

Dissertação de Mestrado

Análise Computacional da Biomecânica Corneal para Diagnóstico de Ceratocone

Kempes Jacinto

malkavk@gmail.com

Orientador:

Dr. Aydano Pamponet Machado

Maceió, Março de 2018

Kempes Jacinto

Análise Computacional da Biomecânica Corneal para Diagnóstico de Ceratocone

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-Graduação em Modelagem Computacional de Conhecimento do Instituto de Computação da Universidade Federal de Alagoas.

Orientador: Dr. Aydano Pamponet Machado

Maceió, Março de 2018

Catálogo na fonte
Universidade Federal de Alagoas
Biblioteca Central

Bibliotecária Responsável: Janis Christine Angelina Cavalcante

- J12a Jacinto, Kempes.
 Análise computacional da biodinâmica corneal para diagnóstico de ceratocone /
 Kempes Jacinto. – 2018.
 80 : il., color., graf., tabs.
- Orientador: Aydano Pomponet Machado.
 Coorientador: João Marcelo de Almeida Gusmão Lyra.
 Dissertação (Mestrado em Modelagem Computacional de Conhecimento) –
 Universidade Federal de Alagoas. Instituto de Computação. Programa de
 Pós-Graduação em Modelagem Computacional de Conhecimento. Maceió, 2018.
- Bibliografia: f. 74-80.
1. Ceratonone. 2. Biomecânica. 3. Diagnóstico. 4. Inteligência artificial.
 5. Segmentação. I. Título.

CDU: 004.932



Membros da Comissão Julgadora da Dissertação de Mestrado de Kempes Jacinto, intitulada: “Análise Computacional da Biomecânica Corneal para Diagnóstico de Ceratocone”, apresentada ao Programa de Pós-Graduação em Modelagem Computacional de Conhecimento da Universidade Federal de Alagoas, em 13 de abril de 2018, às 14h00min, na sala de reuniões do Instituto de Computação da Ufal.

COMISSÃO JULGADORA

Prof. Dr. Aydano Pamponet Machado

Ufal – Instituto de Computação

Orientador

Prof. Dr. João Marcelo de Almeida Gusmão Lyra

Uncisal – Escola de Ciências Médicas

Orientador

Prof. Dr. Jorge Artur Peçanha de Miranda Coelho

Ufal – Faculdade de Medicina

Examinador

Prof. Dr. Renato Ambrósio Júnior

Unifesp – Departamento de Oftalmologia

Examinador

Maceió, abril de 2018.

Dedico este trabalho

àquela que tem sido meu complemento, minha companheira,
apoiadora, amiga, inspiradora, que foi tão louca quanto eu
(se arriscando em horas de estrada, de dia e de noite, entre o conforto de nosso lar e nossos
locais de estudos),
a toda a energia, confiança e incentivo que me deu,
minha querida esposa.

RESUMO

O objetivo do corrente estudo foi encontrar e modelar representações de características da biomecânica corneal a partir de imagens de exames geradas pelo Corvis ST, a fim de realizar sua aplicação a técnicas de aprendizagem de máquina para o diagnóstico precoce de ceratocone. As imagens foram segmentadas para identificação e conversão em vetores para representação das superfícies anterior, superfície posterior aparente, paquimetria aparente e composição dos dados anteriores. Os vetores foram encadeados (imagens em lote), simplificados com Wavelet e submetidos a MLP, k -NN, Regressão Logística, Naïve Bayes e *Fast Large Margin*, além do arranjo dos vetores como histogramas 2D para aplicação em rede neural com *Deep Learning*. A avaliação das classificações foi feita com o escore igual ao produto da sensibilidade multiplicado pela especificidade, com intervalo de confiança entre 0,7843 e 1 e nível de significância 0,0157. Foram usados exames de 686 olhos normais e 406 olhos com ceratocone em graus de I a IV, provindos de bases de exames da Europa e do Brasil, para treinamento e validação dos dados aplicados. Os melhores modelos identificados ocorreram com paquimetria aparente de imagens em lote, com aplicação de wavelet nível 4 e processada com *fast large margin* na base de dados da Europa, com escore 0,8247, sensibilidade de 89,5% e especificidade de 92,14%; e histograma 2D da paquimetria aparente, com LeNET5, na base do Brasil, com escore 0,8361, sensibilidade de 88,58% e especificidade de 94,39%. Conclui-se que os modelos da biomecânica podem ser usados para diagnosticar ceratocone.

Palavras-chave: Ceratocone. Biomecânica. Diagnóstico. Inteligência Artificial. Segmentação.

ABSTRACT

The objective of the present study was to find and model representations of corneal biomechanical characteristics from images of exams generated by Corvis ST in order to apply them to machine learning techniques for the early diagnosis of keratoconus. The images were segmented for identification and conversion into vectors for representation of the anterior surface, apparent posterior surfaces, apparent pachymetry and composition of the previous data. The vectors were chained (batch images), simplified with Wavelet and submitted to MLP, k -NN, Logistic Regression, Naïve Bayes and Fast Large Margin, in addition the vectors were rearranged as 2D histograms for neural network application with Deep Learning. The evaluation of the classifications was done with the score equal to the product of the sensitivity multiplied by the specificity, with confidence interval between 0.7843 and 1 and level of significance 0.0157. Exams of 686 normal eyes and 406 eyes with keratoconus in degrees from I to IV, from exam bases from Europe and Brazil, were used for training and validation of applied data. The best models identified were apparent pachymetry on batch images, with wavelet level 4 and processed with fast large margin in the european database, with a score of 0.8247, sensitivity of 89.5% and specificity of 92.14%; and 2D histogram of apparent pachymetry, with LeNET5, at the brazilian database, with a score of 0.8361, sensitivity of 88.58% and specificity of 94.39%. It is concluded that biomechanical models can be used to diagnose keratoconus.

Keywords: Keratoconus. Corneal Biomechanics. Diagnosis. Artificial Intelligence. Segmentation.

AGRADECIMENTOS

Agradeço às minhas duas famílias – meus pais e meus sogros – que nunca me deixaram desistir dessa longa, difícil e tortuosa estrada chamada formação.

A cada um que fez parte da minha turma nesse programa de pós-graduação. Aos que tiveram que deixar o curso (que me lembraram que apesar das minhas dificuldades, eu ainda estava em condições privilegiadas); aos que terminaram antes (que me mostraram que eu poderia me empenhar mais); para os que terminarão depois (que a força esteja com vocês). Em especial para turma de Probabilidade e Inferência de 2015.1 (*no pain, no gain*). E mais especial a Ana Maria (a consultora matemática, com pensamento e fala mais rápidos que eu podia acompanhar).

Ao senhor Vitor Torres, sempre prestativo a informações, incentivando a término do curso e com conversas (nerds) para relaxamento do espírito.

A todos vocês sou eternamente grato.

Kempes Jacinto

LISTA DE FIGURAS

| | | |
|------|---|----|
| 2.1 | Sequência de aplicação de filtros para segmentação de imagens: original, mediana 3D, desfoque gaussiano, limiar adaptativo, dilatação e erosão. | 12 |
| 2.2 | Mediana 3D aplicada à sequência de quadros em um vídeo | 14 |
| 2.3 | Representação de neurônio | 19 |
| 2.4 | Rede Neural Artificial | 20 |
| 2.5 | Neurônio Artificial | 21 |
| 2.6 | Decomposição multinível de onda | 32 |
| 2.7 | Wavelet Haar | 32 |
| 2.8 | Família de wavelets Coiflet | 33 |
| 2.9 | Representação da Anatomia Corneal. | 34 |
| 2.10 | Perfil de Córneas: normal (imagem à esquerda) e com ceratocone em estágio avançado (imagem à direita). | 35 |
| 2.11 | Córnea com implante de Anel Intraestromal Corneano em dois segmentos. | 38 |
| 2.12 | Dinâmica da córnea durante exame com Corvis ST | 40 |
| 2.13 | Variação de posicionamento do olho em diferentes exames do Corvis | 42 |
| 2.14 | Áreas anômalas capturadas em diferentes exames do Corvis | 43 |
| 3.1 | Identificação de marcadores no exame | 52 |
| 3.2 | Distribuição de valores largura de região de deflexão de córnea das bases de exames do Corvis ST | 54 |
| 3.3 | Diferentes exames do Corvis, onde a córnea não apresenta continuidade. | 56 |
| 3.4 | Histogramas 2D sobre caracterização do movimento da face corneal anterior para avaliação de procedimentos de diagnóstico com DL. | 58 |
| 3.5 | Histogramas 2D para análise da paquimetria corneal aparente para avaliação de procedimentos de diagnóstico com DL. | 59 |
| 3.6 | Sinal de entrada original (primeira imagem), aplicação de coif6 níveis 4 (segunda imagem) e 7 (terceira imagem) | 61 |
| 4.1 | Histogramas 2D do mesmo exame e com técnicas diferentes de remoção de movimento do olho. | 64 |

LISTA DE TABELAS

| | | |
|-----|--|----|
| 2.1 | Classificação de ceratocone, segundo Krumeich | 36 |
| 3.1 | Exemplo de matriz de confusão | 47 |
| 4.1 | Resultado da segmentação das imagens, por grupo de dados (usando a classificação Krumeich) | 63 |
| 4.2 | Diagnóstico resultante de processamento de histogramas 2D | 64 |
| 4.3 | Aplicação da rede DL treinada com base do Brasil na base da Europa | 65 |
| 4.4 | Escores de modelos preditivos com treinamento na base de exames do Brasil | 67 |
| 4.5 | Escores de modelos preditivos com treinamento na base de exames da Europa | 67 |
| A.1 | Resultado de treinamento da base de exames do Brasil com aprendizagem de máquina sobre wavelet | 73 |
| A.2 | Resultado de treinamento da base de exames da Europa com aprendizagem de máquina sobre wavelet | 73 |

SUMÁRIO

| | | |
|----------|---|-----------|
| 1 | INTRODUÇÃO | 9 |
| 1.1 | Definição do Problema | 9 |
| 1.2 | Contribuições | 10 |
| 1.3 | Objetivos | 10 |
| 1.4 | Organização da Dissertação | 10 |
| 2 | FUNDAMENTAÇÃO TEÓRICA | 11 |
| 2.1 | Segmentação de Imagem | 11 |
| 2.1.1 | Desfoque | 12 |
| 2.1.2 | Binarização | 14 |
| 2.1.3 | Dilatação e Erosão | 14 |
| 2.2 | OpenCV | 15 |
| 2.3 | Aprendizagem de Máquina | 15 |
| 2.3.1 | Regressão Linear e Regressão Logística | 17 |
| 2.3.2 | Redes Neurais Artificiais | 19 |
| 2.3.3 | SVM e <i>Fast Large Margin</i> | 23 |
| 2.3.4 | <i>k</i> -NN | 25 |
| 2.3.5 | Naïve Bayes | 26 |
| 2.3.6 | <i>Deep Learning</i> | 28 |
| 2.4 | Wavelet | 30 |
| 2.5 | Biomecânica | 31 |
| 2.6 | Anatomia Corneal | 33 |
| 2.7 | Ceratocone | 35 |
| 2.8 | Exames para Análise de Biomecânica Corneana | 38 |
| 2.9 | Exames de Tomógrafos e seus Dados | 41 |
| 3 | METODOLOGIA PARA DIAGNÓSTICO DE CERATOCONE | 46 |
| 3.1 | Crítérios de Avaliação de Modelos de Aprendizagem de Máquina | 46 |
| 3.2 | Descrição dos Dados de Exames | 47 |
| 3.3 | Metodologias de Processamento | 48 |
| 3.3.1 | Processamento das Imagens | 49 |
| 3.3.2 | Pré-Processamento dos Dados | 51 |
| 3.3.3 | Processamento dos Dados: Aprendizado de Máquina | 57 |
| 4 | RESULTADOS E DISCUSSÃO | 62 |
| 4.1 | Bases de Exames | 62 |
| 4.2 | Segmentação e Processamento das Imagens | 62 |
| 4.3 | Aprendizado com <i>Deep Learning</i> | 63 |
| 4.4 | Aprendizado sobre Wavelets | 66 |
| 5 | CONCLUSÕES | 70 |
| 5.1 | Considerações Finais | 70 |
| 5.2 | Trabalhos futuros | 71 |

| | | |
|----------|---|-----------|
| A | TABULAÇÃO DE EXECUÇÃO DE TÉCNICAS DE APRENDIZAGEM DE MÁQUINA SOBRE DADOS DE EXAMES | 72 |
| | Referências Bibliográficas | 74 |

1

INTRODUÇÃO

1.1 Definição do Problema

UMA lacuna de conhecimento sobre a forma de funcionamento de um órgão, ou de um pedaço dele, pode trazer à tona problemas ao lidar com doenças nesse órgão. Mesmo que uma doença seja há muito conhecida, isso não significa que a mesma seja completamente compreendida. Nessas perspectivas, encontram-se a córnea e o ceratocone, que são respectivamente uma parte extremamente delgada do olho e uma doença que ainda tem suas causas como foco de estudos. A estrutura da córnea é bem conhecida, mas como diagnosticar precisamente uma desordem que, em alguns casos, pode ter avanço muito lento, ao longo de anos, e em outros, muito rápido?

Para auxiliar nesse trabalho, a medicina tem recorrido a recursos e conhecimentos provenientes de outras áreas de conhecimento, a fim de fazer com que novas luzes incidam sobre conhecimentos ainda obscuros. A adoção de procedimentos de análise com princípios e efeitos físicos é um exemplo disso. É comum nos cursos de medicina o estudo de órgãos *ex vivo*, de forma estática. Mas *in vivo* os órgãos movem e são movidos, sofrem e aplicam forças, tensionam e são tensionados. Tais conceitos da física têm sido estudados na medicina sob o nome de biomecânica.

Um passo nesse sentido é o uso de novas tecnologias para execução de exames. Por exemplo, o uso de *lasers* e câmeras de grande precisão para captação de imagens da córnea em exames não invasivos, mostrando o funcionamento do órgão de forma dinâmica e em funcionamento.

Algumas empresas já estão fornecendo equipamentos com essas capacidades. Equipamentos que geram muitos dados e observações, os quais, por serem de uma área de estudo muito nova, não têm seus valores ainda totalmente vinculados a significados e conhecimentos. A análise e experimentação desses dados, fazendo o vínculo com conhecimentos já consolidados, é uma necessidade tão grande quanto a própria descoberta de uma nova forma de exame

e pode ser proporcionada pela área de Modelagem Computacional de Conhecimento. Isso pode ser feito, por exemplo pela extração de informações de exames e prontuários médicos através de perspectivas de análise diferentes ao seu meio habitual.

1.2 Contribuições

O atual estudo pesquisou – com base em exames, gerados por equipamentos não invasivos e com auxílio de técnicas de aprendizado de máquina – a biomecânica da córnea a fim de realizar o diagnóstico de ceratocone, usando para isso características extraídas de imagens.

Nessa pesquisa foram analisados parâmetros de descrição da córnea e de seu movimento ainda não abordados em outros estudos, possibilitando novas formas de análise do ceratocone.

1.3 Objetivos

Os principais objetivos relacionados ao trabalho são:

- Encontrar características da biomecânica corneal em sequências de imagens de exames não invasivos;
- Identificar uma forma de representação dessas características que sejam aplicáveis a aprendizagem de máquina;
- Aplicar técnicas de aprendizagem de máquina que permitam avaliação das características encontradas e forneçam diagnóstico quanto a ocorrência de ceratocone;

1.4 Organização da Dissertação

Cada um dos próximos capítulos abordará uma fase da pesquisa: no capítulo 2, há o levantamento da fundamentação teórica, trazendo a caracterização da córnea e dos conhecimentos da área de ciências da computações que auxiliarão nas fases seguintes. No capítulo 3, há o detalhamento da metodologia de pesquisa, com descrição dos passos que levaram ao desenvolvimento dos experimentos e sua execução. No capítulo 4, são apresentados os resultados e sua discussão. No capítulo 5 descreveremos a conclusão e sugestões para trabalhos futuros.

O capítulo abordou o problema alvo dessa dissertação, apresentando as contribuições e objetivos da pesquisa. No próximo capítulo, será apresentada a fundamentação teórica em conjunto com justificativas de algumas de suas escolhas.

2

FUNDAMENTAÇÃO TEÓRICA

ESTÃO descritos, neste capítulo, a apresentação de conceitos e técnicas para análises quantitativas e uso de aprendizado de máquina. Em seguida, o conceito de biomecânica, a anatomia corneal e propriedades biomecânicas, o ceratocone e exames para quantificar e avaliar esta doença.

2.1 Segmentação de Imagem

Segmentação de imagem é o ato de transformar uma imagem em outra, normalmente com o objetivo de extrair uma estrutura que, de alguma forma, possa ser entendida pelo computador, o qual dá significância a seu conteúdo. Por ser análogo ao funcionamento da visão humana, isso também pode ser chamado de Visão Computacional (ou em inglês *Computer Vision*, ou ainda CV).

Sistemas de visão de máquina são normalmente considerados ser compostos de dois subsistemas: visão de baixo nível e visão de alto nível. Visão de Baixo Nível consiste, primeiramente, de operações de processamento de imagem executadas na imagem de entrada para produzir outra imagem com características mais favoráveis. Essas operações podem envolver imagens com reduzido ruído ou causar a ênfase em certas características (como bordas). Visão de Alto Nível inclui reconhecimento de objetos e, em um nível mais alto, interpretação de cenas. A ponte entre esses dois subsistemas é a segmentação de imagem. (Spirkovska, 1993, p. 3, tradução nossa).

Tanto os sistemas de visão, quanto a segmentação, como apresentados por Spirkovska, não são teorias universalmente válidas, mas sim estudos sobre as imagens a serem processadas e a seleção de técnicas que possam realizar a tarefa desejada. Tais técnicas (também

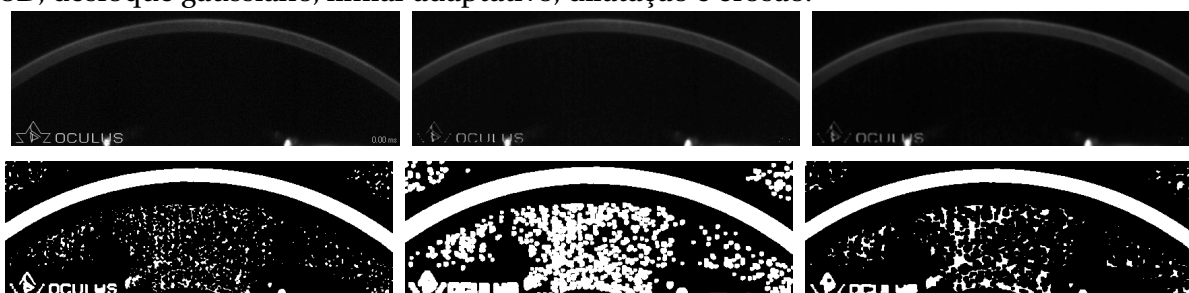
conhecidas como filtros) possuem teorias que as embasam e as tornam válidas. A variedade e a quantidade de filtros a serem aplicados durante um processo de segmentação envolvem algum conhecimento sobre as imagens a serem processadas e os resultados desejados, sendo comum seu “empilhamento”, isto é, o resultado de um filtro é a entrada para um outro (Spirkovska, 1993). A Equação 2.1, demonstra um típico uso de filtros de processamento de imagens.

$$F(i) = f_n \circ f_{n-1} \circ \dots \circ f_m \circ f_l \dots \circ f_1(i) \quad (2.1)$$

onde $F(i)$ é o resultado da pilha de filtros sobre a imagem inicial i , cada componente f é um filtro a ser aplicado sobre o resultado do filtro anterior (recursivamente até o primeiro), o filtro de índice m pode ser o mesmo que o de índice l e a sequência não é obrigatoriamente comutativa.

Mesmo tendo um objetivo em comum, os filtros têm metodologias diferentes e, possivelmente, finalidades distintas. A Figura 2.1 exemplifica o uso de uma sequência de cinco filtros sobre uma imagem original.

Figura 2.1: Sequência de aplicação de filtros para segmentação de imagens: original, mediana 3D, desfoque gaussiano, limiar adaptativo, dilatação e erosão.



Fonte: autor

Nas seções a seguir, são apresentados alguns modelos de filtros, os quais foram usados na corrente pesquisa.

2.1.1 Desfoque

Este não é um único filtro, mas um conjunto de técnicas de desfoque que causam o esmaecimento de imagens. Quando aplicado a imagens, este tipo de filtro causa a perda de acuidade de linhas e pontos. Essa mesma perda pode auxiliar na remoção de ruído de imagens. Um tipo de desfoque bastante utilizado é o Gaussiano. Esse filtro calcula a média da cor para todos os valores dentro de um raio definido, através da convolução da imagem com o kernel Gaussiano especificado (Gomes et al., 2017; OpenCV Development, 2017).

Em processamentos de imagem em que o alvo é encontrar as bordas de objetos, uma alternativa é o uso do tipo *Rolling Guidance*. Esse filtro é caracterizado pela perda de detalhes

da imagem, pela homogeneização de grandes áreas e com definição de bordas. Isso é feito pela execução iterativa de desfoques Gaussianos, com cálculos de cores definidos por médias entre valores de pixels com limiares pré-definidos. Durante a execução desse processamento, a imagem original não é perdida, pois seus valores de cores são usados como guia durante as iterações. Os desfoques ocorrem com o uso do desfoque gaussiano da iteração anterior e as cores são corrigidas para a média das cores dos pixels, também da iteração anterior. Caso a diferença entre a cor resultante e a cor do pixel na mesma posição na imagem original não ultrapasse o limiar definido na configuração do filtro, nova iteração pode ser executada sobre a região. As iterações possuem dois critérios de parada: a não modificação de nenhuma região de uma iteração para outra (todas as regiões já alcançaram a máxima divisão) ou a configuração possuir definição da quantidade máxima de iterações (Zhang et al., 2014).

A definição do filtro pode ser encontrada em Zhang et al. (2014), com implementação em OpenCV Development (2017).

Uma outra alternativa de desfoque é a mediana n -dimensional. Um filtro de mediana de uma dimensão (1D) e tamanho l é aquele que, quando aplicado a um sinal de entrada (I) de tamanho k , gera um sinal de saída (O) de tamanho k , e cujo valor da x -ésima posição de O é dado pela mediana dos valores de I compreendidos no intervalo $(x - l, x + l)$, caso existam.

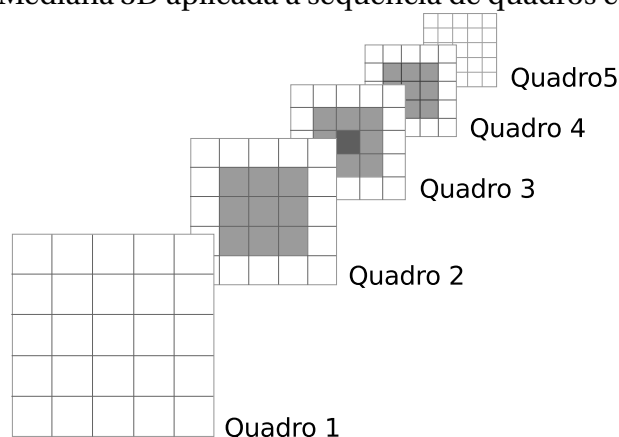
De forma análoga, pode-se aplicar esse filtro em duas dimensões (2D), havendo a possibilidade de delimitar a região (também chamada de *janela*) de abrangência dos cálculos de mediana (OpenCV Development, 2017; Demassieux et al., 1985). As regiões de delimitação são, normalmente, figuras geométricas centradas no valor localizado na posição (x, y) . São exemplos de formas desse filtro em 2D (Demassieux et al., 1985):

Quadrado os valores usados para cálculo de mediana estão compreendidos entre as colunas $x - l$ e $x + l$ e as linhas $y - l$ e $y + l$

Circular todos os valores localizados dentro do círculo de centro (x, y) e raio l

Medianas podem, ainda, ser aplicadas em três dimensões (3D), adquirindo formas análogas às 2D (como Cubo e Esfera), podendo ser aplicadas sobre imagens tridimensionais ou em bidimensionais com distribuição temporal, como quadros em um vídeos. Nesse último caso, é equivalente a montar uma imagem 3D pela justaposição nas imagens 2D, ordenadas pela sequência de exibição (Matyunin et al., 2011). Isso pode ser visto na Figura 2.2, onde estão demarcados o pixel que está sendo calculada a mediana (quadrado cinza escuro no Quadro 3), os pixels a serem usados para calcular a mediana (todos os quadrados cinza dos Quadros 2, 3 e 4) e os pixels que não vão ser usados no corrente cálculo (quadrados brancos em todos os Quadros), representando, assim, um filtro de mediana 3D, tamanho 1, forma cúbica aplicado a uma sequência de 5 imagens.

Figura 2.2: Mediana 3D aplicada à sequência de quadros em um vídeo



Fonte: O autor.

2.1.2 Binarização

Binarização é a conversão de uma imagem ou canal - componentes de uma imagem sobre determinada caracterização, como tipos de cores (como os padrões RGB ou CMYK) ou valores de pixel (como o padrão HSV) - em um conjunto de pixel binários, em que os pixels que atendem às condições configuradas são modificados para branco e todos os outros são definidos para preto. Este tipo de filtro é particularmente útil quando se necessita identificar bordas de objetos em imagens. Existem métodos, como o limiar, em que basta que se defina uma parametrização (luminância máxima, por exemplo) e toda a imagem é modificada (Gomes et al., 2017).

Com limiar adaptativo, por exemplo, a imagem é analisada de forma mais minuciosa e a binarização ocorre por avaliações locais. A imagem é subdividida e os limiares são calculados nas subregiões. É um método que localiza regiões homogêneas (OpenCV Development, 2017).

Em Canny (1986), foi feita definição do filtro que leva seu nome, Canny. Ele é um dos mais usados para identificação de contornos de objetos dentro de imagens 2D. Esse processamento ocorre pela convolução da imagem, com sua versão após a aplicação de um desfoque Gaussiano. Como resultado, as regiões de bordas são destacadas, enquanto que as demais são removidas. Assim, tem-se uma imagem binarizada aparecendo apenas os contornos dos objetos.

2.1.3 Dilatação e Erosão

O filtro de dilatação normalmente é usado para enfatizar regiões já destacadas na imagem, pelo aumento da área dessas regiões. Isso pode causar o desaparecimento de ruído, quando este possui valor tonal inferior ao da área em dilatação (Gomes et al., 2017).

Em imagens binarizadas, por terem regiões distintas em claras e escuras, a dilatação pode causar a união de áreas adjacentes, mas que não se tocavam anteriormente, sendo útil para o

aumento das áreas de bordas com cobertura de imperfeições ou união com áreas espúrias provindas de ruído (Gomes et al., 2017).

Sua implementação é feita pela definição do valor de um pixel como sendo o máximo entre seu valor e os pixels a seu redor. A região de pixels a ser avaliada é equivalente àquela aplicada ao filtro de mediana (Gomes et al., 2017).

Já a erosão é a operação inversa da dilatação. As áreas em destaque diminuem de tamanho, por causa dos valores dos pixels a serem definidos como o mínimo valor entre todos na área afetada ao seu redor (Gomes et al., 2017). Nesse caso, o fundo da imagem (*background*) aumenta sua área sobre as regiões em destaque, podendo ser usado para remoção de ruído sobre as bordas de objetos em primeiro plano ou (em inglês) *foreground* (Gomes et al., 2017).

2.2 OpenCV

OpenCV é um conjunto de bibliotecas de programação para Visão Computacional, a qual abrange centenas de filtros para imagens com finalidades diversas, possuindo implementação em linguagens de programação como C, C++ e Python e disponível para diversos sistemas operacionais de computadores, celulares e tablets. Possui variações para execução em *CPU* (do inglês *Central Processing Unit*, ou em português, Unidade Central de Processamento) e *GPU* (do inglês *Graphical Processing Unit*, ou em português, Unidade de Processamento Gráfico), muitas vezes mais rápida (OpenCV Development, 2017).

Quando em uso, pode ser aplicada em conjunto com outras bibliotecas para processamentos mais complexos e para auxiliar a aceleração de processamentos pela adaptação de uso de conjuntos de valores quaisquer como se fossem imagens. Além dos filtros, como os presentes nas seções anteriores, esse conjunto de bibliotecas possui funções para auxiliar na identificação de objetos em imagens, como obtenção automática de bordas em modelagens vetoriais (mais próprios para compreensão por computadores), além de funcionalidades para exibição, abertura, exportação e edição de imagens (OpenCV Development, 2017).

2.3 Aprendizagem de Máquina

Não existe um conceito único de aprendizagem ou de aprendizado. De forma muito generalista, a plasticidade do cérebro humano permite seu uso para captar experiências do meio natural ou social, extrair diversos tipos de informações e subjetivar conhecimento para aplicação nas próximas experiências (DeVries, 2000).

Aprendizado de Máquina (AM ou, em inglês, *Machine Learning* ou ainda ML) é uma subárea da inteligência artificial que vem a estudar formas de fornecer a um equipamento (computador) uma plasticidade funcional (algumas vezes parecida como a humana e em outras, não), permitindo que, dentro das limitações do *hardware* e do *software*, o equipamento

possa adquirir e usar conhecimento (Haykin, 2009; Valença, 2010).

Em contexto mais restrito:

Aprendizagem de máquina é uma subárea maior dentro da inteligência computacional (também chamada de inteligência artificial). Seu principal objetivo é usar métodos computacionais para extrair informações de dados. (Hsieh, 2009, prefácio, p. ix, tradução nossa)

Um dos principais objetivos da AM é dar ao computador a capacidade de descobrir padrões e generalizar. A partir de conjuntos de dados de treinamento, a máquina deve ser capaz perceber padrões. Quando um novo dado lhe é apresentado, o computador deve ser capaz de identificar a qual padrão de treinamento esse dado se assemelha, gerando respostas compatíveis com o padrão e não com os dados propriamente ditos. Um dos usos para essa forma de AM é a classificação. Problemas, como o de diagnóstico prematuro de ceratocone, podem ser considerados como problemas de classificação difícil. Isso porque a quantidade de fatores que influenciam na classificação é muito grande ou os padrões de distinção são demasiado tênues para serem percebidos pelos humanos, ou combinações dos dois.

Outra finalidade da AM é para regressão. Com objetivo e forma bastante matemáticos, os *modelos de regressão* são aqueles que conseguem definir funções (simples ou compostas por várias funções) que tornam os dados (aparentemente desconexos) em dados estatisticamente analisáveis. Isso é possível pela construção de uma função que pode prever valores a serem fornecidos por uma função, a priori, equivalente e desconhecida. Nesse formato, cada valor que descreve o problema deve ser tratado como uma variável aleatória. Em Haykin (2009), esse tipo de aprendizado é caracterizado pela existência dos seguintes componentes:

Resposta É a variável aleatória que se deseja como resultado da função a ser aprendida (diagnóstico, por exemplo)

Regressores Todas as variáveis aleatórias a serem usadas para identificar o valor da *resposta*

Erro Expectacional ou Erro Explanacional É o grau de incerteza que se tem entre a resposta real e a resposta calculada a partir dos regressores. Durante o aprendizado, o erro pode ser calculado, mas durante a aplicação prática da função aprendida, esse valor pode ser estipulado.

Existem três formas de propiciar o aprendizado à máquina: por reforço, com supervisão e sem supervisão.

O aprendizado por reforço ocorre quando a máquina executa alguma ação sobre o conjunto de possibilidade que lhe são possíveis e “percebe” as consequências disso como avaliações positivas ou negativas, assim ajustando sua conduta de acordo com os objetivos.

A AM com supervisão ou *aprendizado supervisionado* ocorre quando são apresentados conjuntos com um ou mais exemplos, a máquina tenta predizer qual a saída desejada –

classificação ou resultado de função – e então é apresentada a resposta correta. Caso tenha acertado, a máquina pode continuar seu treinamento sem alterações, caso contrário, ajusta seu modelo do problema para tentar superar seu erro.

Já na AM sem supervisão, ou *aprendizado não-supervisionado*, não há uma resposta correta, isto é, espera-se que o computador receba seus exemplos e, ao final do período de treinamento, tenha desenvolvido um sistema de agrupamento ou associação entre os exemplos. Nesse caso, ou se tem uma ideia de algum resultado correto e se espera que o computador identifique algum padrão ainda desconhecido sobre o conhecido, ou não há resultado correto.

A implementação e o uso de cada uma das formas de aprendizado são muito variadas e não existem respostas definitivas para diferentes tipos de problemas. Um mesmo algoritmo de aprendizado pode ter diferentes desempenhos em problemas distintos ou no mesmo tipo de problema, mas com configurações diferentes.

Nas subseções a seguir, estão presentes algumas soluções para problemas de classificação e regressão.

2.3.1 Regressão Linear e Regressão Logística

Tendo seu início com os estudos de Sir Francis Galton, no estudo da relação de transmissão hereditária de características com ervilhas e sido matematicamente aprofundada com as definições fornecidas por Karl Pearson (Stanton, 2001), a *regressão linear* (ou, em inglês, *Linear Regression*) é um método onde, a partir dos dados disponíveis, chamados de dados de treinamento, tenta-se encontrar uma função que represente a variação desses dados. Tal função pode ser usada para prever valores não usados no treinamento ou para separar os dados em classes ou classificações diferentes.

Esse modelo estatístico é baseado no cálculo da *esperança condicional* $E(Y|X = x)$, isto é, tendo um regressor ou variável aleatória X e um conjunto resposta Y aos valores de X , definidos sobre a mesma população, a esperança condicional é o valor esperado para Y , quando a variável $X = x$ (Morettin & Bussad, 2014). Pode ser expresso pela função

$$E(Y|x) = \mu(x) \quad (2.2)$$

onde $\mu(x)$ pode ser qualquer função sobre x .

Segundo Weisberg (2005), a definição mais precisa de regressão linear pode ser dada por

$$E(Y|X = x) = \beta_0 + \beta_1 x \quad (2.3)$$

$$Var(Y|X = x) = \sigma^2 \quad (2.4)$$

sendo essa uma generalização da Equação 2.2, onde β_0 é o valor de $E(Y|X = x)$ quando $x = 0$

e β_1 é a inclinação da reta, a taxa de crescimento da esperança em relação ao valor de X ou, ainda, a taxa de proporção entre as duas variáveis. O valor de β_1 pode ser usado para analisar o quanto X está correlacionado a Y (Weisberg, 2005) – valores próximos a 1 indicam muita correlação entre regressor e resposta; próximo de -1, X é inversamente proporcional Y ; próximo a 0, existe pouca ou nenhuma correlação.

O valor de σ^2 diz respeito a quanto o valor de y_i varia em relação ao seu valor esperado. Isso acontece por causa de fatores como presença de ruído nos valores de X ou modelagem incorreta da função $\mu(x)$. Levando-se em conta a possibilidade de que $\sigma^2 > 0$, isso implica que o valor esperado para y_i com x_i pode apresentar diferenças em relação a \hat{y}_i , chamada de *erro estatístico* (e_i). O valor correto pode, então ser expresso por $y_i = E(Y|X = x_i) + e_i$ (Weisberg, 2005; Morettin & Bussad, 2014).

No caso onde x é uma variável que pode ter um único valor, esse modelo é chamado *regressão linear simples* (ou RSL). No entanto, x pode ser um vetor tal que $x = \{x_1, x_2, \dots, x_n\}$. Esse modelo passa a ser chamado de *modelo de regressão multilinear*, no qual o valor esperado para a resposta y_i passa a ser dado por

$$y_i = E(Y|X = x) + e_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + e_i \quad (2.5)$$

onde cada um dos β_i são análogos aos ocorrentes na regressão linear simples.

Existem casos onde Y possui uma distribuição binomial, isto é, valores dicotômicos como 0 e 1, falso e verdadeiro, sucesso e falha. Nesses casos, pode-se usar um tipo particular de regressão, chamada *regressão logística* (Hosmer & Lemeshow, 2000). Esse tipo de regressão usa uma variação da Equação (2.3), sendo expressa por

$$\pi(x) = \frac{\exp^{\beta_0 + \beta_1 x}}{1 + \exp^{\beta_0 + \beta_1 x}} \quad (2.6)$$

sendo este o resultado de uma transformação da Equação (2.3) definida como

$$\begin{aligned} g(x) &= \ln\left(\frac{\pi(x)}{1 - \pi(x)}\right) \\ &= \beta_0 + \beta_1 x \end{aligned} \quad (2.7)$$

Outra diferença diz respeito ao erro e , pois na RSL o valor de Y está definido sobre uma reta que, a priori, está definida entre $-\infty$ e $+\infty$ e o erro e é uma distribuição Gaussiana de média 0 e de variância constante (Hosmer & Lemeshow, 2000). Já para regressão logística Y terá resultados definidos como 0 ou 1 e o erro passar a ser

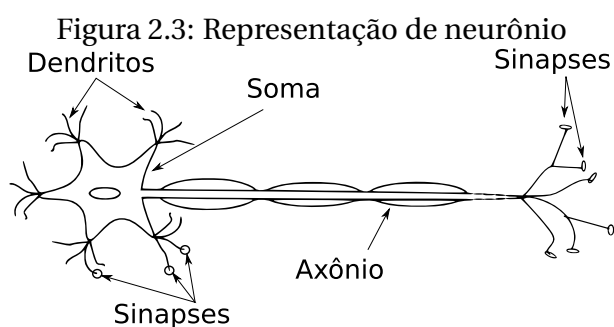
$$e = \begin{cases} -\pi(x), & \text{se } y = 0 \\ 1 - \pi(x), & \text{se } y = 1 \end{cases} \quad (2.8)$$

Não existe uma forma única de encontrar o valor de $E(Y|X = x)$, de modo que métodos

diferentes podem ser tentados na mesmas situações, tal qual o conhecimento específico sobre o problema pode ajudar na escolha de um método (Weisberg, 2005). Por exemplo, em caso de regressão logística um método que pode ser usado é o L-BFGS (ou Limited-memory Broyden–Fletcher–Goldfarb–Shanno, ou, em português, Broyden–Fletcher–Goldfarb–Shanno de memória limitada), que é um algoritmo, computacionalmente viável e iterativo, que funciona bem com variáveis de grandes quantidades de dimensões, fazendo uso de vetores para representar as aproximações de cada iteração, usando derivadas para identificar as próximas otimizações, tendo seu uso em problemas onde a variável X tem seu limites superiores e inferiores definidos (Zhu et al., 1997).

2.3.2 Redes Neurais Artificiais

A inspiração funcional das *redes neurais artificiais* (RNA, ou, em inglês, Artificial Neural Networks) é diretamente o funcionamento do cérebro humano. Para tanto, há o “mapeamento” dos neurônios em funções de comportamento similar. Existem muitas variações de formato de neurônio, mas o interesse foi no funcionamento de forma generalizada. A unidade básica do cérebro humano é o neurônio. Essas células cerebrais possuem prolongamentos que lembram raízes chamados *dendritos*, os quais recebem estímulos nervosos e os propagam para o corpo ou *soma*. De acordo com intensidade e a frequência dos estímulos recebidos, o soma pode ou não gerar um novo pulso e propagá-lo através do *axônio*. Esse, por sua vez, é uma parte alongada que o liga a outras células do mesmo tecido. O ponto de transmissão de estímulo de um axônio para o dendrito é chamado *sinapse* (Haykin, 2009). Uma representação dessas estruturas pode ser vista na Figura 2.3.



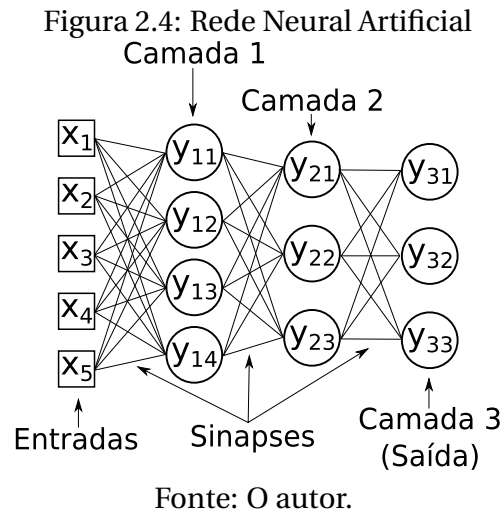
Fonte: Diagrama adaptado de

<https://pixabay.com/pt/neurônio-motor-neurônio-neurone-2040692/> .

De forma análoga, a unidade básica de uma RNA é um *neurônio artificial* (NA). Os dendritos são equivalentes a entradas de uma função (também chamada de *função de ativação* ou ainda FA). Essa, por sua vez, é o equivalente ao soma, enquanto que o resultado da função é a substituição ao axônio. Saídas de um neurônio podem ser encadeadas como entrada dos próximos, formando, assim, uma rede. Na extremidade oposta às entradas, tem-se a saída, que é o resultado do processamento de todos os componentes da RNA. A variação

de intensidade e frequência dos estímulos são simulados pela adição de coeficientes que ponderam cada uma das sinapses, chamados de *coeficientes sinápticos* ou *pesos sinápticos* (Faceli et al., 2011; Haykin, 2009; Valença, 2010; Hsieh, 2009).

Em uma RNA, os neurônios são organizados em camadas. Cada camada tem pelo menos um deles, tendo como entradas, as saídas da camada anterior ou a entrada da rede (Haykin, 2009) – se for a primeira camada – como indicado na Figura 2.4.



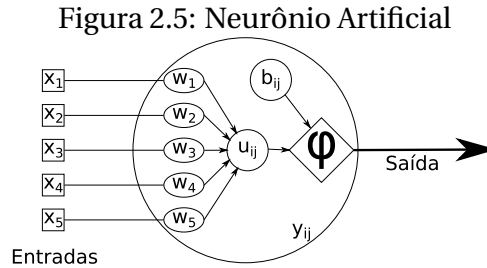
Considerando o j -ésimo neurônio da i -ésima camada, que tem como entrada $x = \{x_1, x_2, \dots, x_M\}$ e coeficientes sinápticos $w = \{w_1, w_2, \dots, w_M\}$, a unificação dessas entradas está representada na Equação 2.9.

$$u_{ij} = \sum_{k=1}^M x_k w_k \quad (2.9)$$

Um dispositivo adicionado artificialmente é o *bias* (em português, viés) b , que modifica o efeito do processamento na unidade. Assim, a saída de uma dessas unidades de processamento pode ser definida como na Equação 2.10, onde φ é uma função contínua e diferenciável, podendo ser linear, limiar, sigmoideal, tangente hiperbólica, dentre outras (Calôba, 1992). Uma demonstração dessa estrutura está presente na Figura 2.5.

$$y_{ij} = \varphi(u_{ij} + b_{ij}) \quad (2.10)$$

A aprendizagem das redes neurais artificiais se dá pelo ajuste dos valores dos pesos sinápticos, que ocorre durante a fase de treinamento. É necessária a apresentação de outro fator artificial que é a *taxa de aprendizado* (em inglês, *Learning Rate*) α . Esse fator controla o quão rápido uma rede neural aprende. Com uma taxa de aprendizado muito alta, a rede neural aprende muito rápido, mas não “entende” o suficiente e a sua capacidade de generalização pode ser comprometida. Por outro lado, se a taxa é muito baixa, o aprendizado é muito lento e pode demorar demais para que a rede aprenda o suficiente com os exemplos e seu



Fonte: O autor.

desempenho final também não será bom o suficiente. Caso se conheça o suficiente do problema, pode-se definir um valor α que garanta não ser alto ou baixo demais. Outra alternativa é que a esse valor comece alto e no decorrer do treinamento, diminua (Calôba, 1992; Haykin, 2009; Valença, 2010).

Considerando w_{ij} como o conjunto de pesos sinápticos do j -ésimo neurônio da i -ésima camada, α a taxa de aprendizado da RNA, d_{ij} como sendo a saída esperada e y_{ij} a saída calculada pelo neurônio, a cada iteração de treinamento supervisionado, o valor de w_{ij} pode ser modificado pela adição do valor Δw_{ij} , calculado a partir da Equação 2.11.

$$\Delta w_{ij} = \alpha(d_{ij} - y_{ij})x_j \quad (2.11)$$

E Δw_{ij} também pode ser dado por

$$\Delta w_{ij} = w_{ij,t} - w_{ij,t-1} \quad (2.12)$$

onde os índices t e $t-1$ representam a iteração atual e anterior, respectivamente.

Igualando as Equações 2.11 e 2.12, pode-se escrever

$$w_{ij,t} = w_{ij,t-1} + \alpha(d_{ij} - y_{ij})x_j \quad (2.13)$$

No entanto, essa forma de aprendizado só pode ser aplicado a redes do tipo *Perceptron*, que possuem apenas uma única camada. Para redes mais complexas (como a da Figura 2.4), isso é resolvido com *MLP* (*Multi Layer Perceptron*, ou, em português, Perceptron de Múltiplas Camadas) e *backpropagation*. Uma rede do tipo MLP é como anteriormente indicado, múltiplas camadas, onde a saída de um neurônio da camada i serve como entrada dos neurônios da camada $i + 1$ e assim por diante, até a saída. A resposta fornecida a uma RNA durante o treinamento não inclui a saída de cada neurônio, apenas o resultado final desejado. Assim, faz-se o uso do *backpropagation*. Nesse algoritmo, o valor esperado como saída de um neurônio no interior da rede é o dado pela Equação (2.14), onde δ_{i+1} é o erro ocorrido na camada seguinte. Esse valor pode ser calculado pela Equação (2.15), onde f'_j é a derivada da j -ésima função de ativação da camada seguinte e $w_{(i+1)j}$ é o peso sináptico dado por esse neurônio a saída do atual (Calôba, 1992). Cada unidade da RNA pode ter sua própria

função de ativação, o que não é uma prática comum.

$$w_{ij,t} = w_{ij,t-1} + \alpha \delta_{i+1} x_j \quad (2.14)$$

$$\delta_{i+1} = \sum_{j=1}^N f'_j w_{(i+1)j} \quad (2.15)$$

Mesmo sendo um modelo de aprendizagem bastante plástico, também pode ser muito problemático, se mal modelado. De forma geral, quanto maior a entrada da rede, maior deve ser o conjunto de exemplos para seu treinamento, não havendo quantidades mínimas ou máximas. Como previamente discutido, a escolha da taxa de aprendizado também deve ser cuidadosamente ajustada (Valença, 2010).

Cada vez que todos os exemplos são usados para treinamento, isso é chamado de *era*. Poucas eras podem ser insuficientes para a rede aprender o suficiente sobre o problema. Por outro lado, eras demais pode levar a RNA a “decorar” os exemplos e perder a capacidade de generalizar (*overfit*) (Valença, 2010; Calôba, 1992).

Um tipo particular de redes neurais são *redes neurais com recorrência*. Nesse tipo, os neurônios apresentam estruturas extra para recorrência, tendo suas saídas como entradas nos próprios neurônios, em vizinhos de camada ou em camadas anteriores. Com esse procedimento, dependendo da estrutura usada, isso pode vir a auxiliar na memorização de padrões ou para processamento em cunho temporal, em que a entrada recorrente serve como fator de influência do instante anterior no processamento atual (Haykin, 2009).

Ao contrário de outros métodos, como as árvores de decisão ou as regressões, olhar o conteúdo de uma RNA não fornece informação a um usuário, pois o aprendizado se resume a valores dentro da estrutura ou parâmetros de uma complexa fórmula, não possuindo uma significação direta (Haykin, 2009).

Outro problema diz respeito ao dimensionamento. Não existe uma regra que indique qual a quantidade de neurônios que deva ter em cada camada ou quantas camadas deve ter uma rede para resolver um determinado tipo de problema. Conhecimento sobre o problema e sobre sua execução podem ajudar nas escolhas dessas quantidades. Em Calôba (1992), há a descrição de um procedimento para a construção de uma RNA, item a item começando pela primeira camada, primeiro neurônio:

1. Treine o neurônio até que consiga distinguir a maior quantidade possível de classes corretamente.
2. Defina um subconjunto de exemplos, contendo apenas aqueles que não puderam ser identificados corretamente.
3. Adicione um novo neurônio e repita o passo 1, usando o novo neurônio como inicial e o conjunto de exemplo sendo o do passo 2.

4. Caso não haja mais exemplos e a quantidade de classes a identificar for maior que a quantidade de neurônios da última camada, repita o passo 1 com uma nova camada com um neurônio e usando todos os exemplos.

Este procedimento gera uma RNA onde a quantidade de neurônios na n -ésima camada é menor que o da camada anterior. No entanto, não há certeza se este é o melhor procedimento para construção e se os resultados são os mais apropriados, apenas garantindo uma solução.

Ainda assim, as redes neurais artificiais possuem robustez para serem tolerantes a falhas (podem operar com alguns componentes não funcionando corretamente), capazes de fazer generalizações e adaptáveis a presença de ruído nos dados de entrada (Haykin, 2009; Calôba, 1992; Valença, 2010).

2.3.3 SVM e *Fast Large Margin*

Uma SVM (*Support Vector Machine*, ou, em português, Máquina de Vetor de Suporte) é um de tipo de aprendizado de máquina em que o objetivo é encontrar a melhor função de separação de classes. Para isso, a medida de desempenho é a capacidade de generalização a ser maximizada. Isto difere dos processos de regressão como os da seção 2.3.1, pois nesses últimos, o objetivo é encontrar uma função qualquer que atenda aos critérios de separação ou previsão. Enquanto que para as SVMs, o intuito é encontrar a melhor função possível dentro das suas parametrizações (Russel & Norvig, 2010).

Isso é feito pela análise estatística dos elementos a serem usados como exemplos. São escolhidas aquelas amostras que melhor representam as “fronteiras” de separação entre as classes de dados. Essas amostras formarão a base para definição de vetores que são usados para identificar o hiperplano que melhor separa as classes. Esses pontos são chamados de margens (Arreola et al., 2007; Lorena & Carvalho, 2007).

O hiperplano gerado em uma SVM pode ser equivalente àquele em uma regressão linear. No entanto, em uma regressão é encontrado um hiperplano baseado em cada exemplo adicionado ao treinamento, já com uso de SVM, a separação é feita com base estatística dos dados e caso os novos exemplos de treinamento não estejam entre as margens, o hiperplano não muda (Arreola et al., 2007).

O desenvolvimento das SVMs se baseou na *teoria do aprendizado estatístico* (TEA) de Vladimir N. Vapnik.

Essa teoria estabelece uma série de princípios que devem ser seguidos na obtenção de classificadores com boa generalização, definida como a sua capacidade de prever corretamente a classe de novos dados do mesmo domínio em que o aprendizado ocorreu (Lorena & Carvalho, 2007, p. 1).

Tendo-se F como o conjunto de todos os classificadores f ($f \in F$) que podem ser gerados via aprendizado de máquina, usando um conjunto de treinamento T – formado de pares

(x_i, y_i) – onde x_i é uma amostra dentro do conjunto de exemplos disponível X ($x_i \in X$) e y_i é um rótulo que incide sobre x_i designando a que classe pertence e y_i pertencendo ao conjunto de rótulos Y ($y_i \in Y$), uma SVM objetiva encontrar $\hat{f} \in F$, sendo \hat{f} a função que apresenta máxima verossimilhança com a função ótima de separação das classes presentes em X e rotulados por Y . Para medir a proximidade de \hat{f} ao melhor classificador, é necessária a medida do erro $R(f)$ de cada uma das funções f a ser avaliada. Tal erro é chamado de risco e pode ser dado por:

$$R(f) = \int c(f(X), Y) P(X, Y) \quad (2.16)$$

onde $c(f(X), Y)$ é uma função de custo relacionada a $f(X)$ e $P(X, Y)$ é a distribuição de probabilidade que relaciona o conjuntos X e Y .

No entanto, como a função ótima é desconhecida e, por consequência, a sua distribuição de probabilidade também, $R(f)$ não pode ser calculada diretamente, mas pode ser estimado de forma empírica – pela designação $R_{emp}(f)$ – com o cálculo do custo baseado nas amostras disponíveis.

$$R_{emp}(f) = \frac{1}{n} \sum_{i=1}^n c(f(x_i), y_i) \quad (2.17)$$

onde n é a quantidade de exemplos em X . Com isso

$$R(f) \leq R_{emp}(f) + R_{est}(f) \quad (2.18)$$

e $R_{est}(f)$ é o *risco estrutural* ou *termo de capacidade*, que é o risco referente a todo F e é dado por

$$R_{est} = \sqrt{\frac{h(\ln \frac{2n}{h} + 1) + \ln \frac{\theta}{4}}{n}} \quad (2.19)$$

tendo h como a dimensão Vapnik-Chervonenkis (VC) “definida como o número máximo de exemplos que podem ser particionados em duas classes pelas funções contidas em F , para todas as possíveis combinações binárias desses dados” (Lorena & Carvalho, 2007, pp.48); quanto maior o valor h , mais complexas são as funções em F . A garantia de $R(f)$ (na Equação 2.18) é dada com probabilidade $1 - \theta$, com $\theta \in [0, 1]$. Encontrar $R(f)$ é – por consequência das Equações 2.18 e 2.19 – minimizar o risco empírico e o erro estrutural.

Outra medida de qualidade de uma SVM é a *largura da margem*. As margens são as distâncias entre as classes de dados e o classificador. Quando se busca uma solução em que não haja possibilidade de existência de amostras ocorrerem entre as margens, essas são ditas *margens rígidas*, caso contrário, *margens suaves*. Nesse último caso, as amostras entre as margens são chamadas de *ruído* (Arreola et al., 2007).

Ao encontrar uma função de separação, uma SVM encontrou estatisticamente a solução

baseada no conjunto de exemplos disponível, mas o universo de todas as amostras possíveis a um problema pode ser desconhecido. Isso implica que alguma amostra não usada no treinamento pode estar mais distante dos exemplos da classe a que pertence e mais próxima de outra classe. Ao buscar uma solução com margens mais largas (distantes umas das outras), está-se buscando uma solução com maior capacidade de generalização e aumentando a possibilidade da existência de elementos que ocorram entre as margens e a fronteira que as divide, mas sem cruzá-la (Arreola et al., 2007).

As restrições de uma SVM são, dessa forma, encontrar a função mais simples que execute essa tarefa, que diminua o risco empírico e que maximize as margens. A busca de solução é iniciada com funções em baixa dimensionalidade (h baixo). A proporção que se identifica que não há possibilidade de uma separação linear em uma certa quantidade de dimensões, estas podem ser mapeadas em um espaço dimensional maior até que se encontre um hiperplano n -dimensional que melhor satisfaça as restrições (Arreola et al., 2007).

Esta técnica de AM pode ter seu desempenho gravemente degradado quando aplicado em problemas cuja quantidade de dados por amostra é muito grande. Para tanto, existem as técnicas de *Fast SVM* (ou, em português, SVM Rápidas). Nessas técnicas, são usadas outras abordagens em conjunto, a fim de particionar os dados ou o conjunto de treinamento. Um exemplo é o uso de árvores de decisão para subdivisão das amostras de treinamento, tendo como solução a união de vários hiperplanos, cada um sobre um subconjunto (Arreola et al., 2007; Segata & Blanzieri, 2011).

Um subgrupo dentro das *Fast SVMs* é a *Fast Large Margin* (ou, em português, Margens Largas e Rápidas), em que a principal restrição é a maximização das margens aplicada em união a técnicas de particionamento do conjunto de amostras para treinamento (Arreola et al., 2007; Lorena & Carvalho, 2007).

Apesar dos princípios de minimização dos riscos, como toda técnica de AM, SVM não possui garantias de encontrar o melhor de todos os resultados, mas garante resultados ótimos dentro da configuração de implementação. Um exemplo disso pode ser visto ao se aplicar SVM a um conjunto de treinamento muito pequeno e com dimensionalidade muito grande. Nessa condição, a capacidade de generalização é comprometida devido falta de exemplos suficientes. Por outro lado, conjuntos de treinamento muito grandes podem ser computacionalmente muito custosos (Russel & Norvig, 2010; Arreola et al., 2007; Lorena & Carvalho, 2007; Fan et al., 2008).

2.3.4 k -NN

Um classificador que pode ser considerado simples é o k -NN (do inglês, k -Nearest Neighbors, ou, em português, k -Vizinhos mais Próximos). Apesar da prova estatística ter sido apresentada em Cover & Hart (1967), a base para classificação de padrões a partir de exemplos próximos já tinha sido lançada no século 11 por Alhazen em seus estudos sobre a visão

humana e o reconhecimento de objetos, como indicado em Pelillo (2014).

Ele se baseia em encontrar a classificação mais apropriada, levando em consideração os k elementos já classificados mais semelhantes ao elemento em análise (O’Neil, 2013). Essa similaridade é dada por uma função auxiliar, como a *distância Euclidiana* ou a *correlação de Pearson* (Segaran, 2008), que faz uso de um par de vetores de valores e dá como resultado um valor que indica o quão próximos (com a distância Euclidiana) ou correlacionados (com a correlação de Pearson) os vetores são.

Os parâmetros da função de similaridade são dados componentes desse par de vetores. Esses dados possuem limitações de uso, como, por exemplo, dados não-escalares não têm contexto quando aplicados na função de similaridade (nomes de cidades, por exemplo). Outro problema é que os dados componentes do vetor não podem ter muita discrepância na escala de grandeza, pois variações de um dado com maior escala de grandeza podem “dominar” o cálculo de similaridade sobre dados com menor escala, além que dados em uma escala muito pequena podem não surtir muito efeito entre dados de escala muito maior (O’Neil, 2013), havendo necessidade de normalização dos dados. Outro problema é o relativo a redundância, pois a presença de muitos dados que se referiam a mesma informação podem tornar outros não-redundantes menos expressivos na classificação, mesmo que sejam realmente mais apropriados a isso (O’Neil, 2013).

Em uma situação ótima, cujos itens a serem classificados são completamente díspares, pode-se fazer $k = 1$. No entanto, nas outras situações, é necessário fazer algum tipo de ponderação entre os vizinhos e suas classificações. A forma mais simples é por frequência de votação, isto é, a classificação mais comum entre os vizinhos é escolhida como a correta para o item. Quando a quantidade de vizinhos é grande e a classificação possui muito ruído, pode-se usar a ponderação por distância, no qual o resultado é ponderado entre o quão semelhante é um vizinho com sua classificação; o somatório de todas as distâncias dos vizinhos é usado para ponderar a influência das classificações (Segaran, 2008; Zhang & Zhou, 2007).

2.3.5 Naïve Bayes

Tendo-se n características em uma amostra A e que essa possui uma classificação C_j , pode-se criar um classificador que calcule $P(C_j|A)$. Isso pode ser feito através do uso do *Teorema de Bayes*, como descrito em Magalhães & Lima (2007), e usando um conjunto de treinamento que contenha C_i amostras classificadas. De acordo com Langley et al. (1992), a probabilidade de uma amostra A pertencer a uma classe C_i é dada pela composição das probabilidades de cada uma de suas n característica de pertencer a essa classe, isto é:

$$P(C_i) = \prod_{k=1}^n P(A_k) \quad (2.20)$$

Esse valor pode ser aplicado, então, para calcular a probabilidade de classificação de uma

amostra pelo uso do *Teorema de Bayes*, através da inversão das suposições e usando a base de treinamento para tentar inferir uma classificação de uma nova amostra por meio de:

$$P(C_j|A) = \frac{P(A|C_j)P(C_j)}{\sum_{i=1}^k P(A|C_i)P(C_i)}, j = 1, 2, \dots, k \quad (2.21)$$

Esse modo de aprendizado é genérico, podendo ser aplicado a qualquer classificador *Bayesiano*. De forma específica, esse tipo de classificador é usado em situações cujas características são conhecidas e ditas probabilisticamente independentes. Um subtipo é *Naïve Bayes*, ou *Bayes Ingênuo*, ou *Classificador Ingênuo*, em que todas os componentes de A são considerados independentes, mesmo que não o sejam. Classificações baseadas em séries temporais (por exemplo), cujos valores possuem ordenação temporal, a sucessão perde o contexto e torna-se irrelevante.

Em [Frank et al. \(2000\)](#), é possível perceber que se um problema puder ser resolvido por técnicas que envolvam regressões lineares, *Naïve Bayes* pode até ter bons resultados, mas o uso de regressão é mais apropriado. Isso pode ser deduzido pelo fato de que se um problema possui soluções previsíveis por regressão linear, isso implica que seus dados de treinamento possuem dependência e podem ser agrupados como uma função. Essa dependência é uma informação que é descartada com o classificador ingênuo.

Uma vantagem desse método é que o crescimento da quantidade de dados por evento, leva ao aumento linear do tempo computacional, ao contrário de outras técnicas, como as redes neurais, que têm seu crescimento computacional elevado de forma exponencial. Outra vantagem, é que pode usado em problemas que haja dependência semântica entre o dados, mas os mesmos não possuem representação numérica apropriada para métodos regressores. É um classificador que apresenta resultados reais com boa acurácia, como a classificação de textos realizada em [Prinyakupt & Pluempitiwiriyawej \(2015\)](#).

Esse método de aprendizagem não requer, também, o esforço computacional de remoção de características irrelevantes, visto que isso pode ser computacionalmente custoso, além de que essas características não interferem no cálculo das probabilidades. Enquanto que para outros métodos, a seleção dentre as características relevantes pode auxiliar na diminuição do tempo de execução e no melhoramento de resultados, está demonstrado em [Frank et al. \(2000\)](#) e [Langley et al. \(1992\)](#) que este método dispensa esse tipo de tratamento de dados.

Em [Frank et al. \(2000\)](#) também está demonstrado que, se as características possuem pouco ruído, este é um método que requer poucos exemplos para se ter acurácias de classificação alta e pouco erro. Em comparação com outros métodos de aprendizagem, o treinamento com *Naïve Bayes* requer poucos exemplos para apresentar alta acurácia de classificação, mas a melhoria desse resultado não pode ser garantida pelo aumento da quantidade de exemplos, ocorrendo o inverso com outros métodos de aprendizagem.

2.3.6 *Deep Learning*

Nos estudos iniciais sobre redes neurais artificiais, como os de Pitts e McCulloch ou Hopfield, as redes eram modeladas tendo muitos neurônios em poucas camadas, havendo propostas de modelagem com compatibilidade entre a quantidade de entradas de cada exemplo com neurônios por camada. Em entradas com tamanho muito grande, o custo computacional é muito alto e não há garantias de resultados ótimos para esse procedimento. Em 2006, pesquisadores do Canadian Institute for Advanced Research (CIFAR) passaram a demonstrar interesse em redes com maior profundidade (mais camadas) (Schmidhuber, 2014).

Os pesquisadores introduziram procedimentos de aprendizado não-supervisionado que podiam criar camadas de detectores de características sem requerer dados rotulados. O objetivo em aprender cada camada de detector de características era ser capaz de reconstruir ou modelar as atividades de detectores de característica (ou entradas cruas) nas camadas abaixo. (LeCun et al., 2015, p. 4, tradução nossa).

Isso quer dizer que cada tipo de detector de característica possui um funcionamento e finalidade diferente, recebendo dados das camadas anteriores, identificando as características que lhe são cabíveis e repassando para as camadas seguintes características identificadas (LeCun et al., 2015). As características (ou artefatos) a serem detectados são quaisquer subconjuntos de dados dentro de uma amostra que podem ser destacados para identificação total ou parcial de uma amostra, como bordas de figuras, áreas de cor homogênea, picos de frequência em uma onda sonora, preposições dentro de uma frase, etc.

A sucessão desses procedimentos faz com que as camadas posteriores façam processamento de dados pré-processados, tendendo a exigir menor esforço computacional que a camada anterior. O termo *deep learning* (ou DL ou, em português, aprendizagem profunda) vem da consideração que a camada de entrada é a mais superficial e cada detector é adicionado como uma camada abaixo. Assim, quanto mais camadas, maior a introspecção sobre os dados, ou ainda, mais profundo é o aprendizado sobre as entradas (Schmidhuber, 2014). Esse tipo de estrutura tem sido usada principalmente para análise de sinais (como áudio), imagens, vídeos e texto.

A camada de *convolucional*, por exemplo, realiza convolução de uma entrada n -dimensional em n -dimensões, gerando um mapa de características de bordas.

A camada de *pooling* (ou, em português, agrupamento) agrupa dados n -dimensionais de acordo com critérios estatísticos, como média, mediana, máximo ou mínimo. Ela auxilia na identificação de áreas com maior relevância local. No caso de imagens, impõe *invariância*¹ aos artefatos destacados e remoção de áreas sem importância. Impõe também a quantidade

¹Considerando a existência de um artefato a ser identificado em uma imagem (um objeto qualquer, por exemplo), a invariância é a capacidade de identificar o artefato ignorando sua rotação ou translação dentro da imagem.

máxima de características a serem processadas na próxima camada, podendo diminuir drasticamente o volume de dados.

De acordo com LIS (2015), a união dessas duas camadas tem inspiração biológica nos estudos de Hubel e Wiesel (1968) sobre a visão dos gatos. O arranjo de células do córtex visual é sensível a sub regiões do campo visual. Essas células atuam como filtros de imagens preparados para explorar as correlações de artefatos presentes em uma imagem, identificando bordas e aplicando invariância para entender o conteúdo do que está sendo recebido.

Isso pode ser constatado em Karpathy et al. (2014), quando esses conceitos são aplicados a reconhecimento de vídeo e visão computacional.

Primeiro, características visuais locais que descrevem uma região do vídeo são extraídas de conjuntos de pontos de interesse densos ou esparsos. Em seguida, as características são combinadas em uma descrição de tamanho fixo em nível de vídeo. (...). Finalmente um classificador (...) é treinado em uma representação em uma “sacola de palavras” para distinguir ao longo de classes visuais de interesse (Karpathy et al., 2014, p. 2, tradução nossa).

As LSTMs (*Long-Short Term Memory* ou, em português, Memória de Longo e Curto Prazo) são camadas que funcionam como redes neurais recorrentes. As LSTMs têm sido usadas principalmente para identificação de objetos em cenas, geração de descrição de conteúdo de vídeos, tradução e reconstrução de texto. Por mais distintos que esses usos pareçam ser, na verdade, são muito semelhantes. Quando se usa camadas de convolução e *pooling*, várias características podem ser extraídas. Informação do mesmo tipo pode ser extraída de seqüências de palavras. Assim, para uma LSTM, é uma questão de realizar o mapeamento de palavras ou artefatos em conjuntos de classes e, posteriormente, mapear as classes de entrada em outras classes a serem usadas diretamente (como no processo de classificação de conteúdos de vídeos) ou processadas por camadas mais profundas (Wu et al., 2016; Karpathy et al., 2014; Venugopalan et al., 2015).

As camadas de *reshape* (ou, em português, remodelagem) e *flatten* (ou, em português, achatamento) mudam a forma dos dados. Enquanto a de *flatten* diminui uma entrada de n -dimensional para uma única dimensão, a de *reshape* muda para qualquer dimensão e formato (por exemplo, se a entrada for uma matriz de 4×16 , pode ser mudada para 2×32 ou para uma matriz cúbica $4 \times 4 \times 4$) (Schmidhuber, 2014).

Também é comum o uso de MLPs nas camadas mais profundas de um algoritmo de DL. Caso se aplique apenas essas camadas sobre as entradas, pode haver custo computacional muito alto, o mesmo não ocorre quando processa dados pré-processados ou se realiza classificação dos artefatos identificados/separados pelo restante da estrutura de aprendizado (Schmidhuber, 2014; LeCun et al., 2015).

Como em outras estruturas de aprendizado de máquina, DL não possui implementação ou configuração única para todos os tipos de problemas. Apesar de melhores resultados em

diversos tipos de problemas de processamentos de entradas, as redes de *deep learning*, tal qual outras redes neurais, não possuem garantia de resultados. Estudos na área têm demonstrado que estruturas mais profundas e/ou com presença de camadas com processamento paralelo (um mesmo resultado é aplicado a conjuntos distintos de camadas e posteriormente reunidos para fornecer resultados em conjunto) apresentam melhor identificação de artefatos e classificação (Szegedy et al., 2015).

Por outro lado, está demonstrado em Zeiler & Fergus (2013), que conhecimento sobre o que se deseja classificar e como um algoritmo de *deep learning* aprende com os exemplos é fator preponderante para construção de uma estrutura mais apropriada para desenvolvimento de uma boa solução de classificação. Nesse estudo é, evidenciado, também, um outro tipo de camada, chamada de *deconvolução*. Essa camada constrói imagens que representam as características identificadas em uma entrada.

O custo computacional de certas DLs pode ser muito alto, sendo comum que alguns *frameworks* trabalhem apenas com GPU, para diminuir o tempo de processamento. Estruturas mais complexas exigem, inclusive, o uso de *clusters*² ou computadores com múltiplas CPUs e GPUs (como as estruturas da GoogleNet apresentada em Szegedy et al., 2015).

Apesar da maioria dos estudos em DL envolver conjuntos de amostras provindos de fontes como a Internet – com milhares ou até milhões de exemplos para cada classificação desejada – ao aplicar em tarefas mais específicas e de menor escala, a quantidade de amostras necessárias para treinamento pode ser menor do que para fazer o mesmo serviço em RNAs ou SVMs (Szegedy et al., 2015). Pela mesma necessidade de quantidade de elementos para treinamento é comum o uso de camadas de *dropout* (ou, em português, “cair fora”) que descarta o conteúdo de camadas anteriores e insere zeros, de forma estatística, tendo como objetivo evitar o *overfit* (Schmidhuber, 2014).

2.4 Wavelet

Como indicado na Seção 2.3, o aprendizado de máquina pode ser usado na descoberta de padrões. No entanto, alguns padrões são baseados em conjuntos muito grandes de dados e alguma simplificação se faz necessária.

Funções complexas e de característica estacionárias (como funções que possuem frequência de repetição) podem ser reproduzidas através de *transformadas de Fourier*, isto é, uma função temporal é usada como *base* e outras funções são adicionadas a ela, de forma a compor a função desejada cuja formulação pode ser desconhecida, mas seu gráfico é conhecido. Essa estrutura também pode ser usada para codificação de ondas estacionárias. Para situações menos triviais – como funções temporais sem frequência ou ondas não estacionárias, mas ainda com características temporais – pode-se usar *transformada de Fourier de Tempo-Curto*.

² *Clusters* são agrupamentos de computadores interconectados que realizam processamento em paralelo como se fosse um único super-computador.

Nesse procedimento, a frequência é inserida de forma artificial pelo uso de janelas de tempo (Daubechies, 1990).

A *transformada de wavelet* (ou TW, ou, em inglês, *wavelet transform*, ou WT, ou apenas *wavelet*) tem a mesma função de uma transformada de Fourier de Tempo-Curto, mas com funcionamento diferente. Uma TW aplica transformações de tempo e de escala, via contração ou dilatação de ondas. A diferença permite a TW ser usada em outras áreas e de formas diferentes (Daubechies, 1990).

É possível o uso de wavelet com dados discretos (em inglês *discrete wavelet transform* ou DWT), permitindo o uso sobre codificação/compressão de imagens e vetores de dados discretos (Calderbank et al., 1998). Principalmente para DWT, é comum o uso famílias de wavelets que são ondas “construídas” em estudos para usos em diversas áreas como a Haar (Figura 2.7) e Coiflet (Figura 2.8). Ao ser aplicada a conjuntos de dados, uma DWT provoca a simplificação pela criação de uma representação com menor quantidade de componentes. Esse processo pode ser feito repetidamente.

O processo de composição pode ser iterado com sucessivas aproximações sendo decompostas por vez, então aquele sinal original é quebrado em muitos componentes de menor resolução. Isso é chamado decomposição wavelet de múltiplos níveis (Kaur et al., 2013, p. 671, tradução nossa).

A Figura 2.6 representa esse processo de decomposição. No canto superior esquerdo encontra-se um mapa do processo de decomposição e, a direita, as ondas resultantes. O componente assinalado com S é a onda original. Cada uma dos cA_i são as sucessivas aproximações, enquanto que os cD_i são os detalhes removidos na iteração.

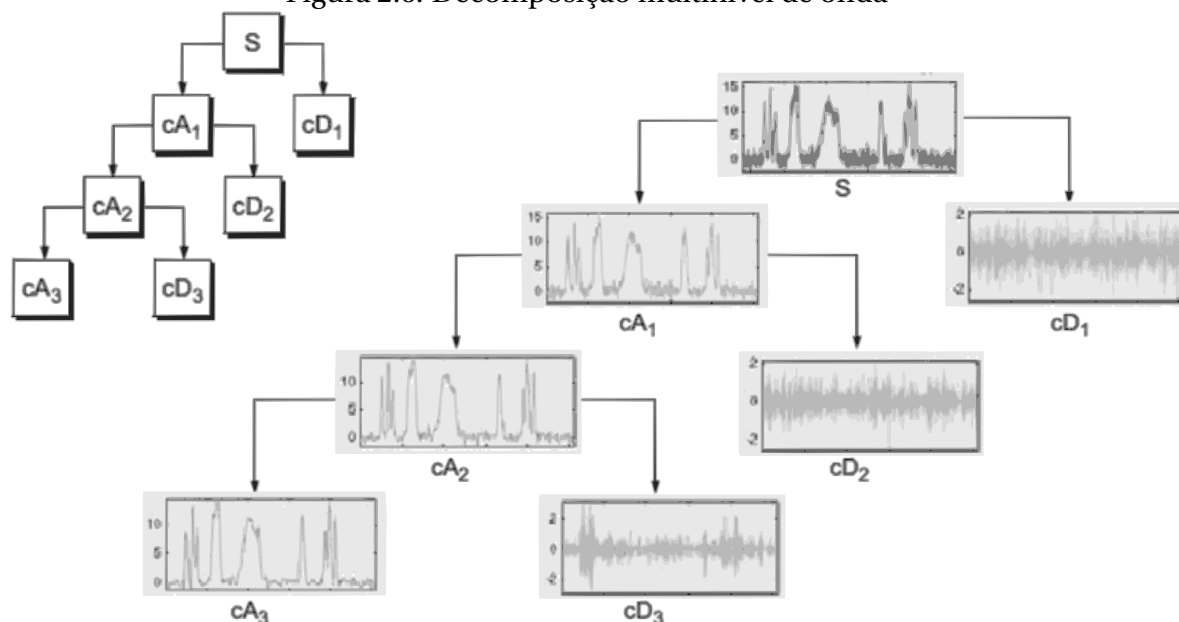
2.5 Biomecânica

Apesar dos avanços tecnológicos terem permitido à medicina novas formas de examinar e tratar pacientes, algumas vezes isso pode não ser suficiente para compreensão de uma doença e suas formas de tratamento. Dessa forma, conhecimentos de outras áreas de estudo são usados para contribuir, possibilitando maior entendimento do funcionamento do corpo humano e de doenças, permitindo a criação de novos conhecimentos, como os da *biomecânica*.

“Biomecânica é o estudo das estruturas e funções de sistemas biológicos através dos métodos da mecânica” (Hatze, 1974, pp.1, tradução nossa). Isto é, o uso de conhecimentos da área de mecânica, aplicados para estudar sistemas biológicos em seu comportamento normal e prever mudanças devido alterações das forças envolvidas ou de estruturas componentes, além de propor métodos de intervenção artificial (Fung, 1993).

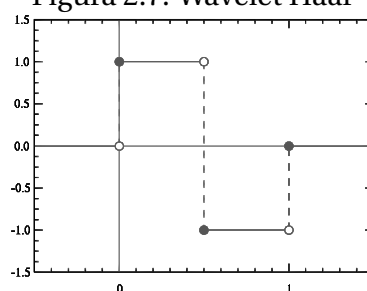
É uma área de estudo que tem recebido cada vez mais impulso na medicina, servindo como um meio para integração com outras áreas de conhecimento, como física e ciências

Figura 2.6: Decomposição multinível de onda



Fonte: Diagrama adaptado de <https://www.mathworks.com/help/wavelet/gs/discrete-wavelet-transform.html?reqbaseadostedDomain=www.mathworks.com>

Figura 2.7: Wavelet Haar

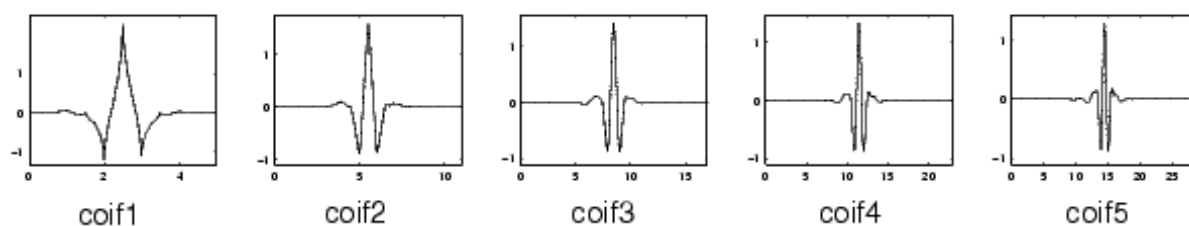


Fonte: https://en.wikipedia.org/wiki/Haar_wavelet.

da computação. Um exemplo disso é o uso de *elasticidade* e *viscosidade* para caracterização dos tecidos que compõem as diferentes partes de um organismo vivo. Elasticidade é a capacidade de um corpo de se deformar, ao sofrer atuação de forças externas, e retornar ao seu estado inicial, quando as forças param de atuar. Já a viscosidade é a característica física dos fluidos de resistir ao escoamento. No corpo humano, muitos tecidos são constituídos de grandes quantidades de água, substâncias elásticas e organização tissular que permitem deformação. Assim, é comum encontrar tecidos com característica *viscoelástica* (viscosa e elástica ao mesmo tempo). Não incomum, também, é que um mesmo tecido contínuo possua diferente viscoelasticidade em porções diferentes, devido a variações de espessura, concentração ou organização de seus elementos constituintes.

Por exemplo, [Shin et al. \(1997\)](#), demonstraram que devido à região periférica da córnea ser mais espessa e diferente organização das fibras do estroma (ver seção 2.6), há menor

Figura 2.8: Família de wavelets Coiflet



Fonte: http://matlab.izmiran.ru/help/toolbox/wavelet/ch01_33a.html.

resistência e maior elasticidade no centro da córnea do que na sua periferia.

2.6 Anatomia Corneal

A anatomia corneal é bem conhecida, sendo amplamente discutida na literatura de cursos de medicina, como em [Rodrigues & Dantas \(2001\)](#), [Vaughan & Asbury \(1990\)](#) e [Kanski \(2003\)](#).

A córnea é um tecido transparente, constituindo a cobertura do olho juntamente à esclera (ou branco do olho) e se ligando a esta na porção do olho conhecida como limbo. É avascular e possui terminações nervosas. Em um adulto, a espessura média é de 0,52mm no centro e 0,64mm na periferia. Não é circular, tendo diâmetro vertical aproximado de 11,5mm e horizontal de 12mm. É constituída, em sua maioria, de água (cerca de 78%), que, em conjunto com a organização de suas 5 camadas, é responsável por dois terços da capacidade refrativa do olho.

Da porção anterior até a posterior, a córnea possui as seguintes camadas:

Epitélio constituído de cinco ou seis camadas de células. É responsável por regular o fluxo de entrada de água e nutrientes, vindo do filme lacrimal, além da remoção de dejetos provindos das outras camadas. As células do epitélio estão em constante renovação, tendo suas células-tronco no limbo e descamando pelas suas duas camadas mais externas. Devido a isso, possui rápida recuperação a lesões, que geralmente não deixam cicatrizes.

Camada (ou membrana) de Bowman é uma condensação do estroma, logo atrás do epitélio, possuindo uma função ainda não identificada, mas auxiliando na elasticidade da córnea. Lesões em sua estrutura podem causar cicatrizes.

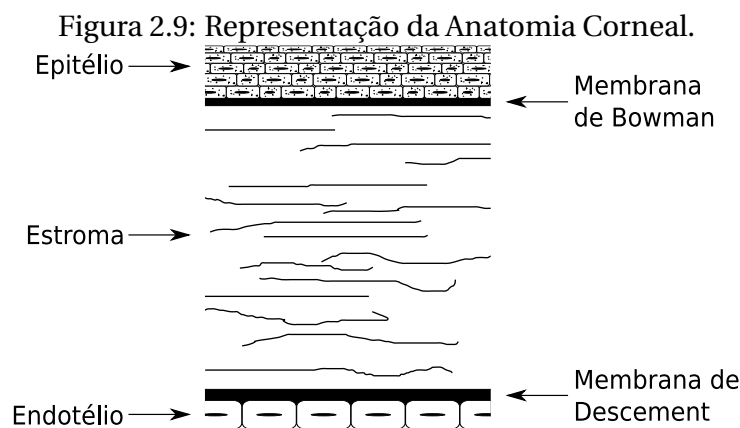
Estroma é uma camada composta de fibrilas de colágeno, proteoglicanos e fibroblastos modificados (ceratócitos), organizados paralelamente à superfície da córnea. É mais compacta nas proximidades da camada de Bowman e é organizada em camadas (lamelas). Além disso, no centro da córnea, a organização é, em sua maioria, no sentido medial-lateral e inferior-superior, enquanto que, na periferia, organiza-se em sentido

circular (Shin et al., 1997). Essa mesma organização é responsável pela transparência e forma da córnea e constituindo 90% de sua espessura. É uma camada que possui muita sensibilidade devido a presença de terminações nervosas.

Membrana de Descemet são fibras de colágeno que separam o estroma do endotélio.

Endotélio é uma única camada de células que não se reproduzem. Sua origem ocorre ainda durante o período de formação do olho. É comum que lesões nessa camada gerem cicatrizes. É responsável pelo controle de hidratação da córnea, removendo o excesso de água.

Uma representação das camadas da córnea está presente na Figura 2.9.



Fonte: Diagrama adaptado de Stein et al. (1997).

Como pode ser percebido, cada camada possui constituição diferente umas das outras, permitindo características físicas próprias a cada uma delas.

Cada uma das camadas da córnea parece ter características mecânicas próprias, conferidas pelas junções intercelulares e complexos de adesão, como no epitélio, ou elasticidade pequena, como a da camada de Bowman, o que faz com que a córnea anterior se deforme proporcionalmente mais intensamente nos casos de edema de córnea. (...). A membrana de Descemet é bastante elástica e o endotélio corneano é pouco resistente (...) (Muccioli et al., 2006, p. 1).

Ainda segundo Muccioli et al. (2006), a viscoelasticidade da córnea é referenciada pelo termo *histeresis* ou módulo de Young, que é a medida da capacidade da córnea de dissipar energia de forças aplicadas a ela. O olho sempre exerce uma força no sentido interior para exterior, que é a *pressão intra ocular* (também conhecida como PIO, *Intra Ocular Pressure* ou, ainda, IOP). A forma de cúpula da córnea modifica a ação da PIO, convertendo-a em uma força tangencial e absorvendo-a nas fibras do estroma (Ribeiro, 2015).

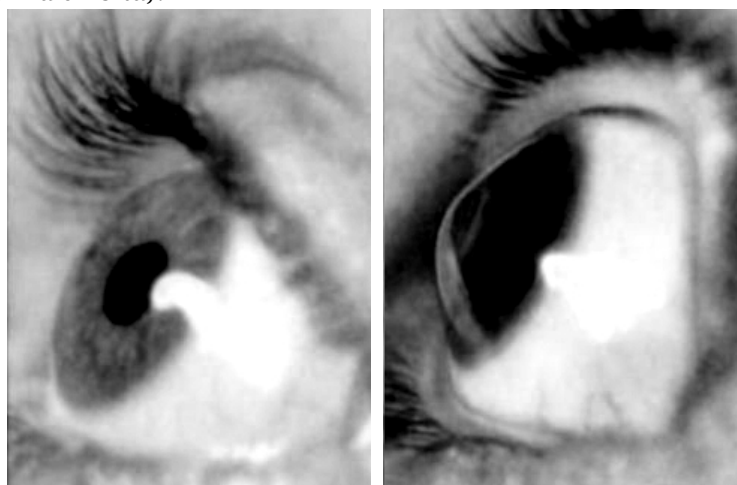
O envelhecimento natural da córnea causa o processo de *crosslink*, isto é, a organização do estroma muda – as fibrilas se tornam mais grossas, alongadas e aderidas, aumentando a rigidez do tecido.

2.7 Ceratocone

Ceratocone é uma doença progressiva, degenerativa, bilateral (se há diagnóstico da doença em um olho, o outro sofre do mesmo mal) e assimétrica (a severidade da doença em cada olho pode ter desenvolvimento diferente), caracterizada pelo afinamento, normalmente, de uma porção central da córnea, além de protusão em formato cônico. Apesar disso, regiões periféricas podem apresentar espessura e curvatura normais (Stein et al., 1997).

Na Figura 2.10, é possível perceber a diferença de curvatura de uma córnea normal (imagem à esquerda) e com ceratocone em estágio avançado (imagem à direita).

Figura 2.10: Perfil de Córneas: normal (imagem à esquerda) e com ceratocone em estágio avançado (imagem à direita).



Fonte: <http://arizonaeyes.net/services/cornea-center/keratoconus/>.

As causas não são claras, havendo associação com origens genéticas (múltiplas ocorrências familiares, síndrome de Down, amaurose congênita de Leber, dentre outras), mas também com alergia ocular. O desenvolvimento da doença ocorre a partir da puberdade, intercalando períodos de estagnação e progressão, prolongando até aproximadamente os 30 a 40 anos de idade, quando se estabiliza (Stein et al., 1997).

A severidade do ceratocone pode ser classificada com a escala de Krumeich, presente na tabela 2.1. Há estudos, como o presente em Romero-Jiménez et al. (2010), que indicam um “5º estado de severidade” que seria a *forma frustra* ou *subclínica*, isto é, o olho possui ceratocone (devido ao diagnóstico da doença no outro olho), mas não há presença de nenhum sintoma que indique isso. Em outros, como em Binder et al. (2005), abordam esse tema como

Tabela 2.1: Classificação de ceratocone, segundo Krumeich

| Severidade (Grau) | Ceratometria Média (D) | Espessura Corneana (μm) | Equivalente Esférico (D) | Cicatriz Corneana |
|-------------------|------------------------|--------------------------------------|--------------------------|-------------------|
| I | <48 | >500 | <-5 | Ausente |
| II | 48 a 53 | 400 a 500 | -5 a -8 | Ausente |
| III | 54 ou 55 | 200 a 400 | >-8 | Ausente |
| IV | >55 | <200 | Não mensurável | Presente |

Fonte: Tabela adaptada de Sinjab, 2012.

sendo um olho que possui uma desordem subclínica com potencial para desenvolvimento de ceratocone. Com grau de severidade I, poucos sintomas são sentidos e procedimentos de investigação clínicos apresentam pouca ou nenhuma diferença em relação a uma córnea normal. Com severidade II, algumas anomalias podem aparecer, mas o diagnóstico muitas vezes é impreciso, ainda confundindo-se com uma córnea normal. A partir do grau III, os sintomas já são mais evidentes e exames clínicos detectam mais facilmente a presença da doença.

A deformação causa progressivo aumento de miopia e astigmatismo irregular, com consequente perda da acuidade e qualidade visual. Para compensar esses sintomas, é recomendado o uso de óculos ou lentes de contatos, nos estágios iniciais. No entanto, devido às mudanças que causam a grande curvatura da córnea em estágios mais avançados, o uso de lentes de contato torna-se impraticável, tanto pelo incômodo, quanto pela possibilidade de provocar ulcerações e edemas, causadas pela sensível queda oxigenação do epitélio (Dome, 2008).

Procedimentos de cirurgias refrativas são contra-indicadas para pessoas com diagnóstico de ceratocone. Nesses procedimentos, a córnea é aplanada pela ação de um *laser* que remove porções de tecido – decomposição fotoablativa – (Manche et al., 1998). Os mais comuns procedimentos são o LASIK (*Laser in Situ Keratomileusis*, ou ainda, em português, Ceratomileusis in Situ a Laser) e o PRK (ou *Photorefractive Keratectomy*, ou ainda, em português, Ceratectomia Fotorrefrativa). O que difere os procedimentos é que, no LASIK, o epitélio e a membrana de Bowman são seccionados para formar uma espécie de *flap* que é levantada, apenas o estroma sofre ablação e o *flap* é recolocado em seu lugar de origem. Já no PRK, o *laser* realiza a erosão na superfície da córnea, incluindo epitélio e membrana de Bowman. Devido à ablação, é possível que esse tipo de procedimento cause uma condição chamada Ectasia pós-LASIK (que também pode ser aplicado ao PRK e outros procedimentos), que possui características compatíveis com o ceratocone (Binder et al., 2005), mesmo em córneas consideradas normais.

Em Rabinowitz (1998), há a explicação de que a deformação corneana ocorre devido à compactação e perda de fibrilas no estroma. Em Sherwin & Brookes (2004), foram realizados estudos (experimentais e bibliográficos), reportando outras modificações estruturais e em todas as camadas da região afetada. O epitélio apresenta, *ex vivo*, perda de células, mas *in vivo*

o mesmo não ocorre, quando também são encontradas dobras na região basal. Na membrana de Bowman, há casos em que esta camada se apresenta descontínua, estando o estroma em contato direto com o epitélio. O estroma apresenta diminuição da quantidade de lamelas e do volume dos proteoglicanos que as separam; a espessura das lamelas permanece estável; acontecem mudanças na orientação das fibrilas e diminuição do volume dos ceratócitos, salienta-se que tais efeitos foram também relatados em [Vellara & Patel \(2015\)](#). Prosseguindo com os estudos de [Sherwin & Brookes \(2004\)](#), tem-se que a membrana de Descemet pode apresentar rupturas, enquanto que no endotélio as células podem ter anormais estruturas escuras, rupturas e perdas celulares. Um problema levantado tanto em [Sherwin & Brookes \(2004\)](#), quanto em [Vellara & Patel \(2015\)](#) é que a maioria das alterações estruturais só puderam ser percebidas em exames *ex vivo*, após o transplante de córnea (*ceratoplastia penetrante*) ou em exames em períodos de estagnação (após o aparecimento de sintomas característicos).

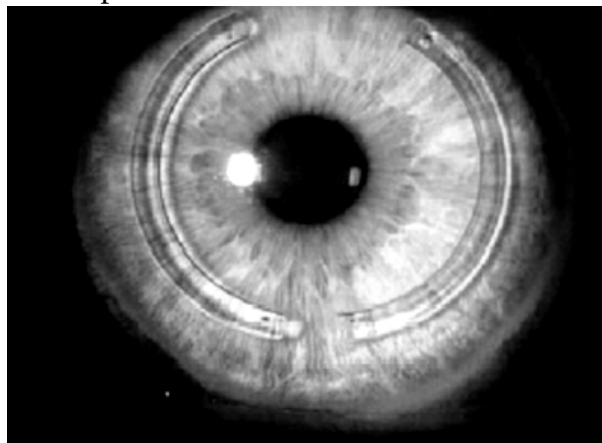
De acordo com a severidade, os sintomas do ceratocone possuem tratamentos que podem retardar e até estagnar o seu desenvolvimento. Em graus I e II, o uso de lentes rígidas ajuda a retardar o desenvolvimento da doença e ajusta a acuidade visual. Ainda nessas fases, é eficiente a aplicação da técnica de *crosslink artificial*, no qual a córnea é irrigada com riboflavina e em seguida exposta a raios ultravioleta tipo A. Isso provoca o aumento do diâmetro e da rigidez das fibrilas de colágeno do estroma, tornando a córnea permanentemente mais rígida e estagnando o desenvolvimento do ceratocone ([Wollensak et al., 2003](#)).

Em graus III e IV um procedimento que apresenta bons resultados é o *anel intraestromal corneano* (ou, em inglês, *Intrastromal Corneal Ring Segment*, ou ainda ICRS) ou *anel de Ferrara*. Nesse procedimento, são inseridos segmentos de anel translúcido no estroma, aplanando e enrijecendo a córnea ([Alio et al., 2014](#)). A aplanação causada por este procedimento modifica a forma cônica para esferoide nas severidades III e IV, mas é problemática com graus I e II, pois a córnea passa de forma esferoide para aplanada, prejudicando a acuidade visual. O resultado desse procedimento pode ser visto na Figura 2.11. Em situações mais graves, pode haver a necessidade de realização de transplante de córnea ou ceratoplastia penetrante, podendo haver complicações, como rejeição do novo tecido, astigmatismo pós-operatório e, mais raramente, reincidência de ceratocone ([Rabinowitz, 1998](#)).

É importante que o diagnóstico de ceratocone seja feito o mais precoce quanto possível, a fim de haver mais possibilidades de tratamento, e que este possa ser menos invasivo e traumático ao paciente. De acordo com as seções 2.5 e 2.6, é esperado que as mudanças estruturais impliquem em alterações da biomecânica. Em [Vellara & Patel, 2015](#) há discussão sobre alguns efeitos, como: a região da córnea afetada pelo ceratocone passa a ter viscoelasticidade irregular, com possível enrijecimento na região central (onde aparece maior stress devido efeito da PIO) e, pela redistribuição do stress, com deformação da região circunvizinha. No entanto, nem de forma qualitativa, nem quantitativa, a biomecânica pode ser considerada uma área de estudo de conhecimentos já devidamente definidos.

Em estudos – como os relatados em [Vellara et al. \(2015\)](#), [Vinciguerra et al. \(2016b\)](#), [Bao](#)

Figura 2.11: Córnea com implante de Anel Intraestromal Corneano em dois segmentos.



Fonte: <http://rosinov.com/catalogo/anillos-intracorneales-ferrara-rings/>.

et al. (2016), dentre outros – há a proposição da análise da biomecânica corneana através de equipamentos com auxílio de computadores, a fim de medir e analisar dados como velocidade corneana, amplitude de deformação e instantes de aplanção, relacionando-os com o diagnóstico de ceratocone.

2.8 Exames para Análise de Biomecânica Corneana

O primeiro aparelho com utilidade para diagnóstico de ceratocone foi o *disco de Plácido*, em que discos concêntricos são projetados na superfície da córnea e a distorção da projeção é usada para avaliar a irregularidade corneana. Com a evolução da tecnologia aplicada aos princípios desse aparelho, tem-se hoje a Videoceratometria Computadorizada (também conhecida como Topografia Corneana), que avalia mais de 8000 pontos projetados com os mesmos tipos de discos, mas processados e analisados por computador (Rodrigues & Dantas, 2001). A projeção gera uma espécie de mapa da superfície da córnea e é um procedimento amplamente usado pela comunidade oftalmológica.

Há também a Tomografia Computadorizada. Nesse exame, uma câmera é rotacionada para capturar o perfil da córnea em 360 graus. Dessa forma, o equipamento captura a curvatura da porção anterior e posterior da córnea. No entanto, tal qual a Topografia, a imagem formada é estática, representando a córnea em repouso e não obtendo informações de sua biomecânica (Kanski, 2003; Rodrigues & Dantas, 2001).

Como estágio I e II do ceratocone interferirem pouco ou em nada na forma da córnea, então a topografia e a tomografia podem não ser suficiente para identificação da doença, nesses estágios. A partir de 2005, começaram a aparecer equipamentos comercialmente viáveis (outros equipamentos foram criados e testados, mas não tiveram viabilidade comercial) para avaliar a biomecânica da córnea através de exames não invasivos *in vivo* (Glass et al., 2008).

O primeiro foi o *ORA* (*Ocular Response Analyzer*, ou em português, Analisador de Resposta

Ocular). Esse aparelho usa um jato de ar direcionado durante o período de 20ms, para (com a ajuda de um feixe e um sensor de infravermelho) perturbar a superfície da córnea no sentido exterior-interior e capturar sua dinâmica. Durante o exame, a córnea se curva para dentro do olho, onde, sobre esse movimento, são feitas diversas medições (chamadas parâmetros) com a finalidade de caracterizar a biomecânica corneana (Glass et al., 2008).

O exame é feito em duas fases: progressiva e regressiva. Na fase progressiva, o jato de ar é soprado, a córnea começa a se dobrar para o interior do olho, se aplanar e a inverte a concavidade na região atingida. Na fase regressiva, o jato de ar é interrompido, daí a córnea passa da concavidade invertida para a forma plana e novamente para o seu estado normal (Glass et al., 2008).

Com os dados, o equipamento é capaz de calcular a PIO e outros 37 parâmetros (Bao et al., 2016). A quantidade de parâmetros varia de acordo com a versão do *software* usado para exportar os dados do aparelho para o computador.

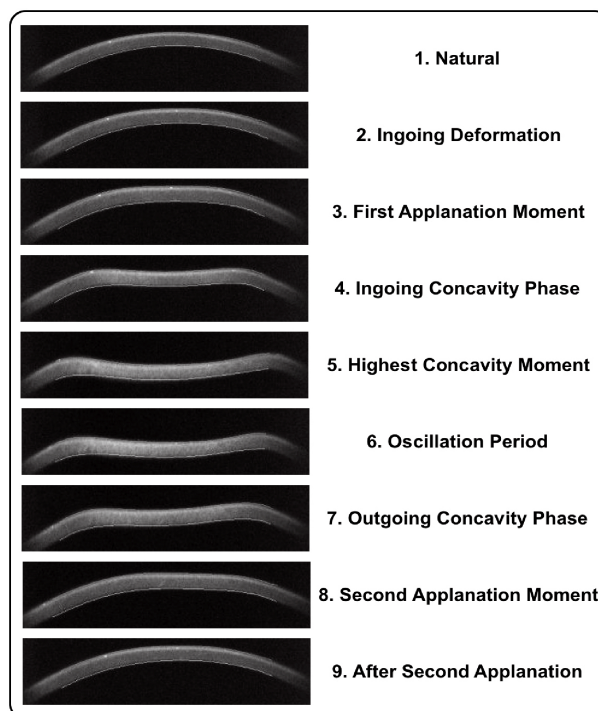
Outro aparelho que tem sido usado para avaliar a biomecânica da córnea é o *Corvis ST* (*Corneal Visualization Scheimpflug Technology*, ou CST, ou, em português, Tecnologia Scheimpflug de Visualização Corneana, ou apenas Corvis). Tal qual o ORA, o Corvis usa um jato de ar para perturbar a córnea, mas usa uma câmera UHS (*Ultra-High-Speed*, ou, em português, Ultra Alta Velocidade) Scheimpflug, com iluminação de uma fonte de LED Azul, para captura de 140 imagens (ou quadros, ou *frames*) planas de 576×200 pixels, no perfil horizontal, cobrindo área de 8,5mm, durante 32,11ms (Ambrósio et al., 2013).

A câmera Scheimpflug permite o foco em um plano diferente do paralelo às lentes de captura limitando a área focada, isto é, o plano de captura de imagens pode ser qualquer um, contanto que a lente possa ser posicionada a focar na área alvo e o restante da imagem possa ficar fora de foco (Rebordao & Cabral, 2008). Esse mesmo procedimento permite que a câmera capture imagens em plano de corte no sentido frontal-traseiro e superior-inferior da córnea, captando em uma mesma imagem as porções anterior e posterior da córnea (Dubbelman et al., 2006). No entanto, por causa de suas características refrativas (como exposto na seção 2.6), a imagem da região posterior pode apresentar distorções.

O decorrer do exame do Corvis é similar ao do ORA e pode ser constatado na Figura 2.12:

1. a córnea inicia em sua posição de repouso
2. o jato de ar é aplicado e a córnea é defletida para o interior do olho
3. a córnea continua deformando e ocorre a primeira aplanção
4. a deformação prossegue
5. até atingir a máxima concavidade
6. o jato de ar permanece atuando e a córnea oscila

Figura 2.12: Dinâmica da córnea durante exame com Corvis ST



Fonte: [Ambrósio et al., 2013](#).

7. o jato de ar para de ser acionado e a córnea começa a voltar a sua posição de repouso
8. acontece a segunda aplanção
9. a córnea retorna ao seu estado normal

O Corvis diferencia-se do ORA, pois a origem dos dados do ORA é unidimensional e são derivados da captura feita por um único sensor ([Glass et al., 2008](#)), enquanto que os do Corvis são originados do processamento da sucessão de imagens ([Ambrósio et al., 2013](#)). O jato de ar usado pelo ORA possui variação de pressão (após a primeira aplanção, o aparelho calibra a pressão de ar para refletir a PIO medida) ([Machado et al., 2011](#)), já o Corvis, usa jato de ar em pressão constante. À exceção da PIO e do tempo para primeira aplanção, nenhum outro parâmetro dos aparelhos podem ser comparados com os do outro, devido usarem métodos diferentes.

O Corvis também utiliza um *software* em computador para analisar os dados internos (as imagens) e gerar parâmetros que quantificam a biomecânica. Durante a captação dos dados para pesquisa fazendo uso do *software* do equipamento, foi percebido que de versão para versão do aplicativo pode haver variação na precisão de cálculo dos parâmetros.

2.9 Exames de Tomógrafos e seus Dados

Seguindo a classificação de Krumeich, quando o ceratocone encontra-se nos estágios I, II ou frustro, os topógrafos podem não apresentar informações suficientes para diagnóstico de ceratocone. Por outro lado, os tomógrafos já podem apresentar indícios da doença (Vinciguerra et al., 2016a), mesmo que a compreensão da biomecânica ainda seja um fator que tem prejudicado o uso mais intensivo desses equipamentos.

As empresas fabricantes dos tomógrafos e seus *softwares* têm feito atualizações em seus produtos. Diversos parâmetros sobre a biomecânica são introduzidos em cada atualização, mas o significado deles ainda não está claro. Outro problema é a incompatibilidade de parâmetros. Por possuírem diferentes métodos para geração de dados, os exames dos equipamentos não são compatíveis ou comparáveis, impedindo uma melhor avaliação sobre qual o mais apropriado (Bao et al., 2016).

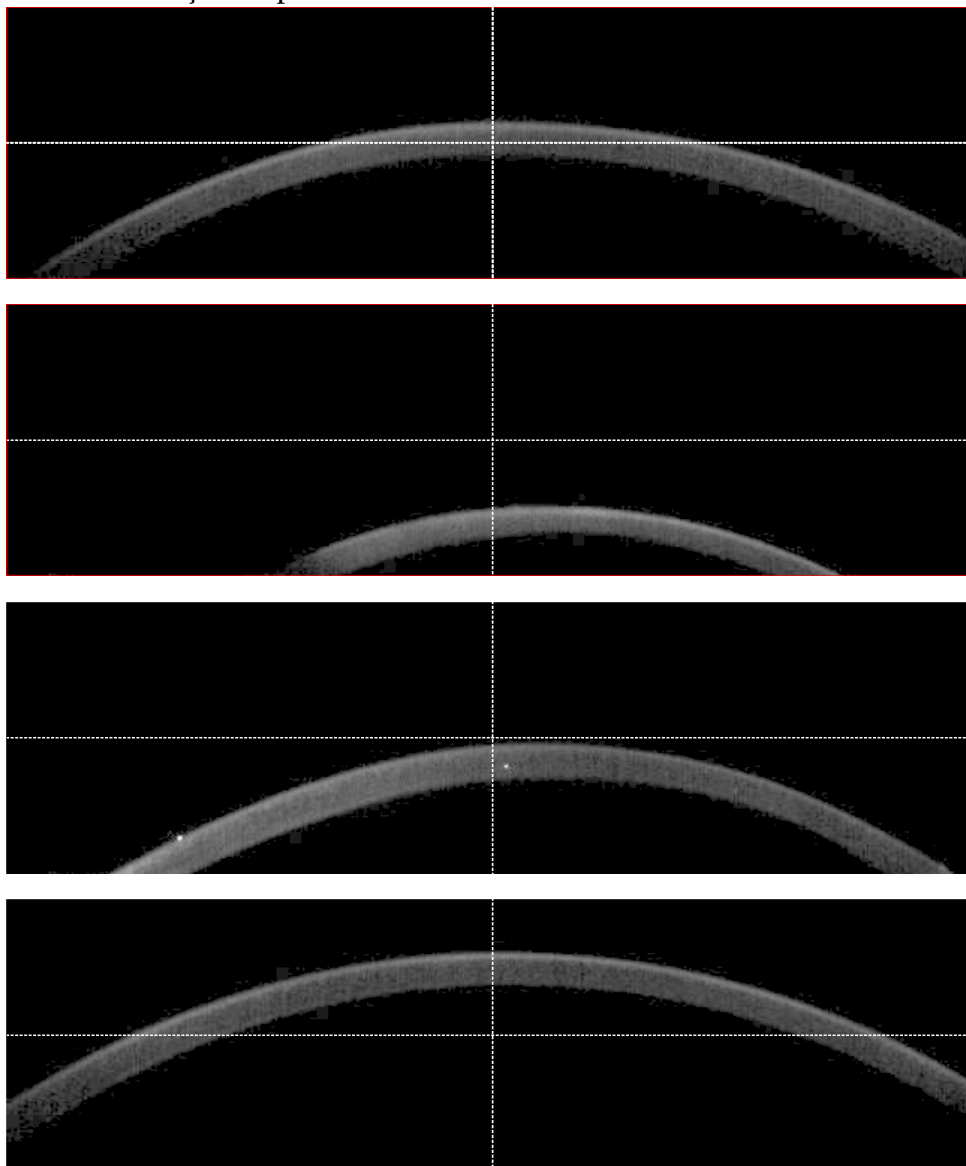
Estudos como os de Tian et al. (2014) indicam que há correlação entre esses dados e o diagnóstico de ceratocone, principalmente aqueles que dizem respeito a amplitude de deformação durante os exames. Resultados compatíveis foram encontrados em Ye et al. (2015), no entanto, os procedimentos metodológicos prejudicam a confiabilidade desse estudo que apresenta apenas 24 pacientes e não há informação sobre os graus de ceratocone abrangidos no estudo. Já em Ali et al. (2014), foi demonstrado que esses exames são consistentes na repetibilidade, isto é, a repetição de um exame ao longo do tempo (várias vezes por dia ou em dias diferentes) gera os dados compatíveis, indicando que não possuem variação aleatória. Em Vellara & Patel (2015), os resultados apontam para que a viscoelasticidade pode sofrer variação ao longo do dia, mas ainda há consistência entre os dados gerados.

Um problema a ser destacado é que os equipamentos são passíveis à imprecisão da captura dos dados. Mesmo que tenham procedimentos automáticos para detecção do posicionamento e alinhamento do olho, que disparam automaticamente o início do exame, ainda assim é possível que o olho não se encontre em uma posição ótima, como exibido na Figura 2.13 (onde foram adicionadas linhas pontilhadas para identificação do centro das imagens). Isso gera dados que podem ser errados. O Corvis, por exemplo, pode gerar imagens que possuam outros objetos que não a córnea, além de apresentar fochos de luz anormais e córnea descentralizada, que podem comprometer o exame (Vellara et al., 2015), como visto na Figura 2.14. O ORA, por possuir análise unidimensional, pode realizar o exame fora da centralização com o ápice da curvatura corneal e não disponibiliza indicadores para destaque desse problema.

O jato de ar de pressão variável do ORA pode ainda trazer aberrações para o diagnóstico, como a insensibilidade aos efeitos de uso de *crosslink* e a falta de relacionamento entre parâmetros ditos relativos a viscoelasticidade e o módulo de Young (Shetty et al., 2015).

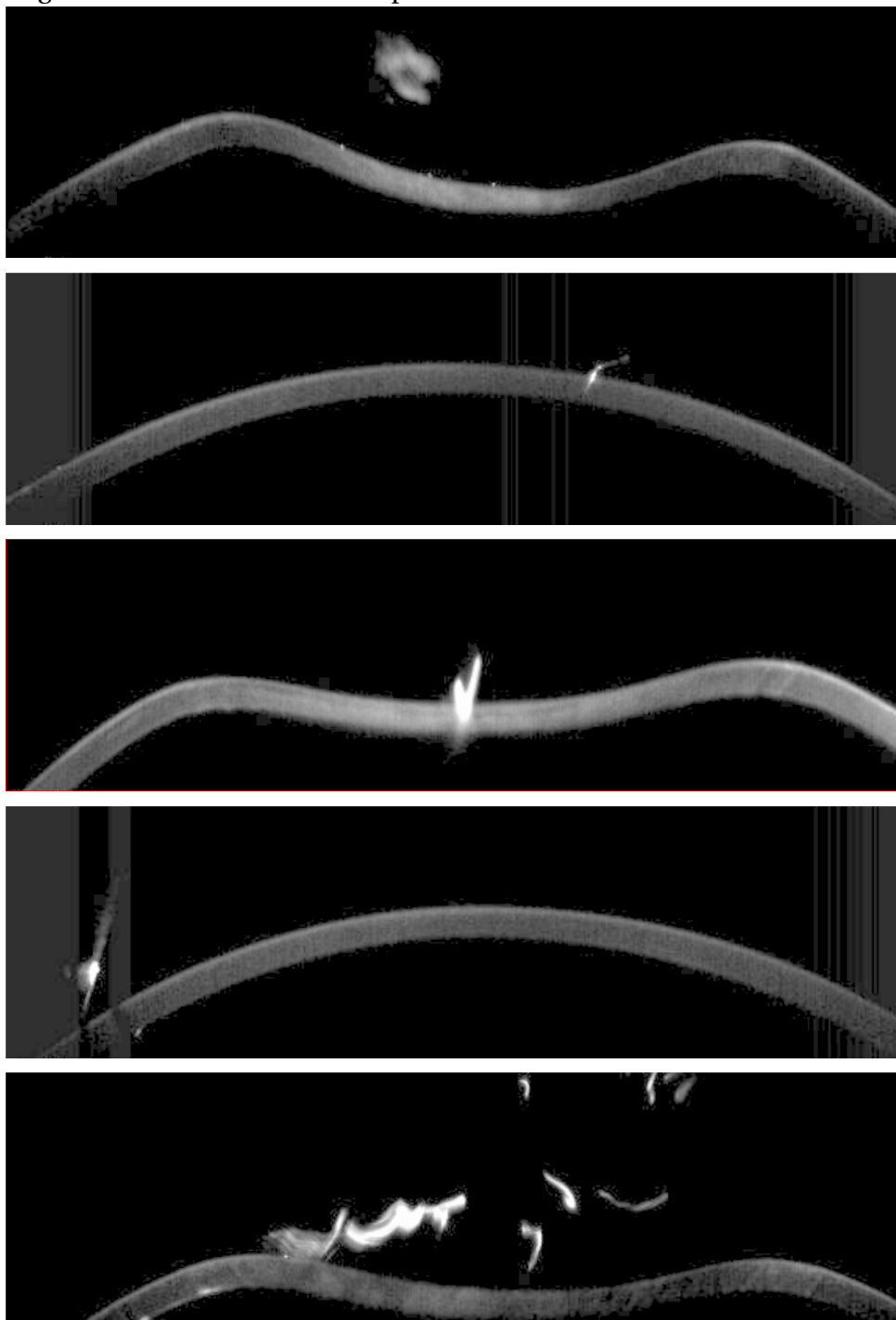
Os dados também não podem ser usados separadamente para diagnóstico. Isso tanto diz respeito a não capacidade dos parâmetros para uma completa separação e classificação

Figura 2.13: Variação de posicionamento do olho em diferentes exames do Corvis



Fonte: O autor.

Figura 2.14: Áreas anômalas capturadas em diferentes exames do Corvis



Fonte: O autor.

dos estágios da doença, quanto que alguns deles ainda podem ser afetados por fatores não englobados pelos aparelhos, como a origem étnica (que tem correlação com a pressão intraocular) ou a idade (que influencia o depósito de gordura na parte posterior do olho e que atua sobre dissipação de forças aplicadas em sua parte anterior) (Vinciguerra et al., 2016b).

Nesse último ponto, o Corvis possui uma vantagem sobre o ORA, que é a possibilidade de analisar o movimento do olho. Durante exame, o jato de ar aplicado pelo Corvis atua apenas sobre uma fração da córnea, mas ela se move por inteiro. A região afetada é apenas a central, mas as imagens geradas cobrem quase a córnea inteira. Então, ao analisar a periferia da córnea, pode-se inferir o movimento do olho (Koprowski & Ambrósio, 2015).

O Corvis ainda possibilita a realização de outras análises sobre seus exames. Isto porque, além de parâmetros já valorados, esse equipamento ainda disponibiliza o conjunto de imagens adquiridas durante cada exame. Em Koprowski et al. (2014) e Koprowski & Ambrósio (2015), está demonstrado que é possível fazer esse procedimento.

Em Koprowski et al. (2014) são indicados os seguintes procedimentos para processamento de um exame:

1. Exportação das imagens a partir do *software* do fabricante
2. Aplicação de filtros de imagens:
 - (a) Mediana 3D $3 \times 3 \times 3$ (aplicado a todo o exame)
 - (b) Canny
 - (c) Dilatação
 - (d) Erosão
3. Conversão das imagens em vetores. Cada imagem corresponde a um vetor unidimensional que possui as distâncias do topo da imagem até o primeiro pixel não preto (topo da córnea)

De posse desses vetores, é possível fazer algumas operações. Uma delas é a normalização de posição, que é a subtração de todos os vetores do exame pelo primeiro deles. Isso faz com que a posição relativa de cada coluna seja processada em relação a seu valor inicial, evitando problemas de posicionamento do olho. Outra operação é a detecção do movimento do olho. Isso é feito pelo cálculo da média dos valores das extremidades de um vetor e subtraindo de todos os valores do mesmo.

Apesar desses procedimentos, não há registro que Koprowski et al. tenha aprofundado o estudo da dinâmica da córnea como um todo. Em vez disso, houve o foco apenas na coluna de dados que apresentou maior deslocamento, aproximando seus valores de um parâmetro já gerado pelo Corvis, a *amplitude de deformação*.

O presente capítulo elencou as motivações oftalmológicas para estudo da biomecânica da córnea. Também foram apresentados conceitos da área de aprendizado de máquina e segmentação de imagem usados para processamento de imagens com intuito de reconhecer padrões e propiciar meios para implementação de visão computacional. No próximo capítulo, estão presentes os experimentos realizados para avaliar a viabilidade de diagnóstico de ceratocone com uso de exames para captação da biomecânica corneal.

3

METODOLOGIA PARA DIAGNÓSTICO DE CERATOCONE

ESTE capítulo apresentará os procedimentos metodológicos do corrente trabalho, descrevendo os critérios de avaliação dos modelos aprendizagem de máquina, as bases de dados de exames (Seção 3.2) e dois conjuntos de modelos implementados (Seção 3.3), a fim de, a partir das imagens extraídas do *software* do aparelho, analisar a biomecânica da córnea e realizar diagnóstico de ceratocone.

3.1 Critérios de Avaliação de Modelos de Aprendizagem de Máquina

Segundo Amancio et al. (2014) e Witten & Frank (2005), não há uma forma única de avaliação de modelos de aprendizagem de máquina ou classificadores. A seleção de um *ranking* sobre os modelos depende do que se deseja enfatizar na avaliação.

Ao fazer classificações para modelos dicotômicos, como de diagnóstico (presença ou ausência de doença), há quatro possibilidades de classificação: verdadeiro positivo (correto diagnóstico de presença da doença), falso positivo (incorreto diagnóstico de presença da doença), verdadeiro negativo (correto diagnóstico de ausência de doença) e falso negativo (incorreto diagnóstico de ausência da doença). Isso pode ser sintetizado através de uma *matriz de confusão*, como a da Tabela 3.1 (Witten & Frank, 2005).

Classificadores de diagnóstico, por serem voltados para a área médica, podem ser qualificados em testes de *sensibilidade* e *especificidade*. A sensibilidade diz respeito à proporção entre predições verdadeiras positivas, dado o total classificações positivas. Enquanto que especificidade é a proporção de predições negativas para o total de classificações negativas. A multiplicação desses testes ($rp = \text{sensibilidade} \times \text{especificidade}$) – método *Precision-Recall* – é uma medida de qualidade de classificação, pois valores de rp próximos a 0 indicam

Tabela 3.1: Exemplo de matriz de confusão

| | | Predição | |
|-----------------------|----------|----------|----------|
| | | Positivo | Negativo |
| Classificação Correta | Positivo | VP | FN |
| | Negativo | FP | VN |

Fonte: o autor

baixa capacidade preditiva – pelo menos um entre sensibilidade e especificidade é igual a 0 –, enquanto que próximos a 1, alta capacidade preditiva em ambos testes (Witten & Frank, 2005).

Assim sendo, esse método foi utilizado como critério para avaliação dos modelos de aprendizagem deste trabalho, sendo utilizado o termo *escore* como sinônimo ao valor de rp .

3.2 Descrição dos Dados de Exames

Este estudo foi realizado na Universidade Federal de Alagoas com apreciação do CEP/CONEP-UFAL, obtendo aprovação para sua execução (CAAE: 55324516.6.0000.5013). Foram usados prontuários médicos fornecidos por grupos colaboradores de duas origens. A primeira base de dados, a mesma usada em Dantas (2017), foi originada em atendimentos realizados em clínica privada de colaboradores (R.A.J.), no Rio de Janeiro, no período entre 2011 e 2014, sendo submetidos à realização do exame do Corvis (Versão 3.01). Nesse grupo, foram excluídos os pacientes que tinham idade inferior a 18 anos, possuíam histórico cirúrgico ocular, enfermidades oculares além do ceratocone, fizeram uso crônico de medicação tópica, tinham cicatrizes (ou opacidades) corneanas ou fizeram uso de lentes de contato até 72 horas antes do exame.

O segundo conjunto é o mesmo usado em Ambrósio et al. (2017). A origem de seus dados são diagnósticos feitos em clínicas em Milão (Itália) e no Rio de Janeiro (Brasil). Há autores em comum com aqueles de Ambrósio et al. (2013). Os critérios de exclusão de exames foram os mesmos da primeira base. Foi identificado que essa base não possui a classificação dos graus de ceratocone. Os pacientes com ceratocone tiveram seus dados exportados como anônimos. Alguns integrantes do grupo de estudos de Milão são também provedores do primeiro grupo de dados, havendo a possibilidade de que alguns pacientes estejam presentes em ambos conjuntos de dados.

Assim, os dados podem ser sintetizados da seguinte forma:

Base Brasil são 226 exames com diagnóstico normal e 222 com diagnóstico de ceratocone (totalizando 448)

Base Europa são 485 exames com diagnóstico normal e 207 com diagnóstico de ceratocone (totalizando 692)

Em momentos preliminares, os exames com diagnóstico de ceratocone com graus 3 e 4 foram excluídos, pois são situações de fácil diagnóstico e como a córnea possui grande flacidez nesses estágios da doença, foi teorizado que esses exames poderiam trazer disparidades que atrapalhassem o processo de aprendizagem. No entanto, a análise dos resultados indicaram que o acréscimo ou a remoção desses exames do conjunto de treinamento não afeta a qualidade do diagnóstico. Por causa da pequena quantidade de exemplos (apenas 29 na base do Brasil), foi, então, estipulado que esses exames não seriam excluídos da base de treinamento.

É comum encontrar na literatura pesquisada (como em [Glass et al., 2008](#)), que a pressão intraocular possui correlação com o diagnóstico de ceratocone. No entanto, ainda no início da pesquisa, foi percebido que o uso dessa informação em conjunto com as técnicas usadas nessa pesquisa pioraram a acurácia do diagnóstico.

3.3 Metodologias de Processamento

Como o problema de diagnóstico de ceratocone é relativo à classificação, a pesquisa foi voltada para procedimentos com aprendizado de máquina nesse sentido. Ao mesmo tempo houveram indicações em várias referências que a biomecânica pode ser uma resposta a isso, visto que, clinicamente, o diagnóstico é difícil. Como foi bem frisado em [Wang et al. \(2016\)](#), [Vinciguerra et al. \(2016b\)](#), [Vinciguerra et al. \(2016a\)](#), [Vellara & Patel \(2015\)](#), [Vellara et al. \(2015\)](#); dentre outros.

Os primeiros estudos ocorreram usando os dados gerados pelo *software* do fabricante do Corvis. Devido ao fato que nenhuma das técnicas escolhidas apresentou score maior que 0 – com todas as predições incorretas para presença de ceratocone – em mais de 184 configurações diferentes (com uso de redes neurais, SVM, regressão logística e *deep learning*), a atenção foi voltada aos dados. Percebeu-se que os mesmos foram exportados de diferentes versões do aplicativo e que apresentavam variação na quantidade de parâmetros dos exames. Os exames foram, então, reexportados usando uma única versão (1.3b1716). Após a exportação, os dados foram comparados com os previamente existentes e se verificou que, de uma versão para outra, o *software* gerava dados diferentes. Assim, os dados numéricos de todos os exames foram descartados, sendo usadas apenas as imagens geradas pelo equipamento.

Devido as capacidades de classificação de *deep learning*, como o exposto em [Zeiler & Fergus \(2013\)](#), [Wu et al. \(2016\)](#), [Szegedy et al. \(2015\)](#) e outras publicações, essa foi a primeira alternativa de solução, com aplicação dos algoritmos diretamente sobre as sequências de imagens. Como essa técnica faz reconhecimento automático de características estáticas (em cada imagem) e dinâmicas (ao longo da sequência de quadros), foi esperado que se encontrasse uma “assinatura” que diferenciava córneas saudáveis, daquelas com ceratocone.

Mesmo com diferentes configurações, devido a muita exigência do *hardware* dos compu-

tadores onde foram aplicados, esse método se mostrou tecnicamente inviável. Vários dos *frameworks* de DL apontam a necessidade da execução dos aplicativos em equipamentos que tenham GPUs para acelerar o processamento. No entanto, não houve disponibilidade de computadores com GPU compatível com nenhum dos *frameworks* testados - Theano (LIS, 2016), Keras (Chollet, 2015) e Caffe (Jia, 2014). O mais rápido usado, Theano, demorou 18 horas para processar um único exame, enquanto que a necessidade seria para todos os 448 exames múltiplas vezes, sendo que a execução sobre todo o conjunto apenas uma única vez demoraria aproximadamente 336 dias, tornado esse procedimento inviável.

A partir das discussões em Koprowski et al. (2014), Koprowski & Ambrósio (2015) e Ribeiro (2015), foi adotada a extração da dinâmica a partir da segmentação das imagens. Para tal, foi utilizado OpenCV (OpenCV Development, 2017) com implementação em C++, sendo a linguagem de programação que apresentou melhor desempenho.

Como indicado em Spirkovska (1993), esse processo foi quebrado em duas fases: Pré-processamento e Processamento. Além dessas fases, houve a necessidade de identificação se o posicionamento do olho perante o exame influenciava os resultados e a validação dos dados. Após esses passos, os dados foram reorganizados e submetidos às técnicas de *machine learning*.

3.3.1 Processamento das Imagens

O início da pesquisa foi baseado nos estudos realizados em Koprowski et al. (2014) e Koprowski & Ambrósio (2015) para identificação das características do movimento da parte anterior da córnea.

Primeiramente o exame tem todas suas imagens carregadas em um único vetor de imagens. Nesse momento, para evitar procedimentos distintos para os olhos esquerdos e direitos, além de excluir o viés de diagnóstico com lateralidade, quando o exame é de olho esquerdo, as imagens são espelhadas horizontalmente em seu carregamento.

Depois é usado o filtro de mediana 3D, em formato $3 \times 3 \times 3$. Os conjuntos de bibliotecas que compõem a OpenCV e as distribuições da linguagem de programação C++ não possuem implementação desse filtro no formato de 3 dimensões, isso originou a necessidade de sua implementação manualmente. O tamanho do filtro foi escolhido para balancear a necessidade de remoção de ruído com manutenção das características da córnea. Com valores menores que 3, não há tanta remoção do ruído. Com valores maiores, há perda na definição da superfície corneal.

Para melhorar a diferenciação da córnea em relação ao fundo da imagem e de ruído, foi aplicado *rolling guidance* (Zhang et al., 2014), (com variância de cor em até 19,61%). O uso de variâncias de cores menores causam o destaque de ruído e, mesmo tornando a identificação do contorno da córnea melhor, ainda destaca indesejadas cavidades causadas por falhas na captura das imagens. Com variâncias maiores acontece perda de curvatura da córnea por

incluir ruído nas áreas da superfície da córnea.

Em seguida, é aplicado Canny (com limiars em 3,92 e 11,75%) a cada uma das imagens. O limite superior ajuda a remover o ruído causado por superexposições, enquanto o limite inferior remove ruídos provindos da compactação da imagem que podem elevar a cor dos pixels do fundo.

Foi identificado que nesse passo é possível excluir a maior quantidade de objetos da imagem que não sejam a córnea (como os vistos na Figura 2.14). Isso é conseguido com a aplicação de bibliotecas nativas da OpenCV para encontrar áreas na imagem (*findContours* e *contourArea*), apagando todos objetos encontrados e redesenhando apenas a maior área (com *drawContours*). Com o uso da função *findContours*, o aplicativo delimita os objetos presentes na imagem que possam identificar contornos e os transforma em objetos do tipo *Area*. O maior objeto que deve estar presente é a córnea, logo, é selecionado apenas o objeto de maior área. A área da figura é, então, coberta com preto em todas as posições e apenas a maior área é pintada em branco.

Para diminuição de imperfeições como manchas diretamente sobre a superfície da córnea, foram aplicadas dilatação e erosão com tamanho 3, em cada imagem. Esse valor de configuração representa metade da tolerância a descontinuidade na superfície da córnea para ainda considerá-la válida. Se uma região possuir descontinuidade de até 5 pixels, a dilatação fará com que as duas extremidades da descontinuidade se unam e, ao aplicar a erosão, a região se permanecerá contínua.

Dos passos indicados, a execução do *rolling guidance* e da remoção de objetos, que não a córnea, não estão presentes nos estudos de Koprowski. Essas alterações ocorreram por causa da identificação da melhoria do score em até 0,10 em todos os modelos em estudo. No entanto, essa mesma sequência de filtros causou descarte de muitos exames por problemas nas fases seguintes do processamento das imagens. Isso ocorreu em mais de 20 configurações diferentes dos filtros, com 1200 configurações de técnicas de aprendizagem de máquina.

Assim, o pré-processamento foi modificado. Dentre os passos que prejudicaram a segmentação, o uso de *rolling guidance* e Canny não auxiliaram o processo de identificação da parte posterior da córnea, pois as imagens com muito ruído em tons próximo ao preto causavam diminuição da área da córnea, enquanto que ruídos em tonalidade próxima ao branco aumentavam muito a mesma área. Foram, então, removidos e substituídos por um conjunto com iterações sobre filtros de desfoque gaussiano e limiar adaptativo, seguidos dos procedimentos de dilatação e erosão.

A cada iteração se tenta aplicar todos os filtros e realizar o processamento dos dados. Caso ocorra algum erro de segmentação ou validação dos dados, os parâmetros do limiar adaptativo são modificados, adaptando a segmentação à variação na luminância das imagens.

O método de remoção dos objetos que não a córnea também foi modificado. Após a erosão, há a possibilidade que grandes objetos compostos por pontos pertencentes a logo marca do fabricante, identificação do *frame* e parte da íris ou outros na parte de cima da

imagem, apresentem área maior que a da córnea. Passou-se a utilizar uma ponderação entre a maior área, a maior largura e a maior altura de objeto para identificar o melhor candidato à córnea. Tal ponderação passou a ser executada após a aplicação de erosão.

3.3.2 Pré-Processamento dos Dados

Ainda seguindo os estudos realizados em [Koprowski & Ambrósio \(2015\)](#), foi, inicialmente, desenvolvido um conjunto de algoritmos e programas para extração da caracterização do movimento apenas da face anterior da córnea, iniciando com a conversão de cada imagem em um vetor, como descrito na Seção 2.9.

A partir desses vetores foram criados conjuntos de vetores com as seguintes características:

Inclinação da córnea calculada com a primeira derivada discreta em cada vetor

Deformação da córnea calculada com a segunda derivada discreta em cada vetor

Velocidade corneal tendo em vista que cada vetor da sequência de vetores corresponde a um instante de tempo e que seus valores são posições no espaço, a velocidade é calculada como a primeira derivada discreta no tempo (subtraindo a mesma posição em vetores consecutivos)

Aceleração corneal definida como a segunda derivada discreta no tempo dos vetores de representação da córnea

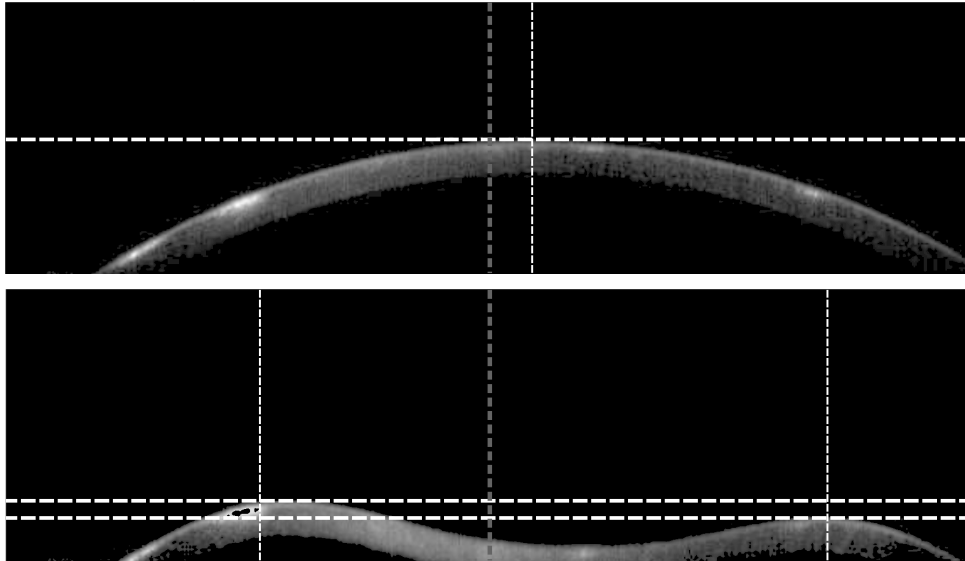
Nos estudos apresentados em [Koprowski et al. \(2014\)](#) e [Koprowski & Ambrósio \(2015\)](#), os vetores que representam o exame da córnea são usados diretamente, não levando em conta o posicionamento do olho perante o jato de ar.

Foram adicionados, então, procedimentos para adicionar esse fator e verificar se esse posicionamento influencia os resultados ao aplicar os dados aos métodos de aprendizagem. Para tanto, foram definidos três marcadores no exame: o centro da córnea, o limite mais à esquerda do pico esquerdo e o limite mais à direita do pico direito. Esses picos laterais são as porções da córnea que apresentam pouco movimento vertical devido ao exame, estando nos limites entre a região da córnea que sofre distorção com o jato de ar e a região que não sofre. A maior parte do deslocamento que se percebe à esquerda do pico esquerdo e à direita do pico direito é exatamente a deflexão do olho devido às forças não dissipadas pela córnea.

Os marcadores estão assinalados na Figura 3.1. Na primeira imagem, é possível perceber que o centro da imagem (marcado com uma linha vertical cinza) não é coincidente com o centro da córnea (linha vertical branca). Na segunda, imagem é possível perceber que a distorção da córnea pode ser irregular, tendo picos em alturas diferentes (distanciamento entre linhas horizontais brancas).

O centro da córnea pode ser encontrado no primeiro *frame* do exame, correspondendo ao índice médio das posições do vetor que possuem menor valor.

Figura 3.1: Identificação de marcadores no exame



Fonte: O autor.

$$Centro(X) = \frac{j-i}{2} \quad \{i, j | i \leq j, (x_i = x_j), (\forall h < i | x_h > x_i), (\forall k > j | x_k > x_j)\} \quad (3.1)$$

Já o marcador de limite esquerdo do pico esquerdo deve ser encontrado com o seguinte procedimento:

```

1 Definir o valor 1000 em "e"
2 Definir 0 em "w"
3 Para cada frame "t" do exame
4     Definir 1000 em "y"
5 Percorrer todos os "i" dados da primeira metade de X
6     Se X[i] for menor que "y"
7         Definir o valor X[i] em "y"
8         Definir "i" em "x"
9 Se "y" for maior que "w"
10     Definir "y" em "w"
11     Definir "x" em "e"
12 Definir pico esquerdo como "e"

```

Podendo ser expresso por

$$PicoEsquerdo(X) = i \quad \{\forall t, i | (x_{i,t-1} < x_{i,t} \geq x_{i,t+1}), (\forall h < i | x_{h,t} > x_{i,t})\} \quad (3.2)$$

De modo análogo, o lado direito do pico direito pode ser dado por

$$PicoDireito(X) = j \quad \{\forall t, j | (x_{j,t-1} < x_{j,t} \geq x_{j,t+1}), (\forall k > j | x_{j,t} < x_{k,t})\} \quad (3.3)$$

Todos os vetores são, então, normalizados em posição, como indicado na Seção 2.9.

Assim, pode-se evidenciar três tipos de centros: o centro da córnea, centro do movimento e centro do exame. Este último é apenas o centro da imagem, correspondendo à coluna 288. Enquanto que o centro do movimento é o ponto médio entre os extremos dos picos previamente encontrados. Apenas um dos três é usado por treinamento. Durante esse estudo o centro usado foi chamado de *centro base*.

Em análises preliminares, o uso dos centros da córnea e do movimento tiveram resultados pouco inferiores aos encontrados usando o centro da imagem (escore variou entre 0,02 e 0,05 nos mesmos modelos). Por terem um custo computacional mais alto, esses procedimentos foram descartados.

Nos estudos apresentados em [Koprowski et al. \(2014\)](#) e [Koprowski & Ambrósio \(2015\)](#), para se ter apenas dados relativos ao efeito do jato de ar, a deflexão do olho é removida. Para tanto, o movimento do olho é interpretado como o deslocamento vertical sofrido pelos pontos extremos das imagens da córnea (*frame a frame*). É calculado, então, a média do movimento desses pontos e esse valor é subtraído de todos os valores que representam o corrente *frame*.

Mas, como visto na Figura 3.1, a deflexão pode ocorrer de forma irregular e existe a possibilidade do movimento do olho ser propagado de forma desigual, aparecendo maior movimento em uma das laterais da córnea.

O movimento da i -ésima posição no vetor – $Pos(x_i)$ – deve ser proporcional aos movimentos das laterais da córnea. Para remover o movimento do olho, é feita a média ponderada aos movimentos dos picos.

$$Pos(x_i) = x_i - \frac{x_{PicoDireito(X)}(PicoDireito(X) - i) - x_{PicoEsquerdo(X)}(i - PicoEsquerdo(X))}{2} \quad (3.4)$$

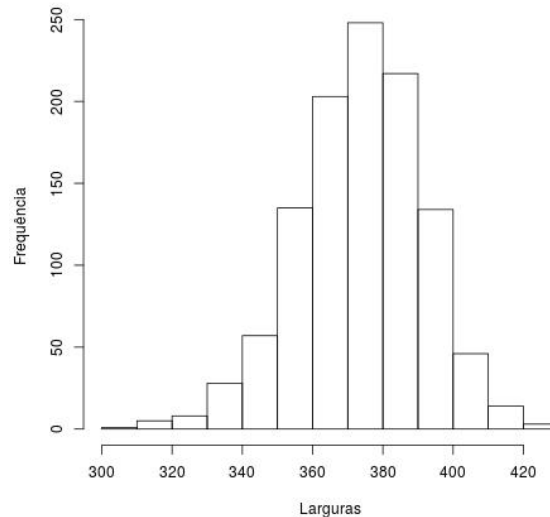
De posse da localização dos extremos dos picos laterais, foi possível calcular a largura da região da córnea que sofre deflexão com o exame. A distribuição desses valores pode ser visto na Figura 3.2.

Devido à grande maioria dos exames possuir região de deflexão com largura de até 400 pixels, essa largura foi definida como o limite lateral do movimento apenas da córnea. Avaliações preliminares demonstraram que esse método causou imprecisão nos cálculos, pois durante a fase de treinamento, o movimento do olho interferiu na qualidade dos dados de vários exames e o escore do diagnóstico foi baixo (próxima a 0,4900).

Com isso, houve a modificação do modo de identificação do movimento do olho. Ao invés do apontamento de locais estimativos de onde a córnea não se move por causa do exame, foi adotado um novo método baseado na inclinação da região periférica da córnea.

Após a identificação dos extremos dos picos laterais, a superfície da córnea é percorrida do pico até o extremo da córnea mais próximo. Nesse percurso, são comparadas as derivadas discretas de colunas com as suas correspondentes na primeira imagem, ou seja, a procura

Figura 3.2: Distribuição de valores largura de região de deflexão de córnea das bases de exames do Corvis ST



Fonte: O autor.

por pontos com a mesma inclinação. Na identificação da primeira sequência de 10 colunas consecutivas cujas derivadas são compatíveis, essa região é identificada como estática perante o exame e a décima coluna é identificada como ponto estático.

É preciso frisar o termo “compatíveis”, pois, como a captura de imagens pode ser imprecisa, a definição dos valores das derivadas como iguais poderia não encontrar córneas válidas. Para evitar isso, é dado o relaxamento na precisão e as derivadas podem variar em até uma unidade (acima ou abaixo).

O cálculo do movimento do olho no ponto i passa então para:

$$x_i = x_i - \frac{x_{ExtremoDir(X)}(ExtremoDir(X) - i) - x_{ExtremoEsq(X)}(i - ExtremoEsq(X))}{2} \quad (3.5)$$

Onde as funções $ExtremoDir$ e $ExtremoEsq$, identificam, respectivamente, o pontos limites à direita e à esquerda que não apresentam movimento devido deflexão da córnea.

Como o volume de dados para aprendizado ainda era muito grande, foram definidas “janelas de corte” dos dados, a fim de se ter a maior quantidade possível de dados, sem comprometimento da capacidade computacional disponível.

O centro base é usado para “recortar” todos os vetores do exame, delimitando-os para os l itens à esquerda até os l itens à direita desse centro. Como a maioria dos exames possuem dados de movimento até a largura de 400 pixels, foi parametrizado que $l = 200$, obtendo-se vetores de 400 dimensões. O aparelho cobre 8,5mm da córnea em 576 pixels, no entanto, a área que é deslocada no exame é de no máximo 6,2mm, o que exige apenas 420. Nesse ponto,

também são marcados como impossibilitados de análise os exames cujas segmentações não conseguiram reconhecer $2l$ pixels da superfície da córnea ou encontrar os 3 marcadores.

Mesmo com a melhoria dos resultados, a caracterização do movimento na face anterior da córnea gerou resultados com acurácia de diagnóstico máxima de 80%. A caracterização foi abandonada e substituída pela análise do movimento da face posterior da córnea, além do acréscimo da paquimetria. Devido a captação da parte posterior da córnea não ser uma imagem real, mas sim refratada e que sofre distorção, as informações do movimento da face posterior e da dinâmica da paquimetria devem ser referidas como aparentes, devido serem imagens captadas e que podem ser diferentes das informações reais.

Tal qual no outro método, o processamento inicia pela identificação da face anterior da córnea em cada imagem e suas conversões em vetores. Em seguida, para cada coluna, é feita a identificação do primeiro pixel preto após o primeiro pixel branco (no sentido superior-inferior), com isso, é identificado o limite posterior aparente da córnea. De forma análoga à porção anterior, são criados vetores correspondentes a essa nova região identificada.

Um terceiro conjunto de vetores é criado pela subtração dos valores da parte anterior, da parte posterior aparente. Isso dá a espessura aparente da córnea em cada coluna em cada imagem, sendo a paquimetria aparente em pixels. Ao mesmo tempo que esse vetor é criado são calculadas a média e o desvio padrão dos valores.

Esse último passo é necessário devido ao fato de que muitas córneas apresentaram problema na identificação da parte posterior por subexposição da córnea. Nesses locais a córnea apresenta afinamento irregular, sendo necessária correção coerente com o seu restante.

O vetor relativo à paquimetria aparente é percorrido a fim de identificar os pontos onde a córnea apresenta espessura inferior à metade do terceiro quartil da distribuição de sua espessura no *frame* atual. Se isso ocorrer, o valor da espessura é definido para o valor médio e o valor da parte posterior aparente é corrigido para o valor da parte anterior adicionado à espessura corrigida correspondentes.

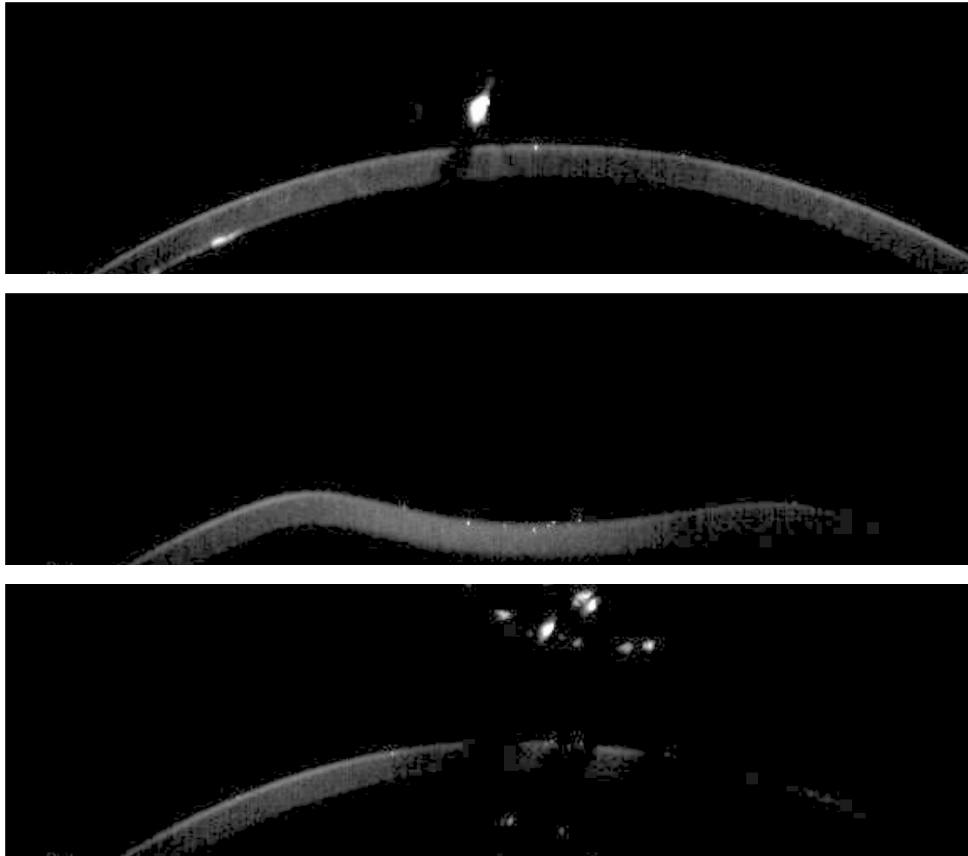
Para que o processamento seja apropriado, deve haver validação dos dados, isto é, se os mesmos quantificam imagens apropriadas para classificação de uma córnea e sua biomecânica. A primeira é uma validação de continuidade, isto é, a superfície da córnea deve apresentar pixels com regularidade esférica. São considerados falhas as imagens que apresentam descontinuidade (como nas imagens da Figura 3.3) ou quando – por causa de ruído na superfície – vários pixels são excluídos ou adicionados, formando cavidades ou picos, onde não deveria haver.

Com os dados nos vetores a continuidade pode ser calculada por

$$C(X) = \sum_{i=1}^{2l} v(x_i) \quad (3.6)$$

onde, $C(X)$ representa a função de validação de continuidade, X é um vetor que possui uma

Figura 3.3: Diferentes exames do Corvis, onde a córnea não apresenta continuidade.



Fonte: O autor.

superfície (anterior ou posterior aparente), x_i é cada um dos componentes de X e $v(x_i)$ é

$$v(x_i) = \begin{cases} 0 & \text{se } i = 1 \text{ ou } \|x_i - x_{i-1}\| \leq 15 \\ 1 & \|x_i - x_{i-1}\| > 15 \end{cases} \quad (3.7)$$

A segunda validação é de coerência. Trata-se da avaliação se a sequência de imagens é apropriada a análise. Para um vetor ser considerado coerente, ele deve ser igual ao vetor anterior ou seus valores terem sofrido variações máximas dentro de um intervalo estabelecido. Ou seja, se ao comparar a mesma coluna de dois quadros consecutivos e a superfície detectada da córnea realiza mudança muito grande de posição, é possível que uma das posições detectadas seja ruído.

A validação de coerência pode ser feita por

$$V(X) = \sum_{i=1}^{2l} o(x_i) \quad (3.8)$$

onde, $V(X)$ é a função de validação de coerência, X é um vetor de superfície e o_{x_i} é dada por

$$o(x_i) = \begin{cases} 0 & \text{se } i = 1 \text{ ou } \|x_{i,t} - x_{i,t-1}\| \leq 15 \\ 1 & \|x_{i,t} - x_{i,t-1}\| > 15 \end{cases} \quad (3.9)$$

onde os índices t e $t-1$ representam o vetor do *frame* atual e anterior, respectivamente.

Dentro do conjunto de amostras de treinamento, a maior distância entre dois pontos da mesma coluna em *frames* consecutivos ou de pixels em colunas consecutivas foi de 12, assim, foi definido o valor 15, para se ter a margem de 30% de que possa haver exames com distância maior sem considerar que seja ruído.

Devido exigir bastante precisão dos dados, o processamento das imagens possui tolerância a falhas, mas em pequena quantidade. Caso o somatório entre $C(X)$ e $V(X)$ supere um limite superior de erros (foi usado o valor 5), o exame é indicado como inválido para segmentação. A parametrização em 5 foi feita para evitar que os dados possuam excessiva quantidade de erros e tornem a avaliação demasiado imprecisa.

3.3.3 Processamento dos Dados: Aprendizado de Máquina

A fim de se comparar a proporcionalidade do movimento nos exames e não seus valores absolutos – como o valor da amplitude que, por observação dos dados dos exames, é proporcional a variação da PIO – os dados foram normalizados na amplitude de sua variação. A única exceção a essa normalização é a paquimetria aparente, que na verdade é processada em dois formatos diferentes: normalizada e não-normalizada.

Nesse formato, os dados de cada exame foram reduzidos a 140 vetores de 400 dimensões (56000 valores, no total). Esses vetores foram arranjados de forma a servir de entrada em duas formas diferentes de pós-processamento: *Deep Learning* e Aprendizado de Máquina sobre Wavelets.

Aprendizado com *Deep Learning*

Em um exame, o tamanho do conjunto de dados nos vetores é menor do que todos os pixels de suas imagens juntas (em um total de 16128000 de pixels). Isto tornou praticável a aplicação de DL, visto que o custo computacional para processar uma única imagem de tamanho 400×140 é bem menor do que uma sequência de 140 de 576×200 .

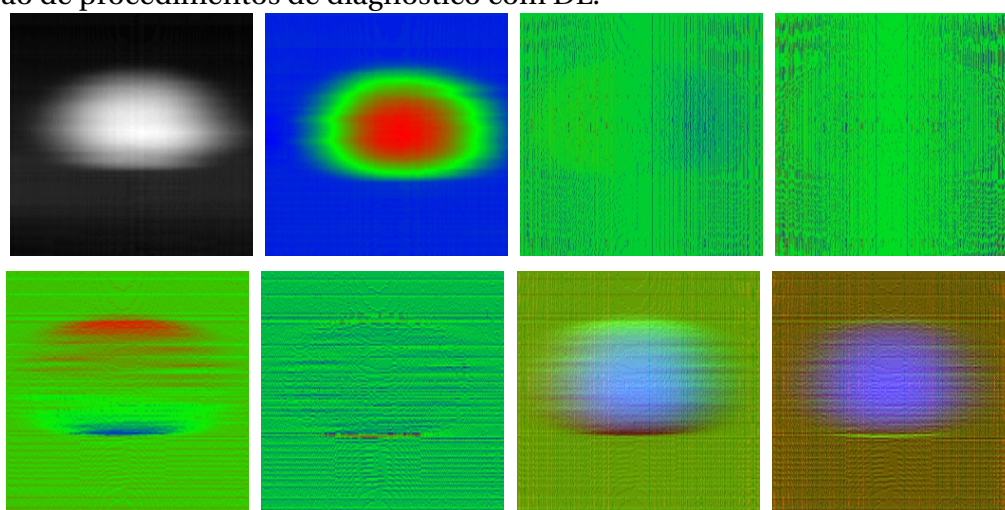
Tal como em [Koprowski et al. \(2014\)](#), isso foi feito pelo rearranjo dos dados, transformando em uma matrizes de 400×140 , as quais puderam ser transformadas em imagens para tentativa de reconhecimento de padrões. Essas imagens podem ser consideradas como visualizações da biomecânica corneal como um histograma 2D, pois do topo para a base se tem o prolongamento do tempo; da esquerda para a direita, a representação dos dados referente à imagem da córnea; as escala de cores, as amplitudes de valores.

Enquanto o estudo estava direcionado para caracterização do movimento da face corneal anterior, algumas formas de histogramas foram avaliados com uso de diferentes configurações de imagens e redes neurais.

A Figura 3.4 apresenta os diferentes formatos usados nesses modelos, sendo da esquerda para direita e de cima para baixo: movimento da face anterior em escala de cinza, movimento

da face anterior em escala RGB, primeira derivada no espaço (mudança na inclinação da córnea), segunda derivada no espaço (deformação), primeira derivada no tempo (velocidade da córnea), segunda derivada no tempo (aceleração corneal), composta 1 (composição do movimento com a inclinação e a deformação) e composta 2 (composição do movimento com a velocidade e a aceleração). A imagem em escalas de cinza apresenta maior amplitude de movimento com tons mais próximos a branco. As imagens em escala RGB apresentam três faixas de valores: azul (baixos), verde (próximos a média) e vermelho (próximos ao máximo).

Figura 3.4: Histogramas 2D sobre caracterização do movimento da face corneal anterior para avaliação de procedimentos de diagnóstico com DL.



Fonte: O autor.

Essas duas últimas foram tentativas de apresentação de mais de uma característica simultaneamente na mesma imagem. Isso foi feito pela composição das características em diferentes canais de cor. Dessa forma, a composta 1 tem no canal azul o movimento; no canal verde, a primeira derivada no espaço; o canal vermelho, a primeira derivada no tempo. De forma análoga, a composta 2 tem no canal azul o movimento; o canal verde, a segunda derivada no espaço; o canal vermelho, a segunda derivada no tempo.

Essas configurações foram executadas sobre redes no formato *deep learning* LeNET 5 e GoogleNET. Com GoogleNET os resultados foram inexpressivos, obtendo score de 0,0000. Já com a LeNET5 houveram resultados inclusive acima de 0,7843.

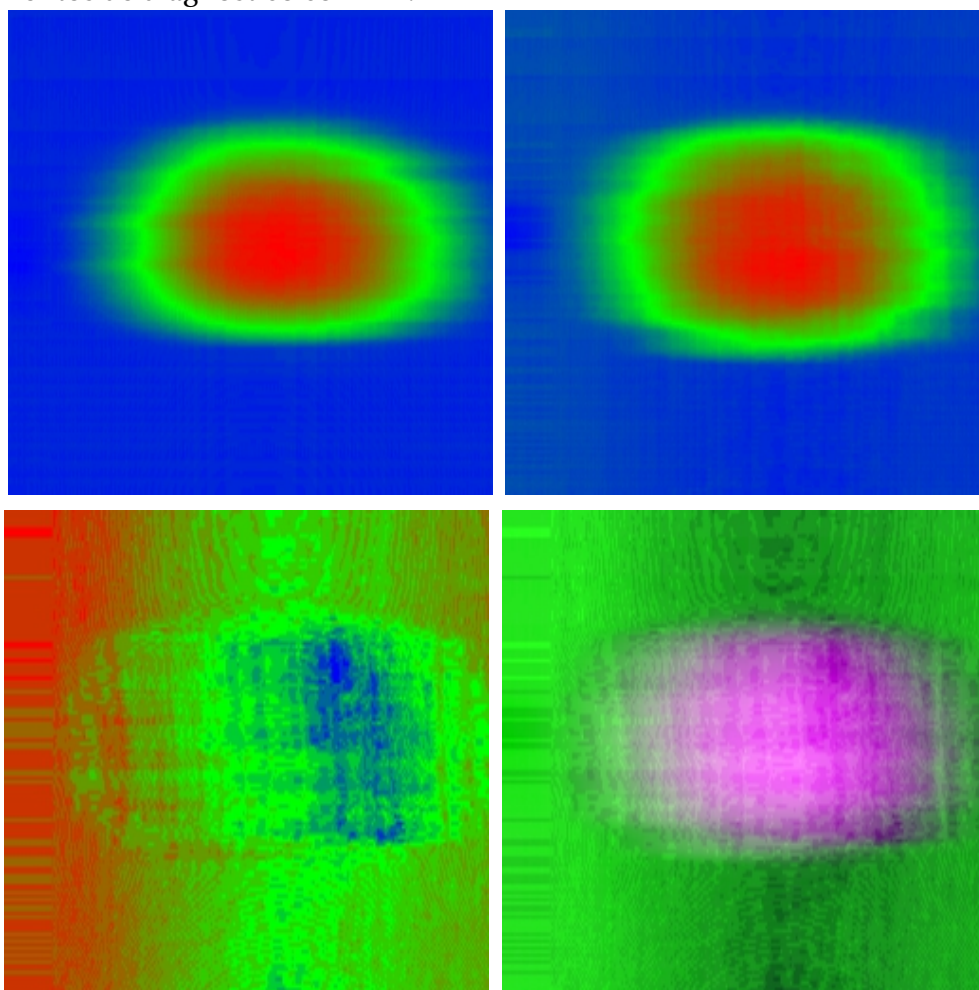
Mesmo não tendo resultados bons (com score variando entre 0,0000 e 0,2601), esses modelos evidenciaram que a combinação de características em uma única imagem podem gerar procedimentos viáveis à avaliação.

As implementações das redes foi feita utilizando o *framework* Caffe em modo *standalone*, sem uso de treinamentos prévios. Esses, por sua vez, foram executados com e sem imagens de média das bases.

Com a mudança das características em análise, foram feitas 4 avaliações: movimento da face corneal anterior, movimento aparente da face corneal posterior, paquimetria aparente e

composta. Esta última, baseou-se nos resultados anteriores, sendo a composição das outras 3 informações em uma única imagem. Essas imagens podem ser vistas na Figura 3.5, onde as imagens correspondem (da esquerda para direita e de cima para baixo): movimento da face anterior, movimento aparente da face posterior, paquimetria aparente e composta.

Figura 3.5: Histogramas 2D para análise da paquimetria corneal aparente para avaliação de procedimentos de diagnóstico com DL.



Fonte: O autor.

Os quatro tipos de histogramas deram origem ao total de 8 modelos processamento: 4 processando apenas as imagens geradas e outros 4 usando também com imagens médias. Essas últimas são um modelo de aprendizado no qual, a partir do conjunto de imagens de treinamento, é gerado uma imagem que possui o valor da média para cada pixel da imagem, gerando uma valores usados para normalizar todas as imagens a serem posteriormente classificadas com o modelo.

Treinamentos com estruturas para aprendizado com DL normalmente requerem uma quantidade muito grande de exemplos, sendo normal o uso de milhares e até milhões de exemplos. Como a quantidade de exames disponível é muito pequena, houve a necessidade de equalização das quantidades de exemplos para melhoramento das fases de treinamento e

testes. Assim, ao final da criação das imagens, foram criadas listas de exames para treinamento e testes contendo o maior número possível de elementos de cada classificação em quantidade igual. Isso foi feito para cada uma das bases.

Wavelets e Classificadores

Esse conjunto de experimentos foi baseado no trabalho de [Ribeiro \(2015\)](#), onde foram testadas 3 formas de wavelets (Coiflet, Haar e Daubuchies), usando o nível de wavelet 7 e aplicados 4 classificadores (regressão linear, RBF, MLP e SVM). Com esse formato, o exame foi tratado como um sinal de entrada. Para isso, os vetores correspondentes à segmentação da córnea foram encadeados em um único vetor que se assemelha a representação de uma onda, a qual foram aplicadas as transformações e os classificadores. Por esse modelo usar os dados das imagens de forma encadeada, o mesmo será referenciado no restante do texto como *imagens em lote*.

No corrente trabalho, por haver diferenças na forma de segmentação e na informação extraída, foram esperados resultados diferentes. Os dados também foram gerados e processados de forma diferente. Foram gerados 4 tipos de dados: movimento da face anterior, movimento aparente da face posterior, paquimetria aparente e composto. Este último é o somatório dos valores dos movimentos das faces da córnea, sendo a composição dessas características.

Após serem gerados no programa desenvolvido em C++, os dados foram exportados para o formato CSV e usados em *script* da linguagem de programação R, que aplicou Coiflet (com onda no formato *coif6* e níveis de wavelet 4 e 7) e o resultado foi exportado para novo arquivo CSV em conjunto com o valor da pressão intraocular, contendo os resultados de processamentos de todos os exames de treinamento. Nessa configuração, o sinal de entrada cai de um vetor aleatório de 56000 dimensões para apenas 3500 ou 437, dependendo do nível de wavelet. A [Figura 3.6](#) exemplifica a aplicação dessas configurações.

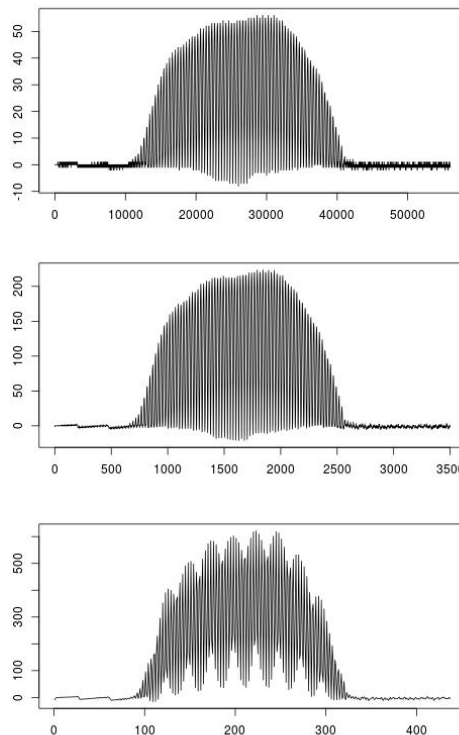
Outros formatos de wavelet foram tentados, mas os resultados apresentaram variação de escore inferior a 0,0010. Dessa forma, a *coif6* foi usado por ser um parâmetro padrão. Foram realizadas, ainda, variações dessas configuração, onde os dados foram exportados sem a presença da PIO, tendo uma análise apenas da biomecânica corneal.

Os arquivos CSV foram importados no aplicativo Rapid Miner e seus dados foram aplicados em modelos classificadores usando validação cruzada ¹ (com 10 grupos). Durante a avaliação apenas da face anterior da córnea, foi verificado que a presença da PIO piorou o escore dos diagnóstico e sua adição aos arquivos CSV, posteriormente, foi excluída.

Os classificadores usados foram regressão logística, árvore de decisão, *random forest*, MLP, *Fast-Large Margin*, *k*-NN e Naïve Bayes. Desses, árvore de decisão e *random forest* foram usados apenas para avaliação do movimento da face anterior da córnea, pois seus resultados

¹Na validação cruzada, o conjunto de amostras é separado em n grupos. A cada iteração, um dos grupos é escolhido aleatoriamente e os outros $n - 1$ são usados para treinamento. No final da iteração, o grupo escolhido é usado para validar o treinamento.

Figura 3.6: Sinal de entrada original (primeira imagem), aplicação de coif6 níveis 4 (segunda imagem) e 7 (terceira imagem)



Fonte: O autor.

não foram satisfatórios (score com valor 0). Enquanto que k -NN e Naïve Bayes foram usados apenas para análise de paquimetria aparente.

O presente capítulo detalhou como os dados foram gerados e preparados para seu uso nos experimentos. Em relação aos experimentos, os mesmos foram, tanto quanto possível, pormenorizados com a apresentação de como e porque foram desenvolvidos em seus formatos. No capítulo a seguir, serão apresentados os resultados dos experimentos e suas análises.

4

RESULTADOS E DISCUSSÃO

NESSE capítulo, são exibidos e discutidos os resultados experimentais da pesquisa. A discussão é iniciada com a análise dos dados, seguido pela segmentação das imagens e continua com ponderações sobre os modelos de aprendizagem de máquina aplicados.

4.1 Bases de Exames

A observação das imagens – exame a exame – permitiu algumas ponderações sobre os seus conteúdos. As imagens da primeira base possuem muita degradação da parte posterior da córnea, há várias ocorrências de objetos estranhos (fachos de luz e pixels com super exposição), além disso, também é predominante a subexposição de luz e algumas não possuem completa captura do perfil corneal.

Já as imagens da segunda base, apresentam melhor qualidade no que diz respeito à exposição de luz e à ocorrência de objetos estranhos. A parte posterior da córnea apresenta, de forma geral, pouca degradação.

Ao consultar os fornecedores das bases, foi constatado que os equipamentos de captura de imagens das bases possuem diferenças de versões e que a qualidade das imagens é tendencial a ser melhor na segunda base.

4.2 Segmentação e Processamento das Imagens

Tendo em vista que a aceitação de um exame como válido só ocorre durante a fase de validação dos dados, o primeiro critério para avaliação do processo de segmentação se deu pela percentual de exames assinalados como válidos.

Exportações de imagens providas de variadas versões do aplicativo do Corvis apresentaram diferentes resultados de validação de segmentação. Isso acarretou que todas as bases

tiveram que ser reexportadas para compatibilização das imagens. É necessário frisar que, no aplicativo, é possível fazer diversas operações sobre a imagem, como mudança de luminosidade e contraste, alterações de resolução e aplicações de filtros. No momento da exportação, todas essas modificações são ignoradas e as imagens voltam a seus estados originais.

Como resultado final, tem-se que a segmentação foi capaz de fornecer dados de 95,79% dos exames disponíveis (as técnicas de segmentação anterior permitiram o reconhecimento de apenas 23,16% dos exames). A discriminação do resultado pode ser encontrado na Tabela 4.1. Essas quantidades também são provindas do processamento das imagens e sua conversão em dados.

Tabela 4.1: Resultado da segmentação das imagens, por grupo de dados (usando a classificação Krumeich)

| Diagnóstico | Base Brasil | | | Base Europa | | |
|-------------|----------------|------------------|-------|----------------|------------------|-------|
| | Exames Válidos | Exames Inválidos | Total | Exames Válidos | Exames Inválidos | Total |
| Normais | 210 | 16 | 226 | 476 | 9 | 485 |
| Ceratocone | 205 | 17 | 222 | 201 | 6 | 207 |
| Total: | 415 | 33 | 448 | 677 | 15 | 692 |

Fonte: autor.

O processo iterativo, unindo no mesmo processamento o ajuste da filtragem de imagem com a avaliação dos dados, permitiu esse aumento, evitando a necessidade de intervenções manuais, como tratamentos de imagens.

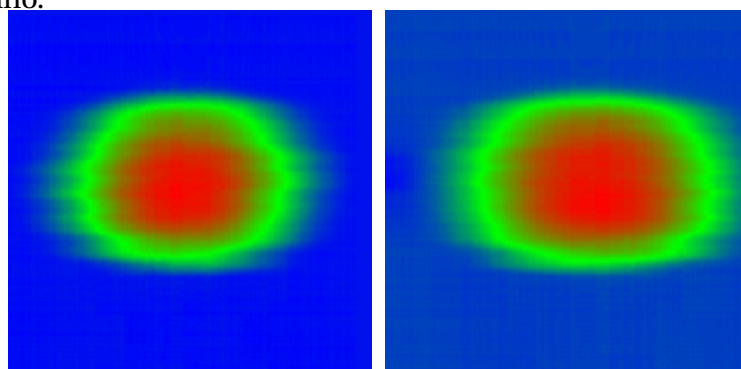
O reconhecimento da deflexão do olho, mesmo sem ter um modo de avaliação anterior à fase de aprendizagem, teve resultados visíveis nos histogramas 2D, com aumento das áreas com movimentos (como visto na Figura 4.1). Isso se deve ao fato de que, quando o algoritmo identificava de forma menos precisa, parte do movimento da córnea era interpretado como sendo deflexão do olho.

Na área azul da primeira imagem da Figura 4.1, pode-se perceber muito movimento irregular, mas em amplitude mais baixa. Na segunda imagem, parte desse movimento já aparece como tendo maior amplitude, passando já para a cor verde, e parte do movimento assinalado com cor verde passou a vermelho. Isso indica que esse dado foi melhor evidenciado na segunda imagem.

4.3 Aprendizado com *Deep Learning*

A execução de aprendizado com DL sobre os histogramas forneceu dois conjuntos de resultados distintos.

Figura 4.1: Histogramas 2D do mesmo exame e com técnicas diferentes de remoção de movimento do olho.



Fonte: O autor.

A primeira execução ocorreu sobre a base do Brasil. Os resultados estão sumarizados na Tabela 4.2.

Tabela 4.2: Diagnóstico resultante de processamento de histogramas 2D

| Característica | FP | VP | FN | VN | Acu. | Sens. | Espec. | RP |
|----------------------|----|-----|-----|-----|--------|--------|--------|--------|
| Paquimetria | 25 | 194 | 11 | 185 | 91,33% | 88,58% | 94,39% | 0,8361 |
| Paquimetria c/ Média | 35 | 191 | 14 | 175 | 88,19% | 84,51% | 92,59% | 0,7825 |
| Composta | 27 | 154 | 51 | 183 | 81,20% | 85,08% | 78,21% | 0,6654 |
| Composta c/ Média | 25 | 178 | 25 | 185 | 87,47% | 87,68% | 88,10% | 0,7725 |
| Anteior | 0 | 0 | 205 | 210 | 50,60% | 0,00% | 50,60% | 0,0000 |
| Anterior c/ Média | 11 | 124 | 81 | 199 | 77,83% | 91,85% | 71,07% | 0,6528 |
| Posterior | 32 | 165 | 40 | 178 | 82,65% | 83,76% | 81,65% | 0,6839 |
| Posterior c/ Média | 10 | 107 | 98 | 200 | 73,98% | 91,45% | 67,11% | 0,6137 |

Fonte: o autor

Na tabela, as colunas FP, VP, FN e VN referem-se aos diagnósticos (respectivamente: falso positivo, verdadeiro positivo, falso negativo e verdadeiro negativo). As colunas Acu., Sens. e Espec. dizem respeito à acurácia, sensibilidade e especificidade dos métodos de diagnóstico.

O segundo conjunto de resultados veio do processamento da base da Europa. Aqui, o resultado foi constante de diagnósticos com escore 0,00, com todos os diagnósticos de ceratocone identificados como falsos negativos (dando 0,00% de sensibilidade e 50,00% de especificidade).

Essas diferenças podem ser apontadas por dois motivos. O primeiro é que algoritmos de DL tentam aprender sobre características que a própria rede identifique como importantes (LeCun et al., 2015; Karpathy et al., 2014; Szegedy et al., 2015; Zeiler & Fergus, 2013). “Auxiliar” o aprendizado direcionando-o para observar características desejadas não só não auxiliou como, possivelmente, prejudicou o processo de aprendizado. O segundo motivo é que a degradação das imagens na base do Brasil são mais predominantes nos exames com ceratocone.

o que não ocorre com as imagens da base europeia. A correção da paquimetria aparente, mesmo que tenha permitido uso de maior quantidade de exames, não forneceu resultados que podem ser usados para diagnóstico.

O uso de *deep learning* não pode, no entanto, ser considerado inválido para diagnosticar ceratocone com as imagens provindas do Corvis. Como indicado na Seção 2.3.6, é comum que os treinamentos de uma estrutura de DL sejam feitos com grandes quantidades de exemplos, podendo a quantidade de exames disponível ter sido insuficiente para isso.

Uma forma alternativa para uso de DL é o uso de técnicas como as tentadas no início do corrente estudo e como corroborado por [Karpathy et al. \(2014\)](#), que é a análise da sequência das imagens originais dos exames (ou mais originais quanto possível) como um vídeo e o diagnóstico como uma classificação dessa mídia, necessitando de *hardware* adequado a isso. Ou, ainda, como [Venugopalan et al. \(2015\)](#), fazendo a tradução de uma sequência de imagens em uma segunda sequência de dados equivalente, neste caso seria o diagnóstico. Para melhorar o treinamento pode-se usar *data augmentation*, que é uma técnica onde são aplicadas pequenas rotações e translações nas mídias de treinamento, com o objetivo de possuir variações de mídias cujas classificações são conhecidas, aumentando o conjunto de dados de treinamento e a capacidade de generalização sobre características aprendidas.

Um teste realizado foi a aplicação da rede treinada com a base do Brasil na base da Europa (Tabela 4.3).

Tabela 4.3: Aplicação da rede DL treinada com base do Brasil na base da Europa

| Característica | FP | VP | FN | VN | Acu. | Sens. | Espec. | RP |
|----------------------|-----|-----|-----|-----|--------|--------|--------|--------|
| Paquimetria | 108 | 168 | 33 | 368 | 79,17% | 60,87% | 91,77% | 0,5586 |
| Paquimetria c/ Média | 172 | 178 | 23 | 304 | 71,20% | 50,86% | 92,97% | 0,4728 |
| Composta | 123 | 167 | 34 | 353 | 76,81% | 57,59% | 91,21% | 0,5252 |
| Composta c/ Média | 120 | 127 | 74 | 356 | 71,34% | 51,42% | 82,79% | 0,4257 |
| Anteior | 0 | 0 | 201 | 476 | 70,31% | 0,00% | 70,31% | 0,0000 |
| Anterior c/ Média | 10 | 21 | 180 | 466 | 71,94% | 67,74% | 72,14% | 0,4886 |
| Posterior | 10 | 24 | 177 | 466 | 72,38% | 70,59% | 72,47% | 0,5115 |
| Posterior c/ Média | 161 | 105 | 96 | 315 | 62,04% | 39,47% | 76,64% | 0,3025 |

Fonte: o autor

Esses resultados não são estranhos levando em conta as diferenças entre as bases e que a quantidade de exemplos na bases não são suficientes para que a rede possua um grau de generalização suficiente para atuar da mesma forma nas duas bases. Além disso, esse resultado é compatível com a discussão ocorrida em [Amancio et al. \(2014\)](#), visto que técnicas de aprendizagem são apropriadas à base de treinamento e dados que foram adquiridos de forma compatível com esta, podendo ter resultados díspares quando aplicadas a outras bases ou com dados adquiridos de forma diferente (mesmo que por aprimoramento de *hardware* ou *software*).

O teste contrário foi realizado (aplicação da rede treinada com a base europeia na base brasileira). O resultado não foi alterado em relação ao treinamento da rede (0,00 de escore, com todos os diagnósticos de ceratocone acusados como sendo falsos negativos).

4.4 Aprendizado sobre Wavelets

Apesar dos resultados apresentados com *deep learning*, a aprendizagem com imagens em lote e aplicando wavelets teve resultados mais variados.

O uso de diferentes níveis de wavelet permitiu a avaliação se o uso de diferentes quantidades de dados influencia nos resultados de diagnóstico. Apesar do uso final com níveis 4 e 7, outros níveis foram tentados. Níveis entre 1 e 3 a informação é representada por muitos dados, o que tornou o processo de aprendizagem muito lento e, algumas vezes, causando fechamento do aplicativo pelo sistema operacional por demasiado consumo de recursos. Em níveis com valores superiores a 7 a perda de informação foi tamanha que não houve aprendizado (escore 0,00, com variações entre especificidade e sensibilidade com 0,00%).

Muitas configurações dos métodos de aprendizagem foram tentados. As árvores de decisão foram abandonadas devido à seleção de atributos cobrir quantidades muito pequenas de parâmetros fazendo as árvores ter poucos níveis e atingindo muita imprecisão (com escore máximo de 0,005).

O uso de MLP foi feito usando 2 camadas com 4 neurônios cada, com taxa de aprendizado 0,1, com decaimento de 0,14 por ciclo e máximo de 250 ciclos. Números maiores de camadas, neurônios, taxa de aprendizado ou eras resultaram em *overfit* ou tendência em avaliar todos os exames como normais.

Entre os modelos de regressores disponíveis no RapidMiner, a regressão logística foi o único a apresentar escores superiores a 0,5. Foi utilizado L-BFGS para estimação dos parâmetros e sem uso de coeficientes negativos.

Diferentes modelos de SVM foram tentados e todos com resultados muito aproximados, dentre eles o Fast Large Margin por conseguir resultados pouco melhores que os outros (com escore diferente em no máximo 0,1), tendo-se melhor resultado com parâmetro $C = 0,6$ e com uso da biblioteca L2 SVM Dual para resolução dos coeficientes.

O método Naïve Bayes foi usado por ter apresentado resultados significativos em estudos na área de medicina (Prinyakupt & Pluempitiwiriwawej, 2015) e por ser empregado em situações de classificação de classes desbalanceadas (Frank & Bouckaert, 2006).

Já o k -NN, como em Zhang & Zhou (2007), teve bons resultados. Mesmo que em aplicação diferente, com entradas muito grandes, no caso dos autores foram usados textos. Foi utilizado $k = 1$, valores maiores apresentaram leve diminuição no escore (máximo de 0,05), mas com grande demora.

A tabulação dos resultados das execuções desses métodos com as melhores configurações

– nas condições do corrente estudo – de aprendizagem computacional podem ser visto no Apêndice A (nas Tabelas A.1 e A.2). Os escores dessas execuções estão expostos nas Tabela 4.4 e 4.5

Tabela 4.4: Escores de modelos preditivos com treinamento na base de exames do Brasil

| | Característica | MLP | Regress. Log. | FLM | k -NN | Naïve Bayes |
|------------|----------------|--------|---------------|--------|---------|-------------|
| W. Nível 4 | Anterior | 0,0000 | 0,7366 | 0,7239 | 0,6449 | 0,5323 |
| | Posterior | 0,0000 | 0,7479 | 0,7194 | 0,6238 | 0,5294 |
| | Paquimetria | 0,4172 | 0,5053 | 0,7442 | 0,6331 | 0,5816 |
| | Composta | 0,0000 | 0,6684 | 0,7240 | 0,4875 | 0,4400 |
| W. Nível 7 | Anterior | 0,0000 | 0,7701 | 0,6290 | 0,6095 | 0,5121 |
| | Posterior | 0,0000 | 0,7426 | 0,6760 | 0,5653 | 0,5301 |
| | Paquimetria | 0,0000 | 0,5097 | 0,7275 | 0,6928 | 0,5848 |
| | Composta | 0,0000 | 0,6579 | 0,7221 | 0,4665 | 0,4481 |

Fonte: o autor

Tabela 4.5: Escores de modelos preditivos com treinamento na base de exames da Europa

| | Característica | MLP | Regress. Log. | FLM | k -NN | Naïve Bayes |
|------------|----------------|--------|---------------|--------|---------|-------------|
| W. Nível 4 | Anterior | 0,0000 | 0,3796 | 0,7656 | 0,5454 | 0,6540 |
| | Posterior | 0,7058 | 0,6997 | 0,7040 | 0,4710 | 0,3209 |
| | Paquimetria | 0,7371 | 0,8113 | 0,8247 | 0,7133 | 0,5582 |
| | Composta | 0,0000 | 0,7987 | 0,7817 | 0,5965 | 0,5574 |
| W. Nível 7 | Anterior | 0,0000 | 0,6467 | 0,6337 | 0,5189 | 0,6053 |
| | Posterior | 0,6956 | 0,7245 | 0,6402 | 0,4333 | 0,3486 |
| | Paquimetria | 0,7458 | 0,8019 | 0,7815 | 0,7133 | 0,5358 |
| | Composta | 0,0000 | 0,7894 | 0,7098 | 0,5439 | 0,5398 |

Fonte: o autor

Observando as duas tabelas, pode-se perceber que a mudança de nível de wavelet produz resultados diferentes, com tendência a melhores escores para o nível 4. Houve piora de escores em apenas 7 casos na base do Brasil, contra 4 na base europeia. Em ambas bases 3 escores não sofreram modificações com a mudança de nível de wavelet. Em 10 configurações na base do Brasil e 13, na da Europa, os resultados foram pouco melhores que seus equivalentes em nível 7.

O uso de MLP expõe uma grande diferença entre as bases. Enquanto que seu uso na base do Brasil foi o pior dos modelos, obtendo o menor escore (0,0000), menor média de escore (0,0522) e menor escore máximo (0,4172 com paquimetria aparente e wavelet nível 4), na base da Europa, houveram 4 resultados com escores superiores a 0,69 – mesmo com pior escore mínimo, este teve escore máximo melhor que os modelos k -NN e Naïve Bayes. Mesmo não tendo o melhor escore dentre os modelos estudados, ainda assim o MLP obteve a melhor sensibilidade (96,77%) com a análise da paquimetria aparente com mesmo nível de wavelet 7.

Observando a Tabela A.1, pode-se perceber que a aplicação de MLP sobre os dados de algumas características apresenta sensibilidade de 0,0%, mas acurácia de 70,31% nos

mesmos casos, sendo isso devido ao grande desbalanceamento da quantidade de exames com diagnóstico normal sobre ceratocone. A avaliação apenas da acurácia poderia levar a acreditar que, nessas situações, o modelo não seria mal sucedido.

As outras formas de aprendizagem apresentaram maior plasticidade, visto que, excetuando alguns casos, tiveram variação menos bruscas – comparando os mesmos modelos, nos mesmos níveis de wavelet, entre uma base e outra – e tendo maior capacidade de diagnósticos em ambas classificações.

O segundo pior modelo foi Bayes ingênuo, com o segundo pior valor mínimo de escore em ambas as bases, com a segunda pior média de escore de ambas as bases, mas com o pior escore máximo entre os modelos. Por ter tido resultados piores que o da regressão logística e do *fast large margin*, isso é um indicativo de que os dados de treinamento não são independentes, como visto na Seção 2.3.5. Esse fato leva ao aumento da crença da correlação entre a biomecânica corneal e o diagnóstico de ceratocone, pois duas importantes relações dos dados são descartadas com esse modelo: a dependência de localização de um ponto da córnea com sua vizinhança e sua mudança com o decorrer do tempo do exame.

Em ambas as bases, o modelo k -NN obteve menor escore, média de escore e maior escore um pouco maiores os valores do Naïve Bayes, não tendo valores escore, acurácia, sensibilidade ou especificidade com destaque entre as outras técnicas de aprendizagem de máquina.

Os melhores resultados foram alcançados com regressão logística e *fast large margin*. Na base brasileira, a regressão logística apresentou melhores resultados com a análise dos movimentos de ambas as faces da córnea – obtendo escore 0,7701 na análise do movimento aparente da região posterior, usando wavelet nível 7. Já o *fast large margin* obteve o melhor escore (com valor de 0,8247, usando wavelet nível 4, sobre dados da paquimetria aparente), maior especificidade e acurácia (com valores 92,14% e 91,44%, sobre os mesmos dados).

O uso de análise com composição de características não apresentou os melhores resultados, mas os valores foram, no pior caso, e à exceção do uso com k -NN e Bayes ingênuo na base do Brasil, pelo menos iguais a uma das características que o compõem, tendo muitos resultados melhores.

A paquimetria aparente, levando em consideração que é a subtração de duas ondas, também é uma composição. Apresentou melhores resultados com quase todos os modelos pesquisados, considerando comparações usando o mesmo modelo, com os mesmos níveis de wavelet e nas mesmas bases. Com essa característica atingiu o melhor escore, acurácia e especificidade – 0,8247, 91,44% e 92,14%, respectivamente – (na base europeia, com wavelet em nível 4 e com uso de *fast large margin*), além da melhor sensibilidade de (96,77%, na base europeia, com wavelet nível 4, usando MLP).

Apesar da aplicação da aprendizagem com *deep learning* na base da Europa ter sido inviável, na base do Brasil, apesar dos problemas de qualidade da imagem, foi possível fazer o processamento com escore

Esses resultados são compatíveis com discussões feitas em [Daxer & Fratzl \(1997\)](#) e [Shin](#)

et al. (1997), pois a maior parte da espessura corneal é dada pelo estroma. No ceratocone, como há modificações no arranjo de sua estrutura e há modificação na forma de dissipação das forças aplicadas na córnea, é esperado que haja modificação na forma com a córnea compensa a pressão intraocular e forças externas.

A paquimetria informada pelo Corvis aparenta ser uma característica estática, pois é fornecida como um único valor pelo aparelho. No entanto, como visto nos histogramas 2D desenvolvidos para o processamento com *deep learning*, a paquimetria aparente é um valor dinâmico e pode ser relacionado diagnóstico do ceratocone.

Mesmo que a técnica de Scheimpflug não apresente imagens reais da porção posterior da córnea e que a medida da paquimetria não seja real – ambos interferidos pela refração da luz, seja pela curvatura da córnea ou outras questões anatômicas –, as imagens capturadas pelo Corvis permitem que se identifique padrões no movimento da córnea, capazes de relacionar a biomecânica desse tecido ao diagnóstico do ceratocone.

Esse capítulo apresentou as vantagens e desvantagens do ambiente *Reproducible Research* durante todo o processo de desenvolvimento e acompanhamento deste trabalho. Finalizamos com o próximo capítulo, no qual serão relatados os principais impactos e contribuições deste trabalho e os passos futuros.

5

CONCLUSÕES

ESTE capítulo abordará o avanço que o presente trabalho traz para o cenário científico. Além disso, apresentará sugestões para trabalhos futuros.

5.1 Considerações Finais

Mesmo com parâmetros que permitem boa acurácia no diagnóstico de ceratocone, a variação de implementações nas consecutivas versões do *software* do Corvis comprometem a qualidade dos dados gerados pelo aparelho. Seus parâmetros de descrição da biomecânica corneal gerados não cobrem todos os aspectos dessa área de estudo e as sequências de imagens capturadas durante os exames ainda são uma fonte de informações não completamente explorada pelo fabricante.

Com o processo de segmentação configurado durante o desenvolvimento do corrente estudo, foi possível delinear 4 tipos de caracterizações da biomecânica corneal, que apresentaram resultados diferentes, havendo evidências que o estudo da dinâmica da paquimetria aparente propicia diagnóstico do ceratocone.

Usando representação dos dados ora na forma de vetores e histogramas 2D, ora na forma de imagens em lote, os escores obtidos permitiram a identificação de técnicas de aprendizagem de máquina aplicáveis a esses modelos de dados e que adquiriram capacidade de classificação com seus padrões. Desses, o processamento de imagens em lote com *fast large margin* (com wavelet nível 4 na base da Europa) e a avaliação de histogramas 2D com processamento em *deep learning* (na base do Brasil) apresentaram melhores resultados.

É possível que as bases de exames usadas possam apresentar sobreposição, pois há autores em comum a ambas. Isso impossibilita a junção das bases para treinamento, o que poderia propiciar aprendizado dos classificadores de exames com maior capacidade de generalização e melhor diagnóstico. Mas, ao contrário disso, a multiplicidade de exames pode levar a classificadores com menor capacidade de aprendizagem e maior *overfit*.

As imagens estudadas são muito dependentes do *hardware* de captura e a análise da paquimetria aparente é dependente da qualidade dessas imagens. Isso quer dizer que caso a córnea sofra muitas alterações nas suas características refratárias e reflexivas ou o equipamento apresente dificuldade na aplicação da técnica de Scheimpflug, as caracterizações das superfícies da córnea podem ser degradadas e até mesmo não ocorrer. Além disso as imagens ainda possuem ruído que podem influenciar na identificação da região anterior e posterior aparente.

5.2 Trabalhos futuros

Como a biomecânica ainda é uma área de estudo em aberto e que há possibilidade de aplicação do corrente estudo em aspectos diferentes da mesma, é interessante expandir os limites para o uso desses conhecimentos como, por exemplo:

- Uso de *data augmentation*, com geração de variações de imagens dos exames reais, para aumento da quantidade de exames para treinamento.
- Identificação de outras formas de representação dos exames e aplicação delas a técnicas de aprendizado de máquina.
- Identificação de outros métodos de redução de dimensionalidade de vetores que auxiliem o método de processamento de imagens em lote.
- Processamento das sequências de imagens geradas pelo Corvis com pouco ou nenhum tratamento por redes neurais com *deep learning* para classificação direta dos vídeos.

Apêndice A

TABULAÇÃO DE EXECUÇÃO DE TÉCNICAS DE APRENDIZAGEM DE MÁQUINA SOBRE DADOS DE EXAMES

Tabela A.1: Resultado de treinamento da base de exames do Brasil com aprendizagem de máquina sobre wavelet

| Característica | MLP | | | Regressão Logística | | | Fast Large Margin | | | k-NN | | | Naive Bayes | | |
|----------------|-------|---------|--------|---------------------|---------|--------|-------------------|---------|--------|-------|---------|--------|-------------|---------|--------|
| | Acur. | Especi. | Senci. | Acur. | Especi. | Senci. | Acur. | Especi. | Senci. | Acur. | Especi. | Senci. | Acur. | Especi. | Senci. |
| Anterior | 50,60 | 50,60 | 0,00 | 85,73 | 84,79 | 86,87 | 84,83 | 82,10 | 88,17 | 80,24 | 79,09 | 81,54 | 72,99 | 75,00 | 70,97 |
| Posterior | 50,60 | 50,60 | 0,00 | 86,24 | 83,70 | 89,36 | 84,83 | 85,17 | 84,47 | 78,77 | 76,75 | 81,28 | 72,68 | 75,68 | 69,95 |
| Paquimetria | 53,52 | 52,15 | 80,00 | 71,04 | 71,03 | 71,14 | 86,30 | 86,26 | 86,27 | 79,49 | 78,54 | 80,61 | 75,88 | 73,31 | 79,33 |
| Composta | 50,60 | 50,60 | 0,00 | 81,67 | 80,45 | 83,08 | 85,05 | 84,26 | 85,93 | 69,62 | 68,10 | 71,58 | 65,79 | 70,00 | 62,86 |
| Anterior | 50,60 | 50,60 | 0,00 | 87,66 | 86,64 | 88,89 | 79,11 | 82,89 | 75,88 | 78,10 | 78,20 | 77,94 | 71,41 | 75,40 | 67,92 |
| Posterior | 50,60 | 50,60 | 0,00 | 85,99 | 83,93 | 88,48 | 81,28 | 89,29 | 75,71 | 75,15 | 74,88 | 75,50 | 72,63 | 75,54 | 70,18 |
| Paquimetria | 50,60 | 50,60 | 0,00 | 71,28 | 70,40 | 72,40 | 85,34 | 85,65 | 84,94 | 83,10 | 81,53 | 84,97 | 76,13 | 73,62 | 79,44 |
| Composta | 50,60 | 50,60 | 0,00 | 80,91 | 79,11 | 83,16 | 84,83 | 88,08 | 81,98 | 68,17 | 67,11 | 69,52 | 66,00 | 71,70 | 62,50 |

Fonte: o autor

Tabela A.2: Resultado de treinamento da base de exames da Europa com aprendizagem de máquina sobre wavelet

| Característica | MLP | | | Regressão Logística | | | Fast Large Margin | | | k-NN | | | Naive Bayes | | |
|----------------|-------|---------|--------|---------------------|---------|--------|-------------------|---------|--------|-------|---------|--------|-------------|---------|--------|
| | Acur. | Especi. | Senci. | Acur. | Especi. | Senci. | Acur. | Especi. | Senci. | Acur. | Especi. | Senci. | Acur. | Especi. | Senci. |
| Anterior | 70,31 | 70,31 | 0,00 | 70,60 | 71,47 | 53,12 | 88,33 | 89,15 | 85,88 | 78,28 | 82,07 | 66,46 | 81,44 | 81,75 | 80,00 |
| Posterior | 86,12 | 88,35 | 79,89 | 83,46 | 83,33 | 83,97 | 86,71 | 90,38 | 77,89 | 74,73 | 80,08 | 58,82 | 56,28 | 83,58 | 38,39 |
| Paquimetria | 82,90 | 81,47 | 90,48 | 89,96 | 89,84 | 90,30 | 91,44 | 92,14 | 89,50 | 86,13 | 87,75 | 81,29 | 79,48 | 85,62 | 65,20 |
| Composta | 70,31 | 70,31 | 0,00 | 89,07 | 88,80 | 89,94 | 89,80 | 91,45 | 85,48 | 80,34 | 83,40 | 71,52 | 79,48 | 84,76 | 65,76 |
| Anterior | 70,31 | 70,31 | 0,00 | 78,73 | 78,14 | 82,76 | 83,40 | 88,89 | 71,29 | 76,95 | 80,89 | 64,15 | 80,94 | 83,08 | 72,86 |
| Posterior | 85,37 | 87,33 | 79,65 | 83,91 | 83,18 | 87,10 | 84,04 | 90,84 | 70,48 | 72,81 | 78,74 | 55,03 | 61,15 | 84,03 | 41,48 |
| Paquimetria | 78,88 | 77,07 | 96,77 | 89,51 | 89,47 | 89,63 | 89,65 | 91,09 | 85,79 | 86,12 | 87,75 | 81,29 | 78,29 | 85,07 | 62,98 |
| Composta | 70,31 | 70,31 | 0,00 | 88,18 | 87,64 | 90,07 | 86,55 | 89,21 | 79,57 | 78,28 | 82,32 | 66,07 | 78,50 | 84,45 | 63,92 |

Fonte: o autor

REFERÊNCIAS BIBLIOGRÁFICAS

- Ali, N. Q., Patel, D. V., McGhee, C. N. J. & A., K. (2014), 'Biomechanical responses of healthy and keratoconic corneas measured using a noncontact scheimpflug-based tonometer', *Investigative Ophthalmology & Visual Science* **55**(6), 3651. URL <http://iovs.arvojournals.org/article.aspx?doi=10.1167/iovs.13-13715>.
- Alio, J. L., Vega-Estrada, A., Esperanza, S., Barraquer, R. I., Teus, M. A. & Murta, J. (2014), 'Intrastromal corneal ring segments: How successful is the surgical treatment of keratoconus?', *Middle East African Journal of Ophthalmology*. URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3959038/>.
- Amancio, D. R., Comin, C. H., Casanova, D., Travieso, G., Bruno, O. M., Rodrigues, F. A. & da Fontoura Costa, L. (2014), 'A systematic comparison of supervised classifiers', *PLOS ONE* **9**(4), 1–14. URL <https://doi.org/10.1371/journal.pone.0094137>.
- Ambrósio, R. J., Lopes, B. T., Correia, F. F., Salomão, M. Q., Bühren, J., Roberts, C. J., Elsheikh, A., Vinciguerra, R. & Vinciguerra, P. (2017), 'Integration of scheimpflug-based corneal tomography and biomechanical assessments for enhancing ectasia detection', *Journal of Refractive Surgery* **33**(7), 434–443. URL <https://www.healio.com/ophthalmology/journals/jrs/2017-7-33-7/%7B634e00a9-9c1a-4271-8ce5-df466977c8c3%7D/integration-of-scheimpflug-based-corneal-tomography-and-biomechanical-assessments>
- Ambrósio, R. J., Ramos, I., Luz, A., Faria, F. F., Steinmueller, A., Krug, M., Belin, M. W. & Roberts, C. J. (2013), 'Dynamic ultra high speed scheimpflug imaging for assessing corneal biomechanical properties', *Revista Brasileira de Oftalmologia*. URL http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0034-72802013000200005.
- Arreola, K. Z., Fehr, J. & Burkhardt, H. (2007), Fast support vector machine classification of very large datasets, Technical report, Freiburg Albert-Ludwigs-Universität/Institut für Informatik.
- Bao, F., Geraghty, B., Wang, Q. & Elsheikh, A. (2016), 'Consideration of corneal biomechanics in the diagnosis and management of keratoconus: is it important?', *Eye and Vision*. URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4932704/>.
- Binder, P. S. B., Lindstrom, R. L., Stulting, R. D., Donnenfeld, E., Wu, H., McDonnell, P. & Rabinowitz, Y. (2005), 'Keratoconus and corneal ectasia after lasik', *Journal of Refr*. URL <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.619.2046&rep=rep1&type=pdf>.
- Calderbank, A. R., Daubechies, I., Sweldens, W. & Boon-Lock, Y. (1998), 'Wavelet transforms that map integers to integers', *Applied and Computational Harmonic Analysis* **5**(3), 332–369. URL <http://www-sciencedirect-com.ez9.periodicos.capes.gov.br/science/article/pii/S1063520397902384>.
- Calôba, L. P. (1992), Redes neurais, in E. d. P. Ferreira & M. F. Sarcinelli, eds, '9º Congresso Brasileiro de Automática', Vitória/ES, pp. 25–38.

- Canny, J. (1986), 'A computational approach to edge detection', *IEEE Transactions on Pat* **8**(6), 679–698. URL <http://ieeexplore.ieee.org/document/4767851/>.
- Chollet, F. (2015), *Keras Documentation*. URL <https://keras.io/>.
- Cover, T. & Hart, P. (1967), 'Nearest neighbor pattern classification', *IEEE Transactions on Information Theory* **13**(1), 21–27.
- Dantas, P. B. (2017), Utilização da aprendizagem de máquina e seleção de atributos para o diagnóstico de ceratocone a partir de parâmetros biomecânicos, Master's thesis, Universidade Federal de Alagoas, Maceió/AL. URL http://www.ufal.edu.br/unidadeacademica/ic/pos-graduacao/modelagem-computacional-de-conhecimento/dissertacoes/2017/pedro-barreto-dantas-pre-banca/at_download/file.
- Daubechies, I. (1990), 'The wavelet transform, time-frequency localization and signal analysis', *IEEE Transactions on Information Theory* **36**(5), 961–1005. URL <http://ieeexplore.ieee.org/document/57199/>.
- Daxer, A. & Fratzl, P. (1997), 'Collagen fibril orientation in the human corneal stroma and its implication in keratoconus', *Investigative Ophthalmology and* **38**(1), 121–129. URL <http://iovs.arvojournals.org/article.aspx?articleid=2161514&resultClick=1>.
- Demassieux, N., Jutand, F., Saint-Paul, M. & Dana, M. (1985), Vlsi architecture for a one chip video median filter, in 'ICASSP '85. IEEE International Conference on Acoustics, Speech, and Signal Processing', Vol. 10, pp. 1001–1004.
- DeVries, R. (2000), 'Vygotsky, piaget, and education: a reciprocal assimilation of theories and educational practices', *New Ideas in Psychology* **18**(2), 187 – 213. URL <http://www.sciencedirect.com/science/article/pii/S0732118X00000088>.
- Dome, E. F. (2008), *Estudo do Olho Humano Aplicado à Optmetria*, 4 ed., São Paulo/SP.
- Dubbelman, M., Sicam, V. & der Heijde, G. V. (2006), 'The shape of the anterior and posterior surface of the aging human cornea', *Vision Research* **46**(6), 993 – 1001. URL <http://www.sciencedirect.com/science/article/pii/S0042698905004906>.
- Faceli, K., Lorena, A. C., Gama, J. & Carvalho, A. C. P. L. F. (2011), *Inteligência Artificial: Uma Abordagem de Aprendizado de Má*, Livros Técnicos e Científicos, Rio de Janeiro/RJ.
- Fan, R.-E., Chang, K.-W., Hsieh, C.-J., Wang, X.-R. & Lin, C.-J. L. (2008), 'Liblinear : A library for large linear classification', *Journal of Machine Learning Research* **9**, 1871–1874.
- Frank, E. & Bouckaert, R. R. (2006), 'Naive bayes for text classification with unbalanced classes', *European Conference on Principles of Data Mining and Knowledge Discovery*. URL <https://link.springer.com/book/10.1007/11871637#page=524>.
- Frank, E., Trigg, L., Holmes, G. & Witten, I. H. (2000), 'Technical note: Naive bayes for regression', *Machine Learning* **41**(1), 5–25. URL <https://doi.org/10.1023/A:1007670802811>.
- Fung, Y. C. (1993), *Biomechanics: Mechanical Properties of Living Tissues*, 2 ed., Springer, San Diego/CA - U.S.A.

- Glass, D. H., Roberts, C. J., Litsky, A. S. & A., W. P. (2008), 'A viscoelastic biomechanical model of the cornea describing the effect of viscosity and elasticity on hysteresis', *Investigative Ophthalmology*. URL <http://iovs.arvojournals.org/article.aspx?articleid=2164254>.
- Gomes, Willer, J., Nicoletto, E. & Bueno, J. S. (2017), *Programa de Manipulação de Imagens GNU*. URL https://docs.gimp.org/2.8/pt_BR/.
- Hatze, H. (1974), 'The meaning of the term 'biomechanics'', *Journal of Biomechanics* **7**(2), 189–190. URL [http://www.jbiomech.com/article/0021-9290\(74\)90060-8/abstract](http://www.jbiomech.com/article/0021-9290(74)90060-8/abstract).
- Haykin, S. (2009), *Neural Networks and Learning Machines*, 3 ed., PHI Learning Private Limited, New Delhi, India.
- Hosmer, D. W. & Lemeshow, S. (2000), *Applied Logistic Regression*, 2 ed., Wiley-Interscience Publication, New Jersey, USA. URL http://resource.heartonline.cn/20150528/1_3kOQSTg.pdf.
- Hsieh, W. W. (2009), *Machine Learning Methods in the Enviromental Sciences: Neural Networks and Kernels*, Cambridge University Press, New Delhi, India.
- Jia, Y. (2014), *Caffe: Deep Learning Framework*, Berkeley AI Research, Berkeley University. URL <http://caffe.berkeleyvision.org/>.
- Kanski, J. J. (2003), *Oftalmologia Clínica: uma abordagem sistemática*, 5 ed., Elsevier Editora LTDA., São Paulo/SP.
- Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R. & Fei-Fei, L. (2014), Large-scale video classification with convolutional neural networks, in 'CVPR'. URL <http://cs.stanford.edu/people/karpathy/deepvideo/>.
- Kaur, S., Kaur, G. & Singh, D. (2013), 'Comparative analysis of haar and coiflet wavelets using discrete wavelet transform in digital image compression', *International Journal of Engineering Research and Applications* **3**(3), 669–673. URL http://www.ijera.com/papers/Vol3_issue3/DK33669673.pdf.
- Koprowski, R. & Ambrósio, R. (2015), 'Quantitative assessment of corneal vibrations during intraocular pressure measurement with the air-puff method in patients with keratoconus', *Computers in Biology and Medicine* **66**, 170–178. URL <http://linkinghub.elsevier.com/retrieve/pii/S0010482515003157>.
- Koprowski, R., Lyssek-Boron, A., Nowinska, A., Wylegala, E., Kasprzak, H. & Wrobel, Z. (2014), 'Selected parameters of the corneal deformation in the corvis tonometer', *BioMedical Engineering OnLine*. URL <https://biomedical-engineering-online.biomedcentral.com/articles/10.1186/1475-925X-13-55>.
- Langley, P., Iba, W., Thompson, K. et al. (1992), An analysis of bayesian classifiers, in 'Association for the Advancement of Artificial Intelligence', Vol. 90, pp. 223–228.
- LeCun, Y., Bengio, Y. & Hinton, G. (2015), 'Deep learning', *International Weekly Journal of Science* **1**(521), 436–444. URL <http://www.nature.com/nature/journal/v521/n7553/abs/nature14539.html?foxtrotcallback=true>.

- LIS (2015), *Deep Learning Tutorial*, laboratoire d'informatique des systèmes adaptatifs ed. URL <http://deeplearning.net/software/theano/theano.pdf>.
- LIS (2016), *Theano Documentation*. URL <http://deeplearning.net/software/theano/theano.pdf>.
- Lorena, A. C. & Carvalho, A. C. P. L. F. (2007), 'Uma introdução às support vector machines', *Recuperação da Informação e Tecnologias Avançadas* **14**(2), 43–67. URL http://www.seer.ufrgs.br/rita/article/view/rita_v14_n2_p43-67.
- Machado, A. P., Lyra, J. a. M., Ambrósio, Jr., R., Ribeiro, G., Araújo, L. P. N., Xavier, C. & Costa, E. (2011), Comparing machine-learning classifiers in keratoconus diagnosis from ora examinations, in 'Proceedings of the 13th Conference on Artificial Intelligence in Medicine', AIME'11, Springer-Verlag, Berlin, Heidelberg, pp. 90–95. URL <http://dl.acm.org/citation.cfm?id=2040981.2040996>.
- Magalhães, M. N. & Lima, A. C. P. d. (2007), *Noções Básicas de Probabilidade e Estatística*, 6 ed., São Paulo/SP.
- Manche, E. E., Carr, J. D. & S., H. P. (1998), 'Excimer laser refractive surgery.', *Western J* **169**(1), 30–38. URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1305094/>.
- Matyunin, S., Vatolin, D., Berdnikov, Y. & Smirnov, M. (2011), Temporal filtering for depth maps generated by kinect depth camera, in '2011 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON)', pp. 1–4.
- Morettin, P. A. & Bussad, W. O. (2014), *Estatística Básica*, São Paulo/SP.
- Muccioli, C., Campos, M., Goldchmit, M., Dantas, P. E. C., Bechara, S. J. & Costa, V. P. (2006), 'A córnea não é um pedaço de plástico', *Arquivos Brasileiros de O.* URL http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0004-27492006000500001.
- Mérula, R. V. & Trindade, F. C. (2006), 'Degeneração marginal pelúcida da córnea: diagnóstico e tratamento', *Arquivos Brasileiros de Of.* URL http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0004-27492006000300032.
- Müller, L. J., Pels, E. & Vrensen, G. F. J. M. (2001), 'The specific architecture of the anterior stroma accounts for maintenance of cornea curvature', *British Journal of Op* **85**(4), 437–443. URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1723934/>.
- O'Neil, C. H. (2013), 'K-nearest neighbors: dangerously simple'. URL <https://mathbabe.org/2013/04/04/k-nearest-neighbors-dangerously-simple/>.
- OpenCV Development, G. (2017), 'Opencv 3.3.0-dev: Open source computer vision - filters', OnLine. URL http://docs.opencv.org/trunk/da/d17/group__ximgproc__filters.html.
- Pelillo, M. (2014), 'Alhazen and the nearest neighbor rule', *Pattern Recognition Letters* **38**, 34–37.
- Prinyakupt, J. & Pluempitiwiriyaewej, C. (2015), 'Segmentation of white blood cells and comparison of cell morphology by linear and naïve bayes classifiers', *BioMedical Engineering OnLine*. URL <https://biomedical-engineering-online.biomedcentral.com/articles/10.1186/s12938-015-0037-1>.

- Rabinowitz, Y. S. (1998), 'Keratoconus', *Survey of Ophthalmology* **42**(4), 297–319. URL <https://www.ncbi.nlm.nih.gov/pubmed/9493273>.
- Rebordao, J. & Cabral, A. (2008), 'Maskless optical interferometric lithography'. URL <https://www.google.com/patents/US7319551>, US Patent 7,319,551.
- Ribeiro, G. B. O. (2015), Uso da transformada de wavelet e técnicas de aprendizado de máquina para criação de mmodel computacional de auxílio ao diagnóstico de ceratocone baseado em parâmetros biomecânicos da córnea, Master's thesis, Universidade Federal de Alagoas, Maceió/AL. URL <http://www.repositorio.ufal.br/bitstream/riufal/1605/1/Uso%20da%20transformada%20de%20Wavelet%20e%20t%C3%A9cnicas%20de%20aprendizado%20de%20m%C3%A1quina%20para%20cria%C3%A7%C3%A3o%20de%20modelo%20computacional%20de%20aux%C3%ADlio....pdf>.
- Rodrigues, M. d. L. V. & Dantas, A. M. (2001), *Oftalmologia Clínica*, 2 ed., Cultural Médica, Rio de Janeiro/RJ.
- Romero-Jiménez, M., Santodomingo-Rubido, J. & Wolffsohn, J. S. (2010), 'Keratoconus: A review', *Journal of The British Contact Lens Association* **33**(4), 157–166. URL <http://www.sciencedirect.com/science/article/pii/S1367048410000561>.
- Russel, S. J. & Norvig, P. (2010), *Artificial Intelligence: A Modern Approach*, 3 ed., Prentice Hall, New Jersey - USA.
- Schmidhuber, J. (2014), Deep learning in neural networks: An overview, Technical report, Istituto Dalle Molle di Studi sull'Intelligenza Artificiale, University of Lugano & SUPSI, Lugano/Switzerland. URL <http://www.idsia.ch/~juergen/DeepLearning8Oct2014.tex>.
- Segaran, T. (2008), *Programa a Inteligência Coletiva*, Alta Books, Rio de Janeiro/RJ.
- Segata, N. & Blanzieri, E. (2011), 'Fast and scalable local kernel machines', *Journal of Machine Learning Research* **11**, 1883–1926. URL <http://www.jmlr.org/papers/v11/segata10a.html>.
- Sherwin, T. & Brookes, N. H. (2004), 'Morphological changes in keratoconus: pathology or pathogenesis', *Clinical and Experimental Op* **32**(2), 211–217. URL <http://onlinelibrary.wiley.com/doi/10.1111/j.1442-9071.2004.00805.x/abstract>.
- Shetty, R., Nuijts, R. M. M. A., Srivatsa, P., Jayadev, C., Pahuja, N., Akkali, M. C. & Sinha Roy, A. (2015), 'Understanding the correlation between tomographic and biomechanical severity of keratoconic corneas.', *BioMed research international*. URL <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4402575>.
- Shin, T. J., Vito, R. P., Johnson, L. W. & McCarey, B. E. (1997), 'The distribution of strain in the human cornea', *Journal of Biomechanics* **30**(5), 497–503. URL <https://www.ncbi.nlm.nih.gov/pubmed/9109561#>.
- Sii, F., Lee, G. A., Sanfilippo, P. & Stephensen, D. C. (2004), 'Pellucid marginal degeneration and scleroderma', *Clinical and Experimental Optometry* pp. 180–184. URL https://www.academia.edu/18487603/Pellucid_marginal_degeneration_and_scleroderma.

- Sinjab, M. M. (2012), *Classifications and Patterns of Keratoconus and Keratectasia*, Springer, Verlag, Berlin/Heidelberg. DE. URL
http://www.springer.com/cda/content/document/cda_downloaddocument/9783642218392-c1.pdf?SGWID=0-0-45-1260541-p174127563.
- Spirkovska, L. (1993), A summary of image segmentation techniques, Memorando Técnico 104002, NASA, Ames Research Center, Moffett Field, California. URL
https://www.researchgate.net/publication/24285532_A_summary_of_image_segmentation_techniques.
- Stanton, J. M. (2001), 'Galton, pearson, and the peas: A brief history of linear regression for statistics instructors', *Journal of Statistics Education*. URL
<https://ww2.amstat.org/publications/jse/v9n3/stanton.html>.
- Stein, H. A., Freeman, M. I. & Stein, R. M. (1997), *Contact Lenses: fundamentals and clinical use*, 1 ed., Slack Incorporated, Thorofare, NJ - USA.
- Steinberg, J., Katz, T., Lücke, K., Frings, A., Druchkiv, V. & Linke, S. J. (2015), 'Screening for keratoconus with new dynamic biomechanical in vivo scheimpflug analyses', *Cornea* **34**(11), 1404–1412. URL
<http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00003226-201511000-00014>.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erthan, D., Vanhoucke, V. & Rabinovich, A. (2015), Going deeper with convolutions, in 'IEEE Conference on Computer Vision and Pattern Recognition', IEEE, Boston, MA, USA. URL
<http://ieeexplore.ieee.org/document/7298594/>.
- Tian, L., Huang, Y. F., Wang, L. Q., Bai, H., Wang, Q., Jiang, J. J., Wu, Y. & Gao, M. (2014), 'Corneal biomechanical assessment using corneal visualization scheimpflug technology in keratoconic and normal eyes', *Journal of Ophthalmology* **2014**, 8. URL
<https://www.hindawi.com/journals/joph/2014/147516/>.
- Valença, M. (2010), *Fundamentos das Redes Neurais*, 2 ed., Livro Rápido - Elógica, Recife/PE.
- Vaughan, D. & Asbury, T. (1990), *Oftalmologia Geral*, 3 ed., Atheneu, São Paulo/SP.
- Vellara, H. R. & Patel, D. V. (2015), 'Biomechanical properties of the keratoconic cornea: a review', *Clinical and Experimental Opt* **98**(1), 31–38.
- Vellara, H. R., Ali, N. Q., Gokul, A., Turuwhenua, J., Patel, D. V. & McGhee, C. N. J. (2015), 'Quantitative analysis of corneal energy dissipation and corneal and orbital deformation in response to an air-pulse in healthy eyes', *Investigative Ophthalmology & Visual Science* **56**(11), 6941–6497. URL
<http://iovs.arvojournals.org/article.aspx?doi=10.1167/iovs.15-17396>.
- Venugopalan, S., Rohrbach, M., Donahue, J., Mooney, R., Darrell, T. & Saenko, K. (2015), 'Sequence to sequence – video to text', *CoRR*. URL
<https://arxiv.org/abs/1505.00487v3>.

- Vinciguerra, R., Ambrósio, R., Elsheikh, A., Roberts, C. J., Lopes, B., Morengi, E., Azzolini, C. & Vinciguerra, P. (2016a), 'Detection of keratoconus with a new biomechanical index', *Journal of Refractive Surgery* **32**(12), 803–810. URL <http://www.healio.com/doiresolver?doi=10.3928/1081597X-20160629-01>.
- Vinciguerra, R., Elsheikh, A., Roberts, C. J., Ambrósio, R., Kang, D. S. Y., Lopes, B. T., Morengi, E., Azzolini, C. & Vinciguerra, P. (2016b), 'Influence of pachymetry and intraocular pressure on dynamic corneal response parameters in healthy patients', *Journal of Refractive Surgery* **32**(8), 550–561. URL <http://www.healio.com/doiresolver?doi=10.3928/1081597X-20160524-01>.
- Wallang, B. S. & Das, S. (2013), 'Keratoglobus', *Eye* **27**(9), 1004–1012. URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3772364/#>.
- Wang, L.-K., Tian, L. & Zheng, Y.-P. (2016), 'Determining in vivo elasticity and viscosity with dynamic scheimpflug imaging analysis in keratoconic and healthy eyes', *Journal of Biophotonics* (5), 454–463.
- Weisberg, S. (2005), *Applied Linear Regression*, 3 ed., John Wiley & Sons, Inc., Hoboken/NJ-EUA.
- Witten, I. H. & Frank, E. (2005), *Data mining : practical machine learning tools and techniques*, 2 ed., Oxford/UK.
- Wollensak, G., Spoerl, E. & Seiler, T. (2003), 'Riboflavin/ultraviolet-a-induced collagen cross-linking for the treatment of keratoconus', *Am J Ophthalmol.* URL <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.553.9353&rep=rep1&type=pdf>.
- Wu, Z., Yao, T., Fu, Y. & Jiang, Y.-G. (2016), 'Deep learning for video classification and captioning', *CoRR*. URL <https://arxiv.org/abs/1609.06782v1>.
- Yanoff, M. & Fine, B. S. (1991), *Patologia Ocular: Atlas Colorido*, 1 ed., São Paulo/SP.
- Ye, C., Yu, M., Lai, G. & Jhanji, V. (2015), 'Variability of corneal deformation response in normal and keratoconic eyes', *Optometry and Vision Science* **92**(7), 149–153. URL <http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00006324-201507000-00016>.
- Zeiler, M. D. & Fergus, R. (2013), 'Visualizing and understanding convolutional networks', *CoRR* p. 818–833. URL <http://arxiv.org/abs/1311.2901>.
- Zhang, M.-L. & Zhou, Z.-H. (2007), 'M l-knn : A lazy learning approach to multi-label learning', *Pattern Recognition* **40**(7), 2038–2048. URL <http://www.sciencedirect.com/science/article/pii/S0031320307000027?via%3Dihub>.
- Zhang, Q., Shen, X., Xu, L. & Jia, J. (2014), 'Rolling guidance filter', *European Conference on Computer Vision*. URL <http://www.cse.cuhk.edu.hk/leojia/projects/rollguidance>.
- Zhu, C., Byrd, R. H., Lu, P. & Nocedal, J. (1997), 'Algorithm 778: L-bfgs-b: Fortran subroutines for large-scale bound-constrained optimization', *ACM Trans. Math. Softw.* **23**(4), 550–560. URL <http://doi.acm.org/10.1145/279232.279236>.

Este trabalho foi redigido em \LaTeX utilizando uma modificação do estilo IC-UFAL. As referências bibliográficas foram preparadas no JabRef e administradas pelo \BIBTeX com o estilo LaCCAN. O texto utiliza fonte Fourier-GUTenberg e os elementos matemáticos a família tipográfica Euler Virtual Math, ambas em corpo de 12 pontos.