.batista@ufca.edu.br,  
joao.bandeira@ifma.edu.br

**Abstract.** *Dry eye syndrome, one of the most common ophthalmic conditions, poses diagnostic challenges due to its multifactorial nature. Traditional diagnosis, involving manual classification of images of the tear film, is limited by the instability of this film. This study proposes an innovative approach that combines unsupervised learning for region of interest segmentation and deep learning for patch classification. The results are promising, with an accuracy of 99.23% and an F1-score of 99.16%, surpassing other techniques and studies in the literature. It is believed that this methodology could be a valuable tool to assist experts in disease diagnosis.*

**Resumo.** *A síndrome do olho seco, uma das doenças oftalmológicas mais comuns, apresenta desafios diagnósticos devido à sua natureza multifatorial. O diagnóstico tradicional, que envolve a classificação manual das imagens do filme lacrimal, enfrenta limitações devido à instabilidade deste filme. Este estudo propõe uma abordagem inovadora que combina aprendizagem não supervisionada para segmentação da região de interesse e deep learning para classificação de patches. Os resultados são promissores, com acurácia de 99,23% e F1-score de 99,16%, superando outras técnicas e estudos da literatura. Acredita-se que esta metodologia possa ser uma ferramenta valiosa para auxiliar os especialistas no diagnóstico da doença.*

## 1. Introdução

O filme lacrimal, uma estrutura multicamadas, é uma composição essencial para a manutenção da saúde ocular. Este filme é uma fina película que reveste a superfície dos olhos, desempenhando funções cruciais na proteção e no funcionamento adequado do sistema visual. As camadas incluem a externa, responsável por evitar a evaporação da água; a aquosa intermediária, que nutre a córnea e a protege de corpos estranhos; e a mucosa interna, que adere as lágrimas aos olhos formando uma película protetora [Stapleton et al. 2017]. Esse complexo conjunto de camadas desempenha funções específicas, visando manter a homeostase e a integridade da superfície ocular. A análise da camada lipídica é essencial, uma vez que sua deficiência pode levar à síndrome do olho seco evaporativo [Marculino et al. 2022].

A síndrome do olho seco é uma condição crônica e progressiva que afeta a qualidade de vida, manifestando-se em sintomas prevalentes, como desconforto, irritação, ressecamento e visão turva [Marculino et al. 2022, Messmer 2015]. O diagnóstico é atualmente baseado na combinação de sintomas e sinais relatados pelos pacientes, sendo mais comum em mulheres, especialmente com o avanço da idade [Marculino et al. 2022]. A prevalência da síndrome varia consideravelmente, abrangendo de 5% a 50% da população [Stapleton et al. 2017]. Fatores de risco, como o uso de lentes de contato, cirurgias oculares, tabagismo, diabetes e influências ambientais, contribuem para a instabilidade do filme lacrimal [Messmer 2015].

O diagnóstico da síndrome do olho seco é considerado uma tarefa desafiadora para os oftalmologistas devido à sua etiologia multicamada. Uma das abordagens de avaliação mais comuns concentra-se na análise da camada lipídica, uma vez que fornece informações essenciais para identificar as causas subjacentes e orientar o tratamento [Messmer 2015]. Essa avaliação pode ser realizada utilizando instrumentos como o Interferômetro Doane, que capturam imagens da camada lipídica permitindo aos especialistas analisar as alterações morfológicas [Doane 1989].

De acordo com vários estudos, classificar as categorias da camada lipídica é uma tarefa clínica difícil, especialmente nas mais finas que não apresentam características morfológicas e/ou de cor. Porém, métodos computacionais têm sido essenciais para auxiliar no diagnóstico da síndrome do olho seco. O trabalho de [G. Villaverde et al. 2014] utilizaram os filtros de *Butterworth* e *Gabor*, a Transformada Discreta de *Wavelet*, os campos aleatórios de *Markov* e as características de matrizes de coocorrência para extrair características de textura. Essas características foram analisadas em escala de cinza e no espaço de cor  $L^*a^*b^*$ . Em seguida, são utilizados os algoritmos *Correlation-Based Feature Selection* (CFS), filtro baseado em consistência e INTERACT. O método alcançou 91,51% de acurácia, utilizando como técnica de classificação o *Support Vector Machine* (SVM).

O método proposto por [Remeseiro et al. 2015] também utilizou informações de cor e textura. Em comparação com o trabalho de [G. Villaverde et al. 2014], este incluiu as cores oponentes do espaço de cor RGB. Ambos os métodos utilizaram os mesmos extratores de características, mas [Remeseiro et al. 2015] removeram a seleção de características. Os resultados alcançados mostraram 93,4% de acurácia usando os algoritmos de aprendizagem de máquina *Naive Bayes*, *Random Tree*, *Random Forest* e SVM.

No estudo realizado por [Cruz et al. 2018a], foram incorporadas informações de cor e textura em seu método, com a utilização de algoritmos de *machine learning* para seleção e classificação de características. Para descrever a textura foi usada a função K de *Ripley*, aplicando-a nos espaços de cores baseado em cores oponentes,  $L^*a^*b^*$ , HSV e RGB. A representação da imagem foi realizada por meio do *Local Binary Pattern* (LBP). Após extrair as características, realizaram a etapa de seleção utilizando o algoritmo *Greedy Stepwise*. O método alcançou uma acurácia de 95,66%, empregando o classificador *Naive Bayes*.

Os trabalhos conduzidos por [Cruz et al. 2018b, da Cruz et al. 2020] propuseram métodos baseados em diferentes índices de diversidade filogenética para extrair características de textura das imagens em escala de cinza. As abordagens propostas por [Cruz et al. 2018b, da Cruz et al. 2020] apresentaram resultados de acurácia de 96,98%

e 97,54%, respectivamente, ambos empregando o *Random Forest*. O método do trabalho de [da Cruz et al. 2020] abordaram o uso dos extratores de textura K de *Ripley* e diversos índices de diversidade nos espaços de cores HLS e escala de cinza. Para selecionar as características mais relevantes usaram o *Greedy Stepwise*. No melhor cenário, apresentaram resultados de 99,62% de acurácia.

É importante destacar que todos os estudos apresentados incorporam uma etapa de segmentação da região de interesse, seguindo a abordagem proposta por [G. Villaverde et al. 2014]. A técnica adotada utiliza limiarização no canal verde (G) de uma imagem RGB para extrair a região de interesse, seguida da redução de um retângulo mapeado dentro desta região até a ausência de área de fundo. Adicionalmente, é relevante mencionar que a base de imagens VOPTICAL\_GCU [VOPTICAL\_GCU 2017], composta por 106 imagens, foi consistentemente utilizada em todos os trabalhos analisados. Esta escolha deve-se ao fato de ser a única base de imagens pública identificada na literatura para as análises em questão, proporcionando uma base uniforme de comparação entre estudos e contribuindo para a coesão e comparabilidade dos resultados. Ao examinar os estudos relacionados, torna-se evidente a pesquisa na análise de textura para extração de características.

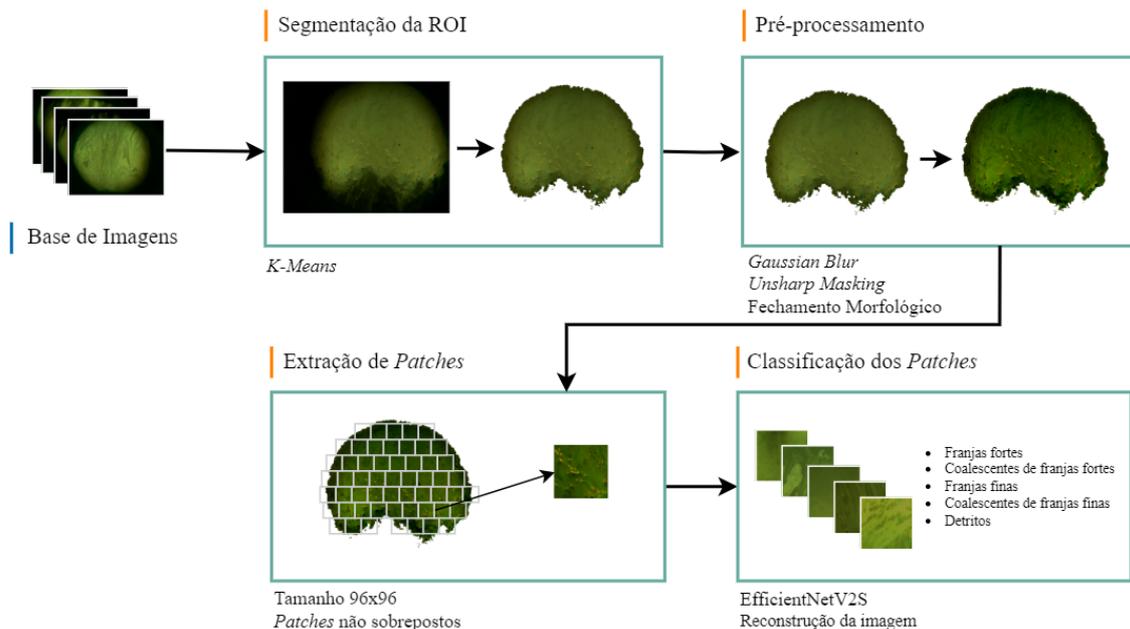
O trabalho proposto apresenta uma abordagem inovadora para a classificação da camada lipídica do filme lacrimal, destacando-se pelas seguintes contribuições: (1) Incorpora técnicas contemporâneas ao utilizar aprendizagem não supervisionada, notadamente o algoritmo *K-Means*, para realizar a segmentação de a camada lipídica; (2) Aplica uma combinação de técnicas de pré-processamento, incluindo filtros, com o objetivo de melhorar a qualidade da imagem antes da análise; e (3) Emprega aprendizagem profunda, especificamente o modelo EfficientNetV2S, na etapa de classificação, oferecendo uma abordagem inovadora para investigar a camada lipídica. Essas contribuições refletem uma abordagem abrangente e inovadora, resultando em avanços significativos na área.

## 2. Método Proposto

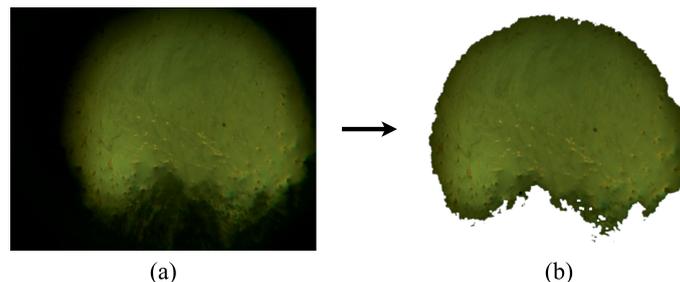
O método proposto consiste em quatro etapas fundamentais, cuja validação é conduzida utilizando uma base de imagens detalhada na subseção 3.2. Na primeira etapa, é realizada a segmentação da região de interesse (ROI) nas imagens interferométricas. A segunda etapa compreende o pré-processamento das imagens segmentadas, visando destacar as informações relevantes. A terceira etapa envolve a extração de *patches* específicos dentro da ROI, permitindo uma análise mais detalhada. Na etapa final de classificação, os resultados são obtidos e, posteriormente, validados. A Figura 1 oferece uma ilustração dessas etapas, enquanto os detalhes são explorados nas seções subsequentes.

### 2.1. Segmentação da ROI

A segmentação consiste em separar objetos de interesse de informações irrelevantes presentes em imagens digitalizadas. As imagens de entrada adquiridas com o Interferômetro Doane incluem áreas irrelevantes presentes nas áreas externas e sombras dos cílios, conforme ilustrado na Figura 2 (a). Assim, a região mais relevante para segmentação está na parte central da área amarelada ou esverdeada da imagem, formada pela superfície anterior do filme lacrimal que recobre a córnea. Portanto, foi realizada uma etapa de extração da ROI utilizando *K-Means* [Davidson 2002] ( $K = 2$ ), que é uma técnica de agrupamento.



**Figura 1. Método proposto.**



**Figura 2. Segmentação: a) Imagem original; b) Imagem segmentada com *K-Means*.**

Essa abordagem permitiu destacar a área de maior importância para análise (Figura 2 (b)), contribuindo assim para a eficácia subsequente das etapas do método proposto.

## 2.2. Pré-processamento

Nesta etapa, as imagens em RGB foram submetidas a técnicas de melhoria usando os filtros *Gaussian Blur* [Gedraite and Hadad 2011] e *Unsharp Masking* [Panetta et al. 2011]. O *Gaussian Blur* é empregado para suavizar a imagem, atenuando detalhes finos e reduzindo o ruído, resultando em uma representação mais uniforme (Figura 3 (a)). Em seguida, foi aplicado o *Unsharp Masking*, que destaca as bordas e realça os detalhes, enfatizando características importantes na imagem (Figura 3 (b)). Essa combinação de técnicas teve como objetivo aprimorar a qualidade visual das imagens, preparando-as para um desempenho aprimorado e generalista do modelo de classificação.

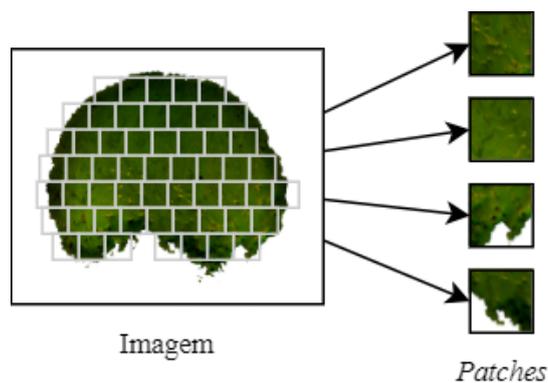
## 2.3. Extração de *Patches*

Na terceira etapa, realizou-se a extração de *patches* no interior da ROI, com dimensões de  $96 \times 96$  e sem sobreposição (Figura 4). Vários experimentos foram realizados para avaliar diferentes tamanhos de *patches*, sendo  $96 \times 96$  identificado como o mais adequado para



**Figura 3. Pré-processamento: (a) Imagem segmentada; (b) Imagem com *Gaussian Blur*, *Unsharp Masking* e fechamento morfológico.**

representar eficientemente a região da camada lipídica. Tamanhos menores levaram à perda de informações relevantes, devido à falta de contexto significativo, comprometendo assim a capacidade de capturar características distintivas da região e resultando em uma representação inadequada. Por outro lado, tamanhos maiores resultaram em uma extração limitada de *patches*, comprometendo a capacidade do modelo em capturar detalhes finos presentes na região [Quintana et al. 2023]. Durante a extração foram descartados *patches* com menos de 30% de pixels pertencentes a ROI, garantindo uma representação mais robusta e informativa da região.



**Figura 4. Ilustração da extração de *patches*.**

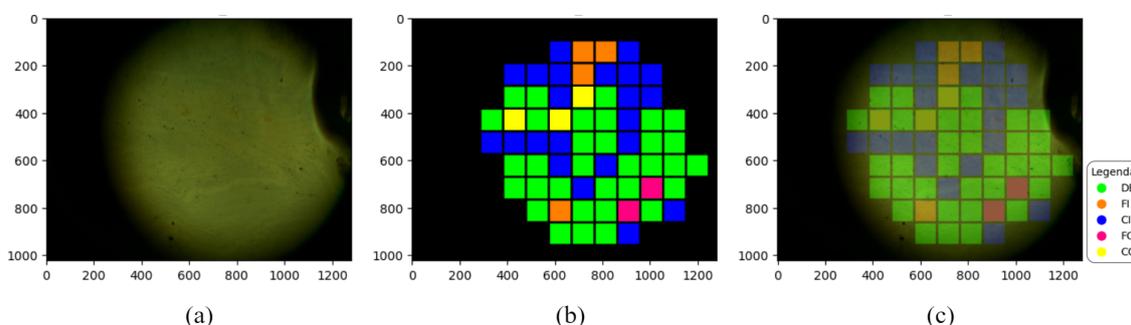
A seleção cuidadosa do tamanho dos *patches* visa otimizar a qualidade e quantidade de informações extraídas, uma vez que, a base de imagens usada para os experimentos possui poucas imagens. Em média foram extraídos 63 *patches* de cada imagem, totalizando 6.679 *patches*.

#### **2.4. Classificação dos *Patches***

Finalmente, os *patches* foram submetidos à classificação nas categorias da camada lipídica do filme lacrimal. Nesse processo, medidas descritivas foram obtidas dos *patches* extraídos para classificar as categorias da síndrome do olho seco. Para a conduzir os experimentos, foram selecionadas cinco arquiteturas de rede: DenseNet-121 [Huang et al. 2016], EfficientNetV2S [Tan and Le 2021], ResNet-50 [Koonce and Koonce 2021], VGG-16 [Simonyan and Zisserman 2014], *Vision Transformer* (ViT) [Dosovitskiy et al. 2020]. Vale ressaltar que para facilitar o treinamento foi usado modelos pré-treinados (com a base ImageNet) com as configurações padrão da

biblioteca Keras. Além disso, para melhorar a capacidade de generalização, técnicas de *data augmentation online* foram aplicadas aos modelos. O processo de *data augmentation* incluiu operações como rotação, zoom, deslocamento horizontal e vertical, cisalhamento e espelhamento horizontal, proporcionando uma maior diversidade ao conjunto de treinamento e contribuindo para um desempenho robusto dos modelos pré-treinados.

Após a etapa de classificação, os *patches* foram remontados para formar a imagem. Os *patches* foram submetidos a uma avaliação para determinar a categoria em que cada um foi classificado. A categoria que apresentou maior frequência foi considerada como a categoria dominante para a imagem em questão, fornecendo a definição da categoria à qual pertence. Para representação visual das categorias, os *patches* foram coloridos de maneira distinta (Figura 5): verde para a categoria detritos (DE); laranja para franjas finas (FI); azul para coalescentes de franjas finas (CI); rosa para franjas fortes (FO); e amarelo para coalescentes de franjas fortes (CO). Na Figura 5, a categoria definida seria CI.



**Figura 5. Patches remontados e determinados: (a) Imagem original; (b) Predição; (c) Predição sobre a imagem.**

### 3. Resultados

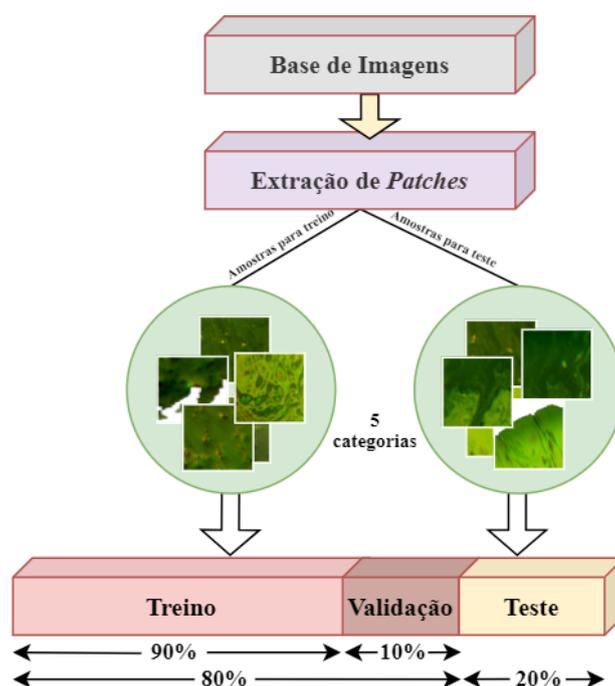
Esta seção aborda a configuração experimental, os procedimentos realizados para validar o método proposto, as métricas de validação adotadas, e detalha o procedimento de preparação da base de imagens.

#### 3.1. Configuração Experimental

O método foi implementado usando a linguagem Python. Utilizou-se principalmente a biblioteca de *deep learning* Keras [Chollet et al. 2015] com TensorFlow-GPU [Abadi et al. 2015] como *back-end*. O computador utilizado para os experimentos consiste em uma CPU Intel Core i5-9300H de 2.40GHz, 24 GB de RAM e placa de vídeo Nvidia GeForce GTX 1650, rodando em um sistema operacional Windows 11.

#### 3.2. Resultados Experimentais

A base VOPTICAL\_GCU [VOPTICAL\_GCU 2017] foi usada para os experimentos. Essa base contém 106 imagens das cinco categorias consideradas: 11 FO, 25 CO, 30 FI, 26 CI e 14 DE. As imagens têm uma resolução de  $1280 \times 1024$  pixels no espaço de cor RGB. Aproximadamente 80% de todo o conjunto de dados foi usado para treinar o modelo e os 20% restantes foram usados para teste (1.336 *patches*). Do total de 80% para treinamento (4.675 *patches*), 20% foram usados para validação (668 *patches*). A Figura 6 ilustra a divisão da base de imagens.



**Figura 6. Distribuição de *patches* durante a fase de treinamento e teste.**

Inicialmente, foi realizado a segmentação da ROI, seguido da aplicação dos filtros e extração dos *patches*. Para validar o método, foi aplicado a técnica de validação cruzada *k-fold*, definindo  $k=5$ . Nessa abordagem, o método é treinado e testado cinco vezes, em cada categoria da camada lipídica do filme lacrimal da base de imagens. Isso permitiu análises mais adequadas do método, uma vez que a eficiência do método é verificada independentemente do conjunto de dados avaliado, garantindo que o método não seja tendencioso [Burman 1989]. Também foi possível evitar o problema de *overfitting*, pois cada categoria foi treinado pelo menos uma vez no procedimento de validação *k-fold*.

Por fim, para avaliar os resultados foram empregadas métricas amplamente reconhecidas no campo de *machine learning*, sendo elas: acurácia (Acc), precisão (Prec), sensibilidade (Sen), *F1-score* (F1) e *Kappa* (Kap) [Leite et al. 2020, Pereira et al. 2016, Júnior et al. 2021]. As métricas foram calculadas com base na matriz de confusão. Como resultado, foi obtido a média de cada métrica junto com seu desvio padrão.

A seguir são apresentados os experimentos realizados para validar o método proposto. Inicialmente são apresentados os resultados obtidos sem a aplicação da etapa de pré-processamento, seguidos dos resultados após a incorporação desta etapa. Para isso, foram utilizadas diversas arquiteturas de redes neurais com as mesmas configurações.

### 3.3. Experimento 1 - Validação da Etapa sem Pré-processamento

A Tabela 1 descreve os resultados produzidos sem a aplicação da etapa de pré-processamento. Observa-se que os resultados obtidos com a arquitetura EfficientNetV2S revelaram um desempenho notável no contexto do experimento. A acurácia atingiu 95,13%, indicando um índice significativo de classificações corretas em relação ao total de instâncias avaliadas. Além disso, a precisão de 92,43% destaca a capacidade do modelo em realizar classificações positivas de forma assertiva. A sensibilidade, avaliada

em 94,00%, demonstra a capacidade do modelo em identificar corretamente instâncias positivas. O *F1-score*, uma métrica que pondera precisão e sensibilidade, alcançou 93,04%, evidenciando um equilíbrio entre a capacidade de prever corretamente ambas as classes. Por fim, o coeficiente *Kappa*, calculado em 93,63%, fornece uma medida adicional da concordância além do acaso, reforçando a consistência e confiabilidade das classificações realizadas pelo EfficientNetV2S no contexto do estudo.

**Tabela 1. Resultados da comparação do método proposto com outras arquiteturas sem a etapa de pré-processamento.**

Arquitetura	Acc(%)	Prec(%)	Sen(%)	F1(%)	Kap(%)
DenseNet-121	88,7 ± 2,13	90,71 ± 2,92	86 ± 1,33	86,29 ± 0,84	85,65 ± 2,69
<b>EfficientNetV2S</b>	<b>95,13 ± 7,75</b>	<b>92,43 ± 13,76</b>	<b>94 ± 10,41</b>	<b>93,04 ± 12,3</b>	<b>93,63 ± 10,2</b>
ResNet-50	91,3 ± 3,89	91,38 ± 5,17	89,33 ± 3,89	89,64 ± 4,54	88,95 ± 4,92
VGG-16	80,38 ± 4,29	84,35 ± 9,67	74,67 ± 4,99	75,52 ± 6,11	75,03 ± 5,33
ViT	79,13 ± 6,96	67,45 ± 3,52	73,33 ± 6,67	69,53 ± 5,29	73,67 ± 8,64

### 3.4. Experimento 2 - Validação da Etapa com Pré-processamento

No segundo experimento (Tabela 2), que incorporou a etapa de pré-processamento, a arquitetura EfficientNetV2S apresentou mais uma vez um desempenho notável, superando não apenas as demais arquiteturas avaliadas, mas também os resultados alcançados na ausência desta etapa. Com uma acurácia de 99,23%, precisão de 99%, sensibilidade de 99,5%, *F1-score* de 99,16% e *Kappa* de 99,05%, o EfficientNetV2S se destaca pela capacidade de capturar padrões complexos nos dados após o pré-processamento. A eficácia desta arquitetura pode ser atribuída à sua estrutura eficiente e otimizada, que permite a extração e aprendizagem de características discriminativas de forma mais eficaz. Além disso, o pré-processamento desempenhou um papel fundamental ao fornecer dados mais refinados e adaptados para a arquitetura da rede neural, eliminando ruídos e realçando características relevantes. A combinação da capacidade de representação profunda do EfficientNetV2S com a etapa de pré-processamento contribuiu para um desempenho superior, destacando a importância crítica desta abordagem para aprimorar a capacidade do modelo de compreender e generalizar padrões nos dados complexos do experimento.

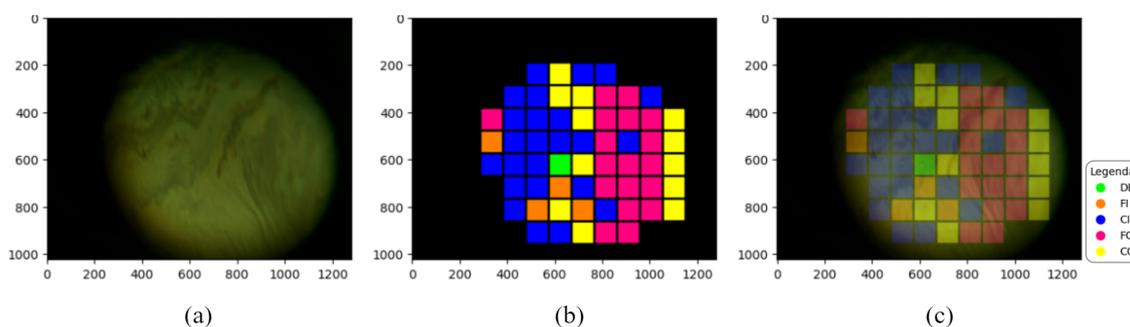
**Tabela 2. Resultados da comparação do método proposto com outras arquiteturas com a etapa de pré-processamento.**

Arquitetura	Acc(%)	Prec(%)	Sen(%)	F1(%)	Kap(%)
DenseNet-121	92,17 ± 1,74	92,71 ± 2,82	90 ± 2,11	90,31 ± 2,35	90,05 ± 2,21
<b>EfficientNetV2S</b>	<b>99,23 ± 1,54</b>	<b>99 ± 2</b>	<b>99,5 ± 1</b>	<b>99,16 ± 1,68</b>	<b>99,05 ± 1,9</b>
ResNet-50	93,91 ± 5,22	94,02 ± 4,91	93,2 ± 5,74	93,07 ± 5,89	92,27 ± 6,62
VGG-16	85,22 ± 4,43	91,73 ± 2,79	82,67 ± 5,33	84,2 ± 5,17	81,28 ± 5,57
ViT	81,3 ± 3,71	79,95 ± 10,59	81,33 ± 6,86	77,03 ± 6,28	76,5 ± 4,7

### 3.5. Estudos de Caso

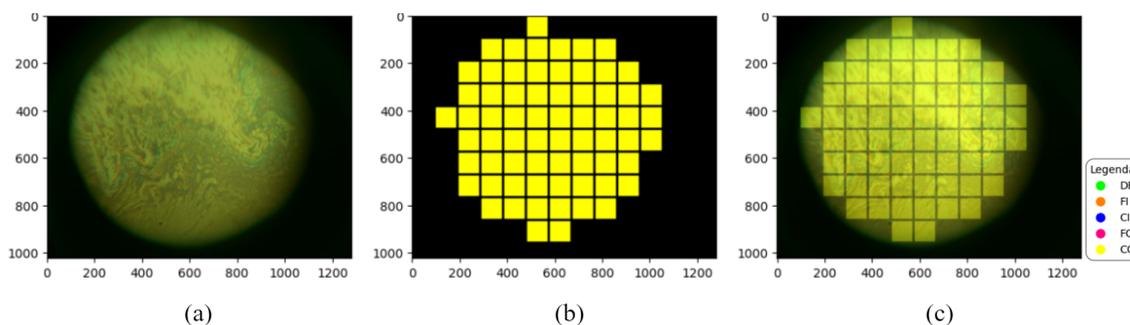
Para avaliar os resultados alcançados no método proposto, foram delineados dois estudos de caso. No primeiro caso, o modelo apresenta algumas deficiências nos resultados de classificação de *patches*. No segundo caso, o modelo classificou eficientemente todos os *patches* da categoria correspondente.

A Figura 7 apresenta o caso que teve o pior resultado do modelo. Nesse caso, o método proposto não foi capaz de definir com maior precisão a categoria da camada lipídica do filme lacrimal. Pode-se notar que foram encontrados predições dos cinco tipos de categorias da camada lipídica do filme lacrimal (falsos positivos). É importante mencionar que a classificação das categorias é uma tarefa clínica difícil, especialmente em camadas lipídicas mais finas que não apresentam características morfológicas e/ou coloridas. A imagem em questão representa a categoria FO (rosa). Porém, a maior frequência das predições foi da categoria CI (azul), classificando-a incorretamente. Este tipo de comportamento foi encontrado em apenas um caso dos 5 *folds* executados.



**Figura 7. Estudo de caso 1. (a) Imagem original; (b) Predição; (c) Predição sobre a imagem.**

O segundo estudo de caso é ilustrado na Figura 8, que apresenta o melhor resultado da classificação do modelo. Neste caso, todos os *patches* foram classificados corretamente na categoria CO (amarelo). Além disso, foram identificados vários casos semelhantes a este comportamento, com poucas predições incorretas. Portanto, esses resultados demonstram que o método proposto é promissor, oferecendo potencial para facilitar a análise dos dados por um especialista, auxiliando em um diagnóstico mais preciso.



**Figura 8. Estudo de caso 2. (a) Imagem original; (b) Predição; (c) Predição sobre a imagem.**

#### 4. Comparação com a Literatura

Ao analisar os trabalhos relacionados (todos apresentam a mesma base VOPTICAL\_GCU) para o diagnóstico da síndrome do olho seco, observa-se uma variedade de abordagens computacionais. Inicialmente, [G. Villaverde et al. 2014] empregaram

técnicas clássicas, como filtros de *Butterworth*, *Gabor*, Transformada Discreta de *Wavelet*, campos aleatórios de Markov, e características de matrizes de coocorrência, explorando escala de cinza e o espaço de cor  $L^*a^*b^*$ . Enquanto alcançaram uma acurácia de 91,51% utilizando SVM, a pesquisa subsequente de [Remeseiro et al. 2015] incorporou cores oponentes do espaço RGB, alcançando 93,4% com algoritmos como *Naive Bayes*, *Random Tree*, *Random Forest* e SVM. Além disso, o estudo de [Cruz et al. 2018a] destacou a incorporação de informações de cor e textura, empregando algoritmos de *machine learning*, resultando em uma acurácia de 95,66% com o *Naive Bayes*.

**Tabela 3. Comparação do método proposto com os trabalhos relacionados.**

Técnica(s)	Acc(%)	Prec(%)	Sens(%)	F1(%)	Kap(%)
Processamento de sinais, modelo e estatístico, seleção de características CFS, consistência, INTERACT e SVM [G. Villaverde et al. 2014]	91,51	-	-	-	-
Processamento de sinais, modelo, estatístico e SVM [Remeseiro et al. 2015]	93,4	-	-	93,13	-
Função K de Ripley, seleção de características <i>Greedy Stepwise</i> e <i>Random Forest</i> [Cruz et al. 2018a]	95,66	-	-	95	93
Índices de diversidade filogenética baseados no caminho mínimo e <i>Random Forest</i> [Cruz et al. 2018b]	96,98	-	-	97	96
Índices de diversidade filogenética baseados em caminhos mínimos e na distância entre pares de espécies e <i>Random Forest</i> [da Cruz et al. 2020]	97,54	-	-	97	96
K de Ripley e diversos índices de diversidade filogenética [da Cruz et al. 2020]	99,62	-	-	99	99
<b>Método proposto com EfficientNetV2S</b>	<b>99,23</b>	<b>99</b>	<b>99,5</b>	<b>99,16</b>	<b>99,05</b>

Por outro lado, os métodos propostos por [Cruz et al. 2018b, da Cruz et al. 2020] inovaram ao explorar índices de diversidade filogenética para extrair características de textura, obtendo acurácias de 96,98% e 97,54%, respectivamente, ambos utilizando *Random Forest*. Notavelmente, o método mais recente de [da Cruz et al. 2020] alcançou impressionantes 99,62% de acurácia, incorporando K de Ripley e índices de diversidade nos espaços de cor HLS e escala de cinza, além de empregar o *Greedy Stepwise* para seleção de características. Em contraste, o presente trabalho propõe uma nova abordagem, integrando técnicas automatizadas, como aprendizado não supervisionado para segmentação da ROI e *deep learning* para classificação. A ausência de uma etapa explícita de extração e seleção de características destaca a robustez e adaptabilidade do método, superando limitações comuns em problemas complexos. Vale destacar também o resultado do *F1-score*, uma medida harmônica entre sensibilidade e precisão, que atingiu 99,16%, demonstrando a eficácia do modelo mesmo em cenários desafiadores. Assim, este trabalho apresenta um avanço na área, fazendo uso de progressos recentes em *machine learning* para aprimorar a precisão na classificação da camada lipídica do filme lacrimal.

## 5. Conclusão

O reconhecimento precoce e a compreensão aprofundada dos sintomas e fatores de risco são fundamentais para um diagnóstico preciso e uma gestão eficaz da síndrome do olho seco. Os resultados deste estudo oferecem um avanço substancial na área, evidenciados pelos 99,23% de acurácia, 99% de precisão, 99,5% de sensibilidade, 99,16% de *F1-score* e 99,05% de *Kappa* alcançados. Essa precisão abre novas perspectivas para aprimorar a

detecção precoce e a gestão clínica da síndrome do olho seco, reforçando a importância da análise da camada lipídica como um componente vital na compreensão abrangente e no cuidado integral da saúde ocular. Para pesquisas futuras, sugere-se explorar a integração de novas bases de imagens. Além de outras técnicas de *machine learning*.

## Referências

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., et al. (2015). Tensorflow: Large-scale machine learning on heterogeneous systems. *Software available from tensorflow.org*.
- Burman, P. (1989). A comparative study of ordinary cross-validation, v-fold cross-validation and the repeated learning-testing methods. *Biometrika*, 76(3):503–514.
- Chollet, F. et al. (2015). Keras. <https://keras.io>.
- Cruz, L. B., Araújo, J. D. L., Sousa, J. A., Almeida, J. D. S., Júnior, G. B., Silva, A. C., and Paiva, A. C. (2018a). Classificação do filme lacrimal usando a função k de ripley como descritor de textura. In *Anais do XVIII Simpósio Brasileiro de Computação Aplicada à Saúde*, Porto Alegre, RS, Brasil. SBC.
- Cruz, L. B. d., Araújo, J. D. L., Souza, J. C., de Sousa, J. A., de Almeida, J. D. S., Junior, G. B., Silva, A. C., and de Paiva, A. C. (2018b). Tear film classification using phylogenetic diversity indexes as texture descriptor. In *2018 IEEE Symposium on Computers and Communications (ISCC)*, pages 00853–00858.
- da Cruz, L. B., Souza, J. C., de Paiva, A. C., de Almeida, J. D. S., Junior, G. B., Aires, K. R. T., Silva, A. C., and Gattass, M. (2020). Tear film classification in interferometry eye images using phylogenetic diversity indexes and ripley's k function. *IEEE Journal of Biomedical and Health Informatics*, 24(12):3491–3498.
- da Cruz, L. B., Souza, J. C., de Sousa, J. A., Santos, A. M., de Paiva, A. C., de Almeida, J. D. S., Silva, A. C., Junior, G. B., and Gattass, M. (2020). Interferometer eye image classification for dry eye categorization using phylogenetic diversity indexes for texture analysis. *Computer Methods and Programs in Biomedicine*, 188:105269.
- Davidson, I. (2002). Understanding k-means non-hierarchical clustering. *Computer Science Department of State University of New York (SUNY), Albany*.
- Doane, M. G. (1989). An instrument for in vivo tear film interferometry. *Optometry and Vision Science*, 66(6):383–388.
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., and Hounsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *CoRR*, abs/2010.11929.
- G. Villaverde, D., Remeseiro, B., Barreira, N., G. Penedo, M., and Mosquera, A. (2014). Feature selection applied to human tear film classification. In *Proceedings of the 6th International Conference on Agents and Artificial Intelligence - Volume 1, ICAART 2014*, pages 395–402, Portugal. SCITEPRESS - Science and Technology Publications, Lda.
- Gedraite, E. S. and Hadad, M. (2011). Investigation on the effect of a gaussian blur in image filtering and segmentation. In *Proceedings ELMAR-2011*, pages 393–396.

- Huang, G., Liu, Z., and Weinberger, K. Q. (2016). Densely connected convolutional networks. *CoRR*, abs/1608.06993.
- Júnior, D. D., Cruz, L., Diniz, J., Júnior, G. B., and Silva, A. (2021). Classificação automática de glóbulos brancos usando descritores de forma e textura e extreme gradient boosting. In *Anais do XXI Simpósio Brasileiro de Computação Aplicada à Saúde*, pages 95–106, Porto Alegre, RS, Brasil. SBC.
- Koonce, B. and Koonce, B. (2021). Resnet 50. *Convolutional Neural Networks with Swift for Tensorflow: Image Recognition and Dataset Categorization*, pages 63–72.
- Leite, F., Almeida, J., Teixeira, J., Cruz, L., Junior, G., and Paiva, A. (2020). Planejamento cirúrgico de estrabismo horizontal utilizando Árvore de regressão de múltiplas saídas. In *Anais do XX Simpósio Brasileiro de Computação Aplicada à Saúde*, pages 356–367, Porto Alegre, RS, Brasil. SBC.
- Marculino, L. G. C., Hazarbassanov, R. M., Hazarbassanov, N. G. T. d. Q., Hirai, F., Milhomens Filho, J. A. P., Wakamatsu, T. H., and Gomes, J. A. (2022). Prevalence and risk factors for dry eye disease: the sao paulo dry eye study. *Arquivos Brasileiros de Oftalmologia*, 85:549–557.
- Messmer, E. M. (2015). The Pathophysiology, Diagnosis, and Treatment of Dry Eye Disease. *Dtsch Arztebl International*, 112(5):71–82.
- Panetta, K., Zhou, Y., Agaian, S., and Jia, H. (2011). Nonlinear unsharp masking for mammogram enhancement. *IEEE Transactions on Information Technology in Biomedicine*, 15(6):918–928.
- Pereira, R., Matos, C., Diniz, J., Junior, G. B., de Almeida, J., Silva, A., and de Paiva, A. (2016). Abordagem deep learning para classificação de lesões mamárias. In *Anais do XVI Workshop de Informática Médica*, pages 2597–2600, Porto Alegre, RS, Brasil. SBC.
- Quintana, G. I., Li, Z., Vancamberg, L., Mougeot, M., Desolneux, A., and Muller, S. (2023). Exploiting patch sizes and resolutions for multi-scale deep learning in mammogram image classification. *Bioengineering*, 10(5):534.
- Remeseiro, B., Oliver, K. M., Tomlinson, A., Martin, E., Barreira, N., and Mosquera, A. (2015). Automatic grading system for human tear films. *Pattern Analysis and Applications*, 18(3):677–694.
- Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Stapleton, F., Alves, M., Bunya, V. Y., Jalbert, I., Lekhanont, K., Malet, F., Na, K.-S., Schaumberg, D., Uchino, M., Vehof, J., Viso, E., Vitale, S., and Jones, L. (2017). Tfos de ws ii epidemiology report. *The Ocular Surface*, 15(3):334 – 365. TFOS International Dry Eye WorkShop (DEWS II).
- Tan, M. and Le, Q. V. (2021). Efficientnetv2: Smaller models and faster training. *CoRR*, abs/2104.00298.
- VOPTICAL\_GCU, V. (2017). optical dataset acquired and annotated by optometrists from the department of life sciences, glasgow caledonian university (uk), 2013. [http://www.varpa.org/voptical\\_gcu.html](http://www.varpa.org/voptical_gcu.html).