

Tomaz Ribeiro Viana Bisneto
Orientador: Antonio Oseas de Carvalho Filho

Determinação automática do glaucoma utilizando índices de diversidade taxonômica

Picos - PI
12 de Novembro de 2018

Tomaz Ribeiro Viana Bisneto
Orientador: Antonio Oseas de Carvalho Filho

Determinação automática do glaucoma utilizando índices de diversidade taxonômica

Monografia submetida ao Curso de Bacharelado em Sistemas de Informação como requisito parcial para obtenção de grau de Bacharel em Sistemas de Informação.

Universidade Federal do Piauí
Campus Senador Heuvídio Nunes de Barros
Bacharelado em Sistemas de Informação

Picos - PI
12 de Novembro de 2018

FICHA CATALOGRÁFICA
Serviço de Processamento Técnico da Universidade Federal do Piauí
Biblioteca José Albano de Macêdo

V614d Viana Bisneto, Tomaz Ribeiro
Determinação automática do glaucoma utilizando índices de diversidade taxonômica / Tomaz Ribeiro Viana Bisneto.– 2018.
CD-ROM : il.; 4 ¾ pol. (37 f.)
Trabalho de Conclusão de Curso (Curso Bacharelado em Sistemas de Informação) – Universidade Federal do Piauí, Picos, 2018.
Orientador(A): Prof. Dr. Antonio Oseas de Carvalho Filho

- Glaucoma-Processamento de Imagens 2. Índices de Diversidade Taxonômica. 3. Reconhecimento de Padrões. I. Título.

CDD 006.37

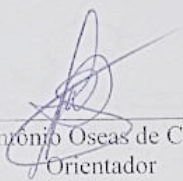
DETERMINAÇÃO AUTOMÁTICA DO GLAUCOMA UTILIZANDO ÍNDICES DE
DIVERSIDADE TAXONÔMICA

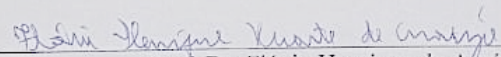
TOMAZ RIBEIRO VIANA BISNETO

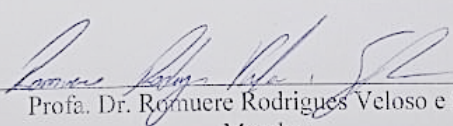
Monografia aprovado como exigência parcial para obtenção do grau de
Bacharel em Sistemas de Informação.

Data de Aprovação

Picos - PI, 26 de novembro de 2018


Prof. Dr. Antônio Oseas de Carvalho Filho
Orientador


Prof. Dr. Flávio Henrique de Araújo
Membro


Profa. Dr. Romuere Rodrigues Veloso e Silva
Membro

Agradecimentos

Agradeço primeiramente a Deus por conceder a realização deste trabalho.

Aos meus familiares, amigos, orientador, professores e todos que fizeram parte da minha formação, o meu muito obrigado.

Por fim, a todos que, de forma direta ou indireta, contribuíram à realização e conclusão desta monografia.

Glória a Deus!

Transmita o que aprendeu. Força, maestria. Mas fraqueza, insensatez, fracasso também.

*Sim, fracasso acima de tudo. O maior professor, o fracasso é... Esse é o verdadeiro
fardo de todos os mestres.*

Mestre Yoda (Star Wars: The Last Jedi)

Resumo

O glaucoma pode ser definido como uma doença neurodegenerativa, que possui uma etiologia multifatorial, que ocasiona na lesão progressiva do nervo óptico. Por ser uma doença crônica e causar cegueira permanente, caso não tratada, seu diagnóstico precoce é necessário, garantindo melhor qualidade de vida e um prognóstico mais favorável para o paciente. O Processamento Digital de Imagens em conjunto com técnicas computacionais de aprendizado de máquina possibilitam a criação de métodos para detecção automática do glaucoma. Assim, esse trabalho apresenta a proposta que busca a criação de uma abordagem para viabilizar o diagnóstico do glaucoma, com o uso de características de textura a partir de índices de diversidade taxonômica. Onde, os resultados indicam que o método é promissor, alcançando um resultado de 98,7% de acurácia com o classificar *Random Forest*.

Palavras-chaves: Glaucoma, Processamento de Imagens, índices de diversidade taxonômica, Reconhecimento de padrões.

Abstract

Glaucoma can be defined as a neurodegenerative disease, which has a multifactorial etiology that leads to progressive damage to the optic nerve. Because it is a chronic disease and cause permanent blindness, if left untreated, its early diagnosis is necessary, ensuring a better quality of life and a more favorable prognosis for the patient. The Digital Image Processing along with computational techniques of machine learning allows the creation of methods for automatic detection of glaucoma. Thus, this paper presents the proposal that seeks to create an approach to enable the diagnosis of glaucoma, with the use of texture feature from taxonomic diversity indexes. Where, the results indicate that the method is promising, achieving a result 98.7 % of accuracy with the Random Forest classifier.

Key words: Glaucoma, Image Processing, Taxonomic Diversity Indexes, Pattern Recognition.

Lista de ilustrações

Figura 1 – Imagem de retina sem glaucoma	16
Figura 2 – Imagem de retina com glaucoma	16
Figura 3 – Etapas do processamento digital de imagens. Adaptado de(GONZALEZ; WOODS, 2002).	17
Figura 4 – Exemplo de arvore filogenéticas de alguns primatas. Disponível em (ARAÚJO, 2003).	19
Figura 5 – (a) Representação da região analisada(b)Árvore filogenética extraída (c) Matriz com distâncias levando em consideração a quantidade de arestas da árvore filogenética que as separa as espécies analisadas. Dis- ponível em (CARVALHO et al., 2018).	20
Figura 6 – Metodologia proposta.	26
Figura 7 – Imagem de retina sem glaucoma disponivel na base DRISHTI-GS . . .	27
Figura 8 – Imagem de retina sem glaucoma disponivel na base RIM-ONE v2 . . .	27
Figura 9 – Exemplo de árvore enraizada na forma de cladograma inclinado. Dis- ponível em (FILHO et al., 2017b).	28

Lista de tabelas

Tabela 1	– Testes com níveis de cinza	30
Tabela 2	– Testes com o canal vermelho (Red)	30
Tabela 3	– Testes com o canal verde (Green)	30
Tabela 4	– Testes com o canal azul (Blue)	30
Tabela 5	– Testes com o todos os canais e nível de Cinza	30
Tabela 6	– Trabalhos relacionados	31

Lista de abreviaturas e siglas

A	Acurácia
AUC	Área sob curva ROC
CAD	Computer-Aided Diagnosis
DO	Disco Óptico
E	Especificidade
FFT	<i>Fast Fourier Transform</i>
FN	Falso Negativo
FP	Falso Positivo
GLCM	<i>Gray Level Co-occurrence Matrix</i>
HOS	Espectros de Ordem Superior
K	Coefficiente Kappa
LDA	<i>Linear Discriminant Analysis</i>
MIAG	Grupo de Análises de Imagens Médicas
MLP	<i>Multi Layer Perceptron</i>
PCA	<i>Principal Component Analysis</i>
PIO	Pressão Intraocular
ROC	<i>Receiver Operating Characteristic</i>
ROI	Região de Interesse
S	Sensibilidade
SMO	<i>Sequential Minimal Optimization</i>
SVM	Máquina de Vetores de Suporte
VN	Verdadeiro Negativo
VP	Verdadeiro Positivo

Lista de símbolos

Δ	Diversidade Taxonômica
Δ^*	Distinção Taxonômica
X_i	Abundância da espécie i
X_j	Abundância da espécie j
Σ	Somatório
N	Número total de Indivíduos
ω_{ij}	Distância da espécie i à espécie j
P_o	Taxa de aceitação relativa
P_e	Taxa hipotética de aceitação.

Sumário

1	Introdução	13
1.1	Objetivos	14
1.2	Organização do trabalho	14
2	Referencial Teórico	15
2.1	Glaucoma	15
2.2	Processamento de Imagens	16
2.2.1	Textura	18
2.3	Índices de Diversidade	18
2.3.1	Índices de Diversidade Taxonômica	19
2.4	Classificação e Reconhecimento de Padrões	20
2.4.1	Classificadores	21
2.5	Validação	22
3	Trabalhos Relacionados	24
4	Método proposto	26
4.1	Aquisição de Imagens	27
4.2	Extração de Características	28
4.3	Classificação	29
4.4	Validação	29
5	Resultados e Discussões	30
6	Conclusão	33
	Referências	34

1 Introdução

O glaucoma pode ser definido como uma doença neurodegenerativa, possuindo uma etiologia multifatorial, que ocasiona a lesão progressiva do nervo óptico, com perda de campo visual correspondente (MELLO; SUSANNA; ALMEIDA, 2016). Por ser uma doença crônica e assintomática, a princípio, é necessário seu diagnóstico precoce para que assim sejam aplicados os procedimentos para amenização de prováveis complicações, evitando a perda parcial ou total da visão.

Mesmo as causas da cegueira variando de acordo com as questões socioeconômicas e geográficas, o glaucoma, independente da população e nível social, se mantém como uma das principais causas de cegueira irreversível (QUIGLEY; BROMAN, 2006). Como visto em Kingman (2004), o glaucoma é a segunda maior causa de cegueira no mundo, está atrás apenas da catarata. Contudo, a perda de visão causada pela catarata pode ser reversível através de cirurgia, enquanto que o glaucoma causa cegueira permanente.

Mediante o quadro clínico apresentado pelo o glaucoma, a detecção nos estágios iniciais aumenta a possibilidade de obter um melhor tratamento. Quando o glaucoma é identificado na fase inicial facilita-se o tratamento por não apresentar grande relevância das lesões, favorecendo ao profissional responsável no tratamento e ao paciente diminuindo custo terapêutico (JÚNIOR et al.,).

Como visto em Nunes (2006), os sistemas de diagnóstico auxiliado por computador (ou simplesmente sistemas CAD, de “*computer-aided diagnosis*”) são sistemas computacionais, muitas vezes acoplados a equipamentos médicos, com a finalidade de auxiliar na tomada de decisão a respeito de um diagnóstico. Aplicações com finalidades diversas têm sido desenvolvidas por vários grupos de pesquisas, visando auxiliar na composição de diagnósticos como uma forma de contribuir para a detecção precoce de doenças.

Com intuito de aumentar a precisão na detecção de lesões, a utilização de técnicas de Visão Computacional e Reconhecimento de Padrões, tem se mostrado efetivas e apresentam bons resultados (LEDERLIN et al., 2013; FILHO et al., 2017a; MUNIA et al., 2017). O desenvolvimento de métodos automáticos para detecção e diagnóstico de doenças estão sendo amplamente exploradas. O que possibilita a detecção de anormalidades em possíveis regiões suspeitas, resultando no auxílio ao diagnóstico de forma antecipada, sendo esta, uma segunda opinião para os especialistas.

Os sistemas CAD geralmente se utilizam das informações extraídas de imagens médicas, que podem ser provenientes de diversos tipos de modalidades de exames. Comumente esses sistemas fazem uso de um conjunto de técnicas computacionais, tais como inteligência artificial, técnicas de extração de características e reconhecimento de padrões, que possibilitam extrair e analisar informações úteis ao diagnóstico de imagens clínicas.

1.1 Objetivos

Em vista do contexto apresentado, este trabalho tem como objetivo desenvolver uma abordagem que viabiliza o diagnóstico automático do glaucoma, utilizando Índice de Diversidade Taxonômico para caracterização da textura da imagem da retina. Realizando a extração de características em imagens do Disco Óptico (DO), tendo como base a diversidade taxonômica dos pixels, a fim de identificar padrões nas texturas, tornando possível a quantificação do glaucoma.

Para a concretização do objetivo geral, tem-se os seguintes objetivos específicos:

1. Desenvolver e adaptar técnicas para caracterizar propriedades de textura;
2. Classificar as regiões de interesse extraídas em glaucoma e normal
3. Avaliar a viabilidade do uso de Índices de diversidade taxonômicos como forma de extração de textura em imagens do DO; e,
4. Construir uma metodologia capaz de sugerir ao especialista uma segunda opinião na detecção e no diagnóstico do Glaucoma.

1.2 Organização do trabalho

Este trabalho está organizado em seis capítulos. O Capítulo 2 apresenta os fundamentos teóricos para o entendimento da metodologia abordada no trabalho. No Capítulo 3 são detalhados os principais trabalhos relacionados. No Capítulo 4 são abordadas as etapas de desenvolvimento da pesquisa, iniciando pela aquisição das imagens, extração de características de textura utilizando os índices de diversidade taxonômica, seguida pela classificação e finalizando com a validação da metodologia proposta. O Capítulo 5 discute os resultados obtidos com a metodologia proposta. E por fim, no Capítulo 6 são apresentadas as conclusões e trabalhos futuros.

2 Referencial Teórico

De modo a obter entendimento claro da metodologia e resultados, aqui apresentados, faz-se necessário conhecer alguns conceitos que darão suporte ao entendimento do tema. Neste capítulo objetivas-se o esclarecimento da teoria para o entendimento deste trabalho, fornecendo o embasamento teórico necessário para a compreensão do mesmo. Desta forma, serão abordados os seguintes tópicos: Glaucoma, processamento de imagens, índices de diversidade taxonômica, classificação e reconhecimento de padrões e Validação dos Resultados.

2.1 Glaucoma

O glaucoma é classificado como neuropatia óptica crônica, progressiva, caracterizada por alterações do DO e da camada de fibras nervosas da retina. Possuindo como anatomia patológica o surgimento de lesões progressivas dos nervos ópticos, compreendendo um conjunto de afecções oculares, ocasionando a perda, irreversível, total ou parcial da visão (MELLO; SUSANNA; ALMEIDA, 2016).

Dentre os fatores de desenvolvimento do glaucoma, a elevação da Pressão Intraocular (PIO) é a mais comum e principal causa da doença. Esta elevação acontece devido a obstrução do canal de drenagem do líquido humor aquoso, resultando em inúmeras afecções oculares em células da retina e do nervo óptico (MELLO; SUSANNA; ALMEIDA, 2016). A fim de detectar essas anomalias causadas pelo glaucoma, são utilizados exames de avaliação de imagens, tais como a Tomografia de Coerência Óptica, Tomografia Retiniana Heidelberg ou por meio de métodos computacionais com análises de imagens da retina.

Por ser assintomático, a princípio, e de etiologia multifatorial o glaucoma acaba apresentando certo grau de dificuldade para sua detecção. Esse cenário contribuiu para o glaucoma se tornar a segunda maior causa de cegueira mundial, como visto em Nugroho et al. (2016), atingindo o número de 60 milhões de casos no ano de 2010, e com estimativas de alcançar 80 milhões de casos até o ano de 2020.

O processo de detecção do glaucoma por meio de exames oftalmológicos se dá principalmente pela avaliação do aumento do PIO. No entanto, o glaucoma pode ocorrer sem aumento da pressão ocular. Assim, outra forma de detecção, é pela análise de danos ao nervo óptico, sendo necessário um especialista treinado, demandando um alto custo, além de ser um processo demorado (ZILLY; BUHMANN; MAHAPATRA, 2017).

Segundo Foster et al. (2002) a principal característica que diferencia o glaucoma de outras causas de morbidade visual e do estado normal da retina é um padrão característico de dano à cabeça do nervo óptico. Assim, é possível a identificação do glaucoma analisando mudanças de formas, cores ou profundidade em imagens do DO, é possível analisar essas

alterações comparando as Figuras 1 e 2. Nesse cenário, a utilização de sistemas CAD auxiliam a detecção precoce de lesões, contrinuindo com tratamento e garantindo um prognóstico mais favorável para o paciente.



Figura 1: Imagem de retina sem glaucoma

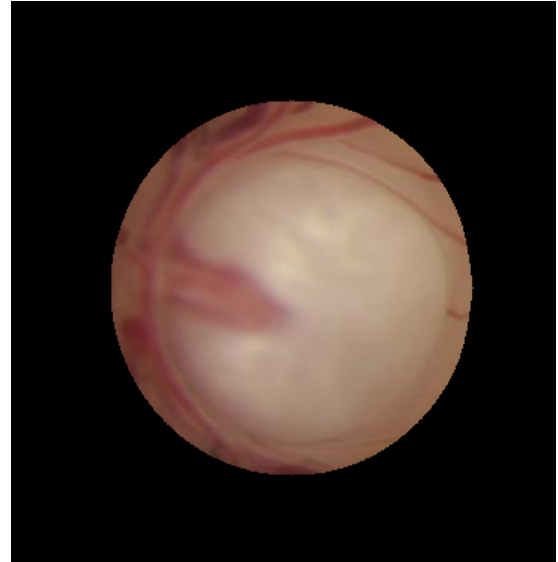


Figura 2: Imagem de retina com glaucoma

2.2 Processamento de Imagens

A história do processamento de imagens digitais como ciência data de 1964, quando durante a corrida espacial, imagens fornecida por sondas espaciais eram processadas por computador para corrigir vários tipos de distorção inerentes as câmeras utilizadas (FILHO; NETO, 1999). Desde o ano de 1964 aos dias atuais, a área de processamento de imagens vem crescendo expressivamente e o contexto de suas aplicações permeiam quase todos os ramos da atividade humana.

Segundo Foster et al. (2002), o processamento de imagens digitais compreende processos cujas entradas e saídas são imagens e, além disso, engloba os processos de extração de atributos a partir de imagens, incluindo o reconhecimento de objetos individuais. Tendo como objetivos, a melhora da informação visual para interpretação humana e representação dos dados para percepção automática através de máquinas.

De maneira geral, como visto na bibliografia (FOSTER et al., 2002; UMBAUGH, 1997; FILHO; NETO, 1999), o processamento de imagens segue um conjunto de 5 passos principais utilizados para demonstrar as etapas do ciclo de vida do processamento de imagem (Figura3). Essas etapas consistem em: aquisição de imagens, pré-processamento, segmentação, representação e descrição, reconhecimento e interpretação.

Ao final de cada umas das etapas do processamento de imagens os resultados gerados funcionam como entrada para a próxima etapa, dessa forma cada um dos passos irá

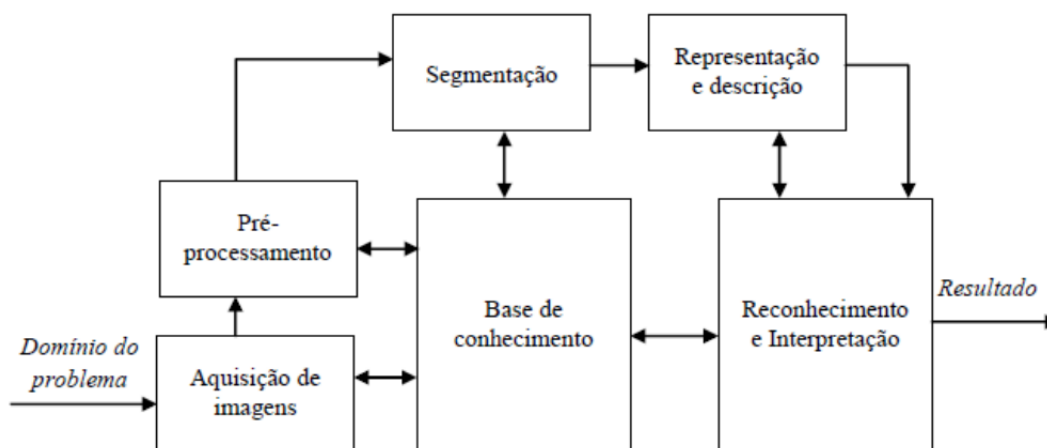


Figura 3: Etapas do processamento digital de imagens. Adaptado de(GONZALEZ; WOODS, 2002).

resultar em informações para alimentar a etapa seguinte. Sendo que ao fim de todas as etapas, o resultado pode ser ou não representado por uma imagem digital.

A primeira etapa desse processo é a aquisição de imagens, onde um digitalizador converte a imagem analógica para digital. Assim, as imagens são capturadas e representadas de forma computacional para serem interpretadas em etapas posteriores.

O pré-processamento consiste na segunda etapa do processo. Essa etapa busca a melhoria dos aspectos visuais e estruturais da imagem, objetivando aumentar as chances de sucesso dos processos seguintes. Nessa etapa podem ser aplicadas técnicas de realce e melhoramento de imagem, por exemplo: remoção de ruído, filtros morfológicos, dentre outras.

A segmentação, por sua vez, trata de extrair da imagem apenas partes que realmente interessam para o processamento, dividir a imagem em regiões disjuntas, permitindo o isolamento de regiões de interesses (ROI). A segmentação é o processo de particionamento de regiões onde todos os elementos de uma região devem ser o mais homogêneo possível, enquanto regiões diferentes devem ser mais heterogêneo (SOUSA et al., 2011). O processo de segmentação pode ser abordado de três formas: manual, semi-automática e automática.

A quarta etapa é a representação e descrição, também conhecida como extração de características. Tendo como objetivo representar, descrever, através de valores, uma imagem ou regiões da mesma. Cada medida extraída de uma determinada área compõe um vetor de características que define um padrão para a determinada área.

E como última etapa, tem-se o reconhecimento de padrões, onde é atribuído um rótulo baseado nas características extraídas na etapa anterior no intuito de diferenciar possíveis padrões contidos em um grupo de imagens, possibilitando a classificação da imagem ou partes dela.

2.2.1 Textura

A extração de características de uma imagem objetiva determinar aspectos que resultem em alguma informação quantitativa de interesse ou que sejam básicas para discriminação entre classes distintas. Como afirma Pedrini e Schwartz (2008), a complexidade na análise de imagens está na definição de um conjunto de características que possa descrever de forma real cada região contida na imagem, de modo a serem utilizadas em processos de alto nível.

Como visto em Gonzalez e Woods (2002), a extração de características em imagens digitais são compreendidas por duas categorias de análises: textura e forma. Uma análise por textura busca descrever suavidade, rugosidade e regularidade da região de uma imagem. Enquanto uma análise por forma mensura informações da configuração e da aparência externa da região avaliada (FILHO et al., 2016).

Segundo Haralick et al. (1973), textura é definida como a informação de uma região relacionada a coeficientes de uniformidade, densidade, aspereza, regularidade, intensidade, dentre outras características da imagem. A análise de textura é relevante em imagens digitais, uma vez que possibilita distinguir regiões da imagem que apresentam características semelhantes (AZEVEDO; CONCI; VASCONCELOS, 2018).

Nas próximas seções serão apresentadas as fundamentações dos descritores de textura baseados em Índices Taxonômicos, utilizados neste trabalho para a extração de características de textura.

2.3 Índices de Diversidade

Uma comunidade pode ser definida como um conjunto de espécies que ocorrem em um determinado lugar e tempo (MAGURRAN, 2013). Diversidade, por sua vez, caracteriza a variedade de espécies presentes na comunidade. A diversidade é um termo muito utilizado na área da ecologia, podendo ser mensurada a partir de duas componentes: riqueza de espécies e a abundância relativa.

A riqueza de espécies consiste no número de espécies encontradas em determinada região. Já a abundância relativa avalia o número de indivíduos de uma determinada espécie existentes numa dada área (PIANKA, 1994). As medidas de diversidade de espécies são geralmente úteis para comparar padrões em diferentes áreas.

Porém, como visto em Clarke e Warwick (1998), diversidade não pode ser medida apenas com a utilização de dados como a abundância e a riqueza de espécies, carecendo de medidas que levem em conta questões evolutivas e de parentesco. Nesse aspecto, as medidas de diversidade filogenéticas, incorporam as relações de parentesco e proximidade entre espécies, aumentando a descoberta de padrões sobre a diversidades em diferentes comunidades (MAGURRAN, 2013).

Dentro do contexto das relações filogenéticas, o conceito de árvores filogenética, ou simplesmente filogenia, é tido com uma árvore onde as folhas representam os organismos, ou espécies, e os nós internos representam supostos ancestrais (Figura 4). As arestas da árvore denotam as relações evolutivas (ARAÚJO, 2003).

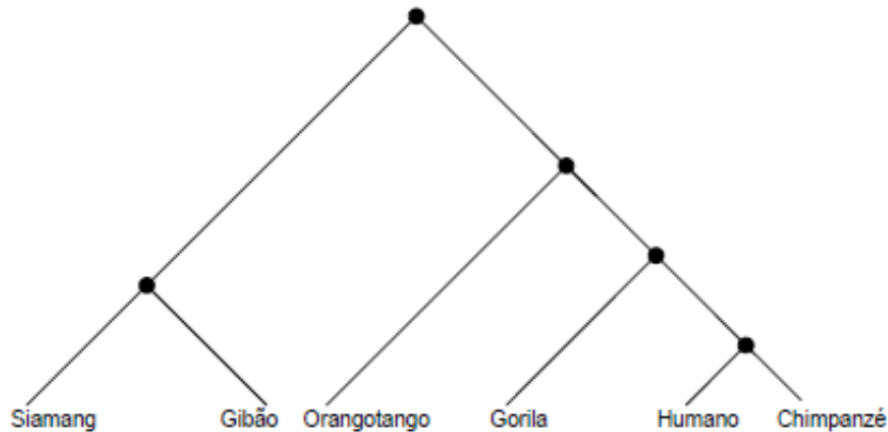


Figura 4: Exemplo de árvore filogenéticas de alguns primatas. Disponível em (ARAÚJO, 2003).

A combinação de abundância das espécies com a proximidade filogenética para gerar um índice de diversidade é denominada diversidade taxonômica (SILVA; BATALHA, 2006). A taxonomia, como ciência, lida com a classificação, identificação e nomenclatura das espécies, baseando se no conjunto das distâncias entre pares de espécies acumuladas a partir das árvores filogenéticas (VANDAMME et al., 1996).

A diversidade taxonômica é muito sensível a perturbações ambientais e apropriado para avaliar as diferenças entre comunidades (SILVA; BATALHA, 2006). Uma comunidade em que as espécies estão distribuídas em muitos gêneros deve apresentar uma diversidade maior que uma comunidade em que a maioria das espécies pertence a um mesmo gênero, dessa forma a análise da diversidade taxonômica proporciona o entendimento de padrões sobre determinada comunidade (RICOTTA, 2004).

A aplicação dos índices de diversidade ao processamento imagens, resulta na discriminação de características em uma imagem. Onde, a imagem ou região da mesma representa a “comunidade” e as “espécies”, nesse contexto, seriam os pixels, como visto em Sousa et al. (2011). A Figura 5 mostra um exemplo de uma região da imagem onde há diversidade de “espécies” em relação aos pixels. Assim, é possível a identificação de padrões em imagens, utilizando as relações taxonômicas.

2.3.1 Índices de Diversidade Taxonômica

O cálculo entre dois organismos escolhidos aleatoriamente em uma filogenia existente em uma comunidade é apresentado por índices de Diversidade Taxonômica (Δ) e Distinção Taxonômica (Δ^*) (CLARKE; WARWICK, 1998). Como afirma Filho et al. (2017b)

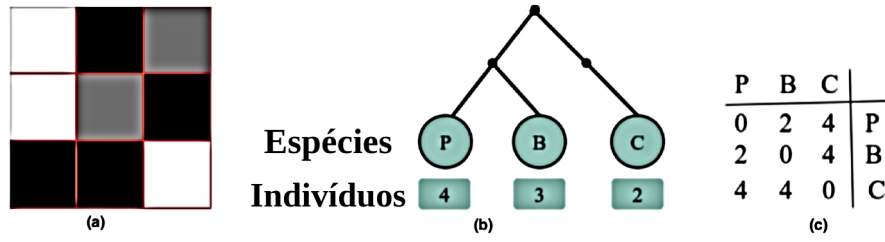


Figura 5: (a) Representação da região analisada (b) Árvore filogenética extraída (c) Matriz com distâncias levando em consideração a quantidade de arestas da árvore filogenética que as separa as espécies analisadas. Disponível em (CARVALHO et al., 2018).

Estes índices possuem três fatores essenciais para aplicação: número de espécies, número de indivíduos e estrutura de conexão das espécies (número de arestas). Neste trabalho são utilizados esses dois índices para descrever a textura de imagens como retina glaucomatosas e retinas normais.

O valor de Δ considera a abundância das espécies e a relação taxonômica entre elas. Portanto, seu valor expressa a distância taxonômica média entre dois indivíduos, escolhidos aleatoriamente de uma amostra (GORENSTEIN, 2009). Podendo ser definido pela equação 2.1, onde X_i é abundância da i -ésima espécie, N é o número total de indivíduos e ω_{ij} é a distância da espécie i à espécie j na classificação taxonômica.

$$\Delta = \frac{\sum \sum_{i < j} \omega_{ij} X_i X_j}{[N(N-1)/2]} \quad (2.1)$$

Já o Índice de Distinção Taxonômica (Δ^*) representa a distância taxonômica média entre dois indivíduos, com a restrição de que sejam de espécies diferentes (GORENSTEIN, 2009). Este índice é definido por 2.2 onde X_i ($i = 0, \dots, s$) é a abundância da espécie i , X_j ($j = 0, \dots, s$) representa a abundância da espécie j e ω_{ij} é a distância da espécie i à espécie j na classificação taxonômica.

$$\Delta^* = \frac{\sum \sum_{i < j} \omega_{ij} X_i X_j}{\sum \sum_{i < j} X_i X_j} \quad (2.2)$$

2.4 Classificação e Reconhecimento de Padrões

Em paralelo ao crescimento das tecnologias da informação, a complexidade dos problemas que essas tecnologias atendem também evoluem de forma igual, de modo que soluções estáticas não oferecem mais suporte a todos as necessidades inerentes a algumas situações, geralmente atreladas a problemas com um alto nível de abstração.

Para Kruse et al. (2016), a inteligência computacional se trata de “compreender conceitos, paradigmas, algoritmos e implementações para desenvolver sistemas que exibem comportamentos inteligentes em ambientes complexos”. Dessa forma, essa metodologia visa fornecer características, inerentes ao comportamento do aprendizado humano, dando

às máquinas a capacidade de resolução de problemas muito abstratos, caracterizando as mesmas como “sistemas inteligentes”.

A técnica de reconhecimento de padrões é uma sub-área da inteligência computacional cujo objetivo é classificar objetos tendo base um conhecimento prévio ou informações estatísticas extraídas dos padrões (OLIVEIRA et al., 2013). Segundo Looney et al. (1997), padrão é tudo aquilo para o qual existe uma entidade nomeável representante, entendido a partir de alguma forma de conhecimento.

O reconhecimento de padrões é um processo que procura a maior aproximação da classificação, buscando a melhor probabilidade de se estar certo. Assim, associando os conceitos de reconhecimento de padrões a resolução de problemas ligados à área de saúde, como afirma Netto et al. (2010), é possível analisar informações úteis ao diagnóstico, determinando a presença ou ausência de uma doença, dessa forma, aprimorando resultados dos diagnósticos.

2.4.1 Classificadores

Para a análise de padrões, na metodologia, foram escolhidas classificadores levando em consideração sua popularidade, eficácia, facilidade de adaptação nos experimentos e sua integração com a ferramenta de extração de características em imagens. Dentre o conjunto de candidatos, três se destacaram por apresentar boa aceitação em todos os aspectos analisados, sendo eles: *Multi Layer Perceptron* (MLP) *Sequential Minimal Optimization* (SMO) e *Random Forest*.

O MLP consiste em uma rede neural perceptron, formada por um conjunto de camadas. Este classificador tem em sua arquitetura uma camada de entrada, uma ou mais camadas ocultas, e uma camada de saída. A camada de entrada consiste em várias unidades que recebem entradas do mundo real, enquanto a camada de saída retorna os resultados para o mundo real, já as camadas ocultas são responsáveis por extrair padrões subjacentes das entradas (MONIKA; VENKATESAN, 2015).

O classificador SMO é um algoritmo para resolver o problema de programação quadrática que surge durante o treinamento de uma SVM. O SMO divide o problema de programação quadrática avaliado em uma série de menores problemas. Esses pequenos problemas são resolvidos analiticamente, o que permite que o SMO manipule conjuntos de treinamento muito grandes e possa distinguir padrões dentre os conjuntos de dados avaliados (PLATT, 1998).

Por fim, o *Random Forest* é uma combinação de preditores de árvores de decisão, de modo que cada árvore depende dos valores de um vetor aleatório amostrado de forma independente e com a mesma distribuição para todas as árvores na “floresta”. Cada árvore dá um voto indicando sua decisão sobre a classe do objeto. Assim, a classe com o maior número de votos é escolhida para o objeto (MA et al., 2018).

2.5 Validação

Para avaliar o desempenho de modelos de classificação, geralmente são utilizadas as seguintes estatísticas: Sensibilidade (S), Especificidade (E) e Acurácia (A). Essas métricas fazem uso da matriz de confusão, que indica a classificação correta ou incorreta das classes em uso, agrupando os resultados em quatro classes, sendo elas: Falso Negativo (FN), Falso Positivo (FP), Verdadeiro Positivo (VP) e Verdadeiro Negativo (VN).

A Sensibilidade (S), é a capacidade de um teste diagnóstico identificar os verdadeiros positivos nos indivíduos que apresentam doença. Tem a função de indicar a eficácia do método, conforme a Equação 2.3.

$$S = \frac{VP}{VP + FN}. \quad (2.3)$$

Especificidade (E), trata-se da capacidade de um teste de diagnóstico identificar os verdadeiros negativos nos indivíduos verdadeiramente saudáveis (Equação 2.4).

$$E = \frac{VN}{VN + FP}. \quad (2.4)$$

A Acurácia (A) calcula a proporção de acertos, ou seja, o total de verdadeiramente positivos e verdadeiramente negativos da amostra. Como demonstra a Equação 2.5.

$$A = \frac{VP + VN}{VP + VN + FP + FN}. \quad (2.5)$$

Além das métricas citadas anteriormente, uma maneira bastante utilizada em avaliação do desempenho quantitativo de uma determinada técnica ou modelo proposto é a análise da curva Receiver Operating Characteristic (ROC) (ERKEL; PETER, 1998). Como visto em Braga (2001), a área sob curva ROC (AUC) caracteriza-se como uma boa forma de avaliar um sistema de classificação binária, pois o mesmo implica em uma descrição empírica da capacidade do sistema de diagnóstico poder discriminar entre dois estados num universo, verificando-se o melhor desempenho, alcançado pelo método avaliado, quando o valor de AUC se aproxima de 1 (equivalente a 100%).

Por fim, como forma de avaliar o nível de concordância do experimento é usado o método estatístico de cálculo do Coeficiente Kappa (K) (COHEN, 1968). Por ser um dado conservador, o Coeficiente Kappa, é utilizado para avaliar a reprodutibilidade dos resultados em fase de validação, onde, quanto mais próximo 1 for o resultado maior será a concordância.

A Equação 2.6 demonstra o cálculo de K, onde P_o é a taxa de aceitação relativa e P_e é a taxa hipotética de aceitação.

$$K = \frac{P_o - P_e}{1 - P_e}. \quad (2.6)$$

A finalidade da utilização destas métricas é medir o desempenho da metodologia utilizada neste trabalho como satisfatório ou não. Com o intuito de apresentar pontos positivos e negativos que permita melhorias futuras deste trabalho.

3 Trabalhos Relacionados

A literatura disponível oferece estudos conhecidos sobre o problema da classificação automática do Glaucoma, bem como a aplicação de técnicas computacionais para extração de características. Nesta seção serão apresentados resumos desses trabalhos.

No sistema desenvolvido por Yu et al. (2005), são usadas técnicas de análise estatísticas *method of moments* para derivar características definidoras de imagens, distinguindo automaticamente imagens CSLT de nervos ópticos saudáveis e doentes. Nessa abordagem foi alcançado valores de acurácia de 86%, utilizando máquina de vetores de suporte (SVM) como classificador.

Foi proposto por Bock et al. (2007), um sistema automatizado de detecção de glaucoma usando imagens. Onde foram analisadas característica de intensidade de pixel usando *Linear Discriminant Analysis* (LDA), de texturas com Filtro de Gabor, características espectrais com *Fast Fourier Transform* (FFT) e comparações com parâmetros de modelos de histograma, visando capturar estruturas glaucomatosas e avaliar os resultados. O sistema proposto alcançou 86% de taxa de sucesso em um conjunto de dados contendo uma mistura de 200 imagens reais de olhos saudáveis e glaucomatosos.

No CAD proposto por Hatanaka et al. (2010), são utilizadas imagem do fundo da retina, calculando a relação entre o Copo e o Disco Óptico, com objetivo de detectar o glaucoma. Nessa abordagem foi atingida uma acurácia total de 84,5% de acurácia, com a sensibilidade de 80% e especificidade de 85%.

Em Acharya et al. (2011), foi apresentado um método para detecção de glaucoma usando uma combinação de textura e espectros de ordem superior (HOS) de imagens de fundo digital. Nesse metodologia de classificação foram usadas imagens fornecidas pela Faculdade de Medicina Kasturba Medical College, e os resultados obtidos apontam uma precisão de 91% de acurácia para a metodologia proposta.

No trabalho de Yadav, Sarathi e Dutta (2014) é proposto uma metodologia de classificação automática de Glaucoma, utilizando *Gray Level Co-occurrence Matrix* (GLCM) e aplicação de *Principal Component Analysis* (PCA). O método de classificação fornece um alto nível de acurácia, que tem como valor de acertos 90% de acurácia.

Na metodologia de Maheshwari, Pachori e Acharya (2017) e apresentado um método de diagnóstico automático do glaucoma. Para aplicação da metodologia foram usadas duas bases de imagens, a primeira privada com 60 imagens. A segunda uma base de domínio público, que conta com 500, adquiridas com o Grupo de análises de imagens médicas (MIAG). Durante a realização da extração de características, nessa metodologia, foram utilizadas as transformadas de *Wavelet* Empírico e *Features* de Correntropia, resultando em uma classificação com sensibilidade de 100%, especificidade de 96,67% e acurácia em 98,33%.

Junior et al. (2018) propõe a detecção do glaucoma através de descritores de textura baseados na diversidade filogenética. Nessa proposta é feita a segmentação do disco optico de maneira automática, usando Otsu e k-means, como metodologias. No estudo foram utilizadas a bases públicas de imagens Drishti-GS e RIM-ONE, e como resultados foi constatado o valor de 99,6% de acurácia durante a etapa de classificação das imagens.

Os trabalhos relatados acima são modelos de metodologias baseadas em características de textura e reconhecimento de padrões para auxílio ao diagnóstico do glaucoma. No intuito de aumentar a acurácia e melhorar os valores de sensibilidade e especificidade, este trabalho pretende explorar um potencial referente a medidas de descrição de textura baseada em índices de Diversidade Taxonômicas.

4 Método proposto

A metodologia proposta neste trabalho teve como objetivo realizar a classificação automática do glaucoma, consistindo em quatro etapas distintas, a primeira, a aquisição de imagens dos bancos de imagens Dristh-GS (SIVASWAMY et al., 2014) e RIM-ONE (TRUCCO et al., 2013); na segunda etapa é realizada a extração de características de textura utilizando os índices Taxonômicos, implementados neste trabalho com a biblioteca de manipulação de imagens OpenCV (TEAM, 2017); na terceira etapa, usa-se algoritmos disponíveis na ferramenta WEKA¹ para realizar a classificação das características extraídas; e por fim, são apresentadas as formas de validação dos resultados obtidos. A Figura 6 demonstra de maneira resumida as etapas seguidas na metodologia que será utilizada nesta proposta.

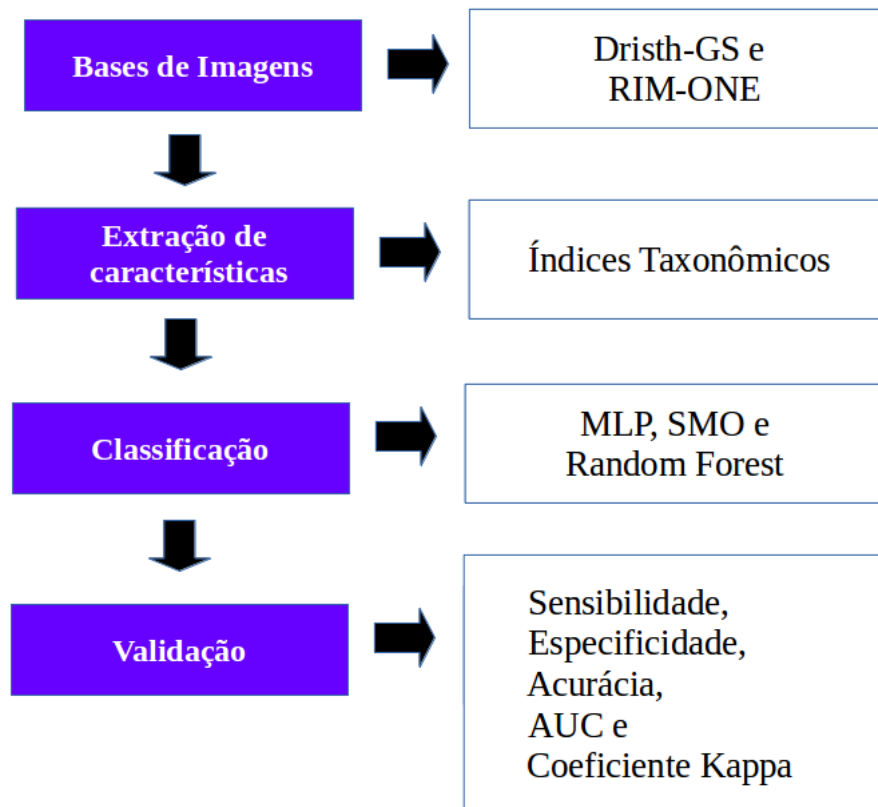


Figura 6: Metodologia proposta.

¹ <http://www.cs.waikato.ac.nz/ml/weka/>

4.1 Aquisição de Imagens

Durante a etapa de aquisição das imagens, na metodologia proposta, foram usadas duas bases de imagens públicas, sendo elas a Dristh-GS e RIM-ONE v2. Ambas as bases são formadas por um conjunto de imagens da retinas, contendo imagens com DO saudáveis e glaucomatosos, apresentando grande aceitação no meio científico e sendo utilizadas em diversas metodologias de análises e detecção do glaucoma.

O conjunto de dados Dristh-GS consiste em um total de 101 imagens, onde se verifica a presença de 70 imagens de olhos glaucomatosos e 31 imagem de olhos saudáveis. Todas as imagens, da base, foram coletadas no hospital de olhos Aravind, Madurai India. No qual, a seleção de pacientes com glaucoma foi feita por pesquisadores clínicos e os pacientes selecionados tinham entre 40 e 80 anos de idade, com um número aproximadamente igual de homens e mulheres. Cada imagem da base possui marcações manuais, obtidas de quatro especialistas, discriminatórias da região do DO e do Copo Ótico (SIVASWAMY et al., 2014).

A base de domínio público RIM-ONE v2, foi disponibilizada pelos hospitais Universitários de Canaria, a Clínica San Carlos e o hospital Universitário Miguel Servet. A base é composta por um total de 455 imagens de retina analisadas por especialistas. Sendo que, são 200 imagens de olhos com glaucoma em vários níveis e 255 sem presença do glaucoma. Cada imagem possui uma máscara correspondente, essa máscara foi utilizada para que o método proposto fosse aplicado apenas sob a região do disco óptico (TRUCCO et al., 2013). As Figuras 7 e 8 apresentam exemplos das imagens contidas nas bases de imagens utilizadas neste trabalho.

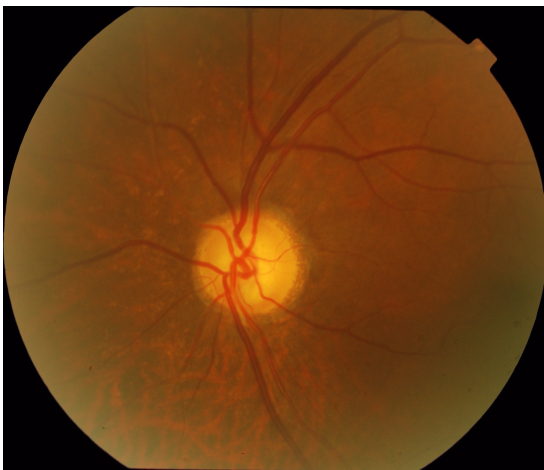


Figura 7: Imagem de retina sem glaucoma disponível na base DRISHTI-GS

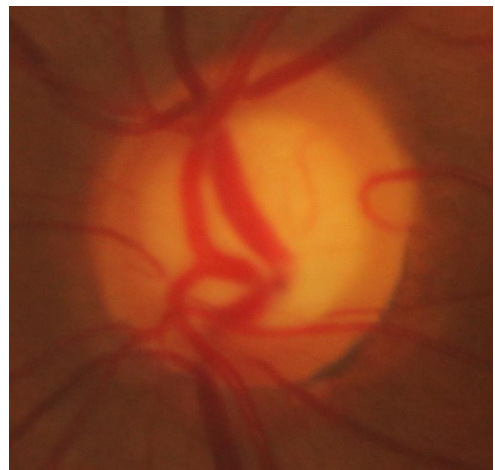


Figura 8: Imagem de retina sem glaucoma disponível na base RIM-ONE v2

4.2 Extração de Características

Embora as redes de aprendizagem profunda tenham mostrado-se promissoras na representação informações em visão computacional, tais modelos de aprendizagem profunda supervisionados, exigem grandes quantidades de dados rotulados e recursos de computação, que são caros (DU et al., 2017). Assim nessa proposta, empregamos uma metodologia tradicional de extração de características, aplicada a um conjunto de imagens de pequeno tamanho, visando demonstrar a eficiência da metodologia, aqui apresentada, na caracterização do glaucoma.

Após a etapa de aquisição de imagens e aplicação das marcações do DO feitas por especialistas médicos, dando continuidade às etapas presentes na metodologia, as imagens foram encaminhadas à fase de extração de características de textura. Para descrição da textura dos objetos foram utilizados os Índices de Diversidade Taxonômica e Distinção Taxonômica, como apresentados na Seção 2.3.1.

Os índices taxonômicos se baseiam na distância filogenética, calculada a partir do número de aresta entre duas folhas em uma determinada árvore. Em nossa proposta, para representar as imagens, utilizamos o modelo de árvore enraizada na forma de cladograma inclinado. A Figura 9 traz um exemplo dessa arquitetura de árvore, onde as espécies são os valores de intensidade dos pixels de um determinado canal e os indivíduos são pixels da imagem.

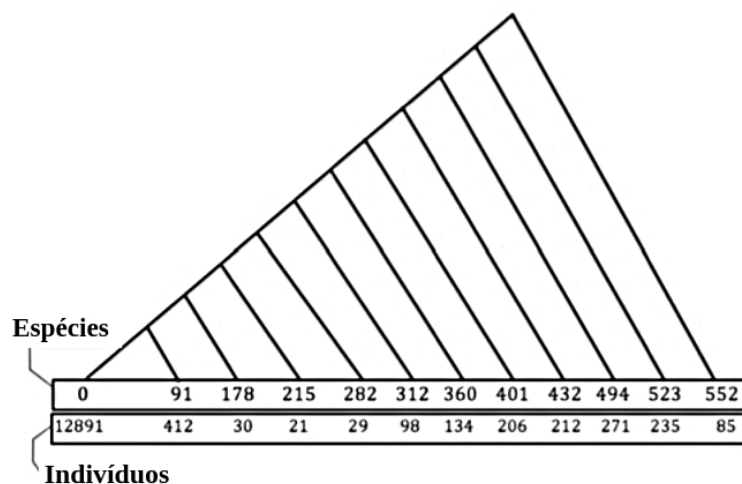


Figura 9: Exemplo de árvore enraizada na forma de cladograma inclinado. Disponível em (FILHO et al., 2017b).

Em cada ROI analisada, foram extraídos os 2 índices de diversidade taxonômica, porém como diferencial da presente metodologia foram aplicados os cálculo de Diversidade Taxonômica e Distinção Taxonômica para os pixels em níveis de cinza, como no trabalho de Filho et al. (2017b), e para cada um dos 3 canais RGB resultando em um total de 8 características por imagem. Ao final do processo, todos os índices aplicados às imagens

caracterizam a textura das mesmas e os atributos extraídos buscam fornecer medidas de homogeneidade ou heterogeneidade de cada ROI.

4.3 Classificação

O processo de classificação, na presente metodologia, foi realizado com classificadores presentes no pacote de software WEKA. Essa ferramenta, agrega um conjunto de algoritmos provenientes de diferentes abordagens voltadas para aprendizagem de máquina e mineração de dados. Também esse pacote de software permite o desenvolvimento de novos esquemas de aprendizado de máquina (HALL et al., 2009)

Dentre os classificadores presentes no pacote WEKA, foram selecionados: MLP, SMO e *Random Forest*. Esses classificadores foram escolhidos por apresentarem bons resultados, como visto na literatura, sendo rigorosos em relação a presença de ruídos no processo de classificação. O método utilizado para classificação foi o *k-fold*, com $k = 10$.

4.4 Validação

Após a finalização da etapa de reconhecimento de padrões, é necessário validar os resultados e discutir prováveis melhorias. Essa metodologia utiliza métricas comumente aplicadas a sistemas baseados em processamento de imagem. Para a validação dos resultados, foram utilizadas as métricas de avaliação baseadas em estatísticas: Sensibilidade (S), Especificidade (E), Acurácia (A), Área sob a curva ROC (AUC) e o Coeficiente Kappa (K).

5 Resultados e Discussões

Este capítulo mostra e discute os resultados obtidos com a metodologia proposta por esse trabalho, obtendo assim, a classificação das regiões de interesse de imagens do DO consideradas em glaucomatosas e normais. Este trabalho apresenta os testes realizados a partir das análises dos resultados provenientes da extração e características de cada canal RGB e níveis de cinza, além da combinação das características extraídas de todos os canais resultando em um vetor de 8 características.

Tabela 1: Testes com níveis de cinza

Classificador	A (%)	E (%)	S (%)	AUC	K
MLP	93,3	89,1	97,7	0,91	0,86
SMO	66,9	90,2	42,2	0,66	0,32
<i>Random Forest</i>	94,2	94	94,4	0,98	0,88

Tabela 2: Testes com o canal vermelho (Red)

Classificador	A (%)	E (%)	S (%)	AUC	K
MLP	93,5	88,4	98,8	0,92	0,87
SMO	83,8	89,1	78,1	0,83	0,67
<i>Random Forest</i>	96,7	97,2	96,20	0,99	0,93

Tabela 3: Testes com o canal verde (Green)

Classificador	A (%)	E (%)	S (%)	AUC	K
MLP	80	68,1	92,5	0,83	0,60
SMO	64,3	90,9	36,3	0,63	0,27
<i>Random Forest</i>	86,1	84,9	87,4	0,94	0,72

Tabela 4: Testes com o canal azul (Blue)

Classificador	A (%)	E (%)	S (%)	AUC	K
MLP	94,2	93,3	95,1	0,96	0,88
SMO	94,4	92,3	96,6	0,94	0,88
<i>Random Forest</i>	96,6	93	94,4	0,99	0,87

Tabela 5: Testes com o todos os canais e nível de Cinza

Classificador	A (%)	E (%)	S (%)	AUC	K
MLP	98,5	99,5	97,4	0,98	0,97
SMO	96	93,7	98,5	0,96	0,92
<i>Random Forest</i>	98,7	98,9	98,5	0,99	0,97

Os valores presentes nas Tabelas 1, 2, 3, 4 e 5 apresentam um resumo dos principais testes realizados, demonstrando o desempenho da metodologia utilizada para descrever características de textura utilizando três classificadores, MLP, SMO e *Random Forest*.

Nos experimentos descritos na Tabela 5 foi constatado que os melhores resultados apresentam valores acima de 98 % na taxa média de acertos, 98.5% para sensibilidade média e 98.9% para especificidade média. Sendo que o maior valor alcançado dentre todos os experimentos foi de 98,7% para acurácia, também apresentando os bons resultados de 0,99 para AUC e 0,97 para o coeficiente kappa, utilizando o classificador *Random Forest* juntamente com o índice de diversidade (Δ) e distinção taxonômica (Δ^*), aplicados a todos os canais de cores da imagem RGB.

Com base nos resultados obtidos, constata-se que o classificador *Random Forest* conseguiu os resultados mais promissores e equilibrados em todos os casos de testes na metodologia. Já o classificador SMO apresentou o menor resultado levando em consideração apenas o canal verde (Tabela 3), tendo o como pior valor para a A 64.3%.

A Tabela 6 apresenta uma breve comparação entre os resultados encontrados utilizando a metodologia propostas e alguns trabalhos citados no Capítulo 3, que realizam a classificação de regiões extraídas do DO em Glaucomatosas e Normais. Porém, algumas bases são diferentes da usada no presente trabalho. Portanto, os dados apresentados na Tabela 6, são apenas para ilustrar/comparar os resultados obtidos com os trabalhos da área, sem desmerecer nenhum dos métodos citados.

Tabela 6: Trabalhos relacionados

Trabalho	Bases de Imagens	Técnica	A (%)	E (%)	S (%)
Yu et al. (2005)	Base Privada	Análise estatística Method of moments	86	-	-
Bock et al. (2007)	Base Privada	linear discriminant analysis (LDA) , Filtro Gabor, Fast Fourier Transform (FFT) e comparações com parâmetros de modelos de histograma	86	0,83	0,88
Hatanaka et al. (2010)	Base Privada	Calculo da relação entre o Copo e o Disco Óptico	84,5	85	80
Acharya et al. (2011)	Imagens fornecidas pela Faculdade de Medicina Kasturba Medical College	Combinação de textura e espectros de ordem superior (HOS)	91	-	-
Yadav, Sarathi e Dutta (2014)	Base Privada	Gray Level Co-occurrence Matrix (GLCM) e <i>Principal Component Analysis</i>	90	-	-
Maheshwari, Pachori e Acharya (2017)	Base Privada e Base Pública MIAG	Transformadas de Wavelet e Features de Correntropia	98,3	96,6	100
Junior et al. (2018)	Dristh-GS e RIM-ONE	Segmentação usando Otsu e k-means e extração de características com índices de diversidade Filogenética.	99,6	99,3	100
Método Proposto	Dristh-GS e RIM-ONE	Índices de diversidade taxonômica.	98,7	98,9	98,5

No que diz respeito às métricas de desempenho utilizadas neste trabalho, o ideal para

um sistema CAD é ter um bom equilíbrio entre as três métricas de avaliação (acurácia, sensibilidade e especificidade), pois, um bom método deve ser capaz de classificar com sucesso, tantos os casos de positivos (que possuem a doença) como os casos de negativos (que não possuem a doença). Fica claro que na presente metodologia além dos valores de validação apresentarem equilíbrio, demonstram resultados robustos, motivando sua aplicação.

6 Conclusão

Este trabalho apresentou um método automático, com o uso dos Índices de Diversidade (Δ) e Distingão (Δ^*) Taxonômicos, junto com reconhecimentos de padrões, capaz de discriminar e classificar, em imagens, anormalidades em possíveis regiões suspeitas, resultando no auxílio ao diagnóstico de forma antecipada do glaucoma, sendo esta, uma segunda opinião para os especialistas.

Os índices de diversidade taxonômicos propostos neste trabalho, mostraram-se eficientes na tarefa de caracterização das regiões do Disco Óptico. Pois, foi possível alcançar resultados promissores, tendo valores expressivos em todas as métricas de avaliação dos resultados, com a taxa de sucesso acima de 98%.

Os resultados obtidos demonstraram o desempenho próspero das técnicas de extração de textura pelos índices taxonômicos em união com o classificador *Random Forest*. Os bons resultados na avaliação da acurácia, sensibilidade, especificidade, AUC e coeficiente de concordância kappa, fornecem bases sólidas para aplicação da técnica em ambientes reais.

Desta forma, este trabalho apresenta contribuições em duas áreas de estudo. Na área da saúde, oferecendo uma metodologia automática para auxílio no diagnóstico do glaucoma. Já na computação, as contribuições consistem na adaptação de técnicas de outras áreas do conhecimento, adequando-as no contexto da computação, especificamente, na área do processamento de imagens propondo novos descritores de textura baseados na diversidade taxonômica.

Assim, acredita-se que a metodologia apresentada neste trabalho poderá integrar uma ferramenta CAD, sendo aplicada em casos reais e atuais na detecção do glaucoma. Assim, trazendo benefícios ao especialista, que poderá contar com uma segunda opinião durante a detecção do glaucoma, bem como garantindo ao paciente melhor qualidade de vida e um prognóstico mais favorável quanto ao tratamento do Glaucoma.

Como trabalhos futuros, pretende-se:

- Aumentar a quantidade e variabilidade das amostras de lesões utilizando mais bases de imagens;
- Desenvolver um método automático para a segmentação do DO; e
- Aplicar índices taxonômicos para investigar outras doenças oculares.

Referências

- ACHARYA, U. R. et al. Automated diagnosis of glaucoma using texture and higher order spectra features. *IEEE Transactions on information technology in biomedicine*, IEEE, v. 15, n. 3, p. 449–455, 2011. Citado 2 vezes nas páginas 24 e 31.
- ARAÚJO, G. S. de. *Filogenia de Proteomas*. Tese (Doutorado) — Universidade Federal de Mato Grosso do Sul, 2003. Citado 2 vezes nas páginas 8 e 19.
- AZEVEDO, E.; CONCI, A.; VASCONCELOS, C. *Computação gráfica: Teoria e prática: geração de imagens*. [S.l.]: Elsevier Brasil, 2018. Citado na página 18.
- BOCK, R. et al. Classifying glaucoma with image-based features from fundus photographs. In: SPRINGER. *Joint Pattern Recognition Symposium*. [S.l.], 2007. p. 355–364. Citado 2 vezes nas páginas 24 e 31.
- BRAGA, A. Curvas roc: aspectos funcionais e aplicações. 2001. Citado na página 22.
- CARVALHO, E. D. et al. Method of differentiation of benign and malignant masses in digital mammograms using texture analysis based on phylogenetic diversity. *Computers & Electrical Engineering*, Elsevier, v. 67, p. 210–222, 2018. Citado 2 vezes nas páginas 8 e 20.
- CLARKE, K.; WARWICK, R. A taxonomic distinctness index and its statistical properties. *Journal of applied ecology*, Wiley Online Library, v. 35, n. 4, p. 523–531, 1998. Citado 2 vezes nas páginas 18 e 19.
- COHEN, J. Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological bulletin*, American Psychological Association, v. 70, n. 4, p. 213, 1968. Citado na página 22.
- DU, B. et al. Stacked convolutional denoising auto-encoders for feature representation. *IEEE transactions on cybernetics*, IEEE, v. 47, n. 4, p. 1017–1027, 2017. Citado na página 28.
- ERKEL, A. R. V.; PETER, M. Receiver operating characteristic (roc) analysis: basic principles and applications in radiology. *European Journal of radiology*, Elsevier, v. 27, n. 2, p. 88–94, 1998. Citado na página 22.
- FILHO, A. O. d. C. et al. Métodos para sistemas cad e cadx de nódulo pulmonar baseada em tomografia computadorizada usando análise de forma e textura. Universidade Federal do Maranhão, 2016. Citado na página 18.
- FILHO, A. O. de C. et al. Computer-aided diagnosis of lung nodules in computed tomography by using phylogenetic diversity, genetic algorithm, and svm. *Journal of Digital Imaging*, Springer, p. 1–11, 2017. Citado na página 13.
- FILHO, A. O. de C. et al. Lung-nodule classification based on computed tomography using taxonomic diversity indexes and an svm. *Journal of Signal Processing Systems*, Springer, v. 87, n. 2, p. 179–196, 2017. Citado 3 vezes nas páginas 8, 19 e 28.

- FILHO, O. M.; NETO, H. V. *Processamento digital de imagens*. [S.l.]: Brasport, 1999. Citado na página 16.
- FOSTER, P. J. et al. The definition and classification of glaucoma in prevalence surveys. *British journal of ophthalmology*, BMJ Publishing Group Ltd, v. 86, n. 2, p. 238–242, 2002. Citado 2 vezes nas páginas 15 e 16.
- GONZALEZ, R. C.; WOODS, R. E. *Digital image processing second edition*. *Beijing: Publishing House of Electronics Industry*, v. 455, 2002. Citado 3 vezes nas páginas 8, 17 e 18.
- GORENSTEIN, M. R. *Diversidade de espécies em comunidades arbóreas: aplicação de índices de distinção taxonômica em três formações florestais do Estado de São Paulo*. Tese (Doutorado) — Universidade de São Paulo, 2009. Citado na página 20.
- HALL, M. et al. The weka data mining software: an update. *ACM SIGKDD explorations newsletter*, ACM, v. 11, n. 1, p. 10–18, 2009. Citado na página 29.
- HARALICK, R. M. et al. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, v. 3, n. 6, p. 610–621, 1973. Citado na página 18.
- HATANAKA, Y. et al. Vertical cup-to-disc ratio measurement for diagnosis of glaucoma on fundus images. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. *Medical Imaging 2010: Computer-Aided Diagnosis*. [S.l.], 2010. v. 7624, p. 76243C. Citado 2 vezes nas páginas 24 e 31.
- JUNIOR, A. S. V. de C. et al. Automatic methods for diagnosis of glaucoma using texture descriptors based on phylogenetic diversity. *Computers & Electrical Engineering*, Elsevier, v. 71, p. 102–114, 2018. Citado 2 vezes nas páginas 25 e 31.
- JÚNIOR, A. S. V. de C. et al. Segmentação do disco óptico para detecção automática do glaucoma utilizando o algoritmo de otsu. Citado na página 13.
- KINGMAN, S. Glaucoma is second leading cause of blindness globally. *Bulletin of the World Health Organization*, SciELO Public Health, v. 82, p. 887–888, 2004. Citado na página 13.
- KRUSE, R. et al. *Computational intelligence: a methodological introduction*. [S.l.]: Springer, 2016. Citado na página 20.
- LEDERLIN, M. et al. Management strategy of pulmonary nodule in 2013. *Diagnostic and interventional imaging*, Elsevier, v. 94, n. 11, p. 1081–1094, 2013. Citado na página 13.
- LOONEY, C. G. et al. *Pattern recognition using neural networks: theory and algorithms for engineers and scientists*. [S.l.]: Oxford University Press New York, 1997. Citado na página 21.
- MA, J. et al. De-anonymizing social networks with random forest classifier. *IEEE Access*, IEEE, v. 6, p. 10139–10150, 2018. Citado na página 21.
- MAGURRAN, A. E. *Measuring biological diversity*. [S.l.]: John Wiley & Sons, 2013. Citado na página 18.

- MAHESHWARI, S.; PACHORI, R. B.; ACHARYA, U. R. Automated diagnosis of glaucoma using empirical wavelet transform and correntropy features extracted from fundus images. *IEEE journal of biomedical and health informatics*, IEEE, v. 21, n. 3, p. 803–813, 2017. Citado 2 vezes nas páginas 24 e 31.
- MELLO, A. A. P.; SUSANNA, J. R.; ALMEIDA, H. G. Glaucoma, 4. ed. *Rio de Janeiro: Cultura Medica*, p. 512, 2016. Citado 2 vezes nas páginas 13 e 15.
- MONIKA, P.; VENKATESAN, D. Di-ann clustering algorithm for pruning in mlp neural network. *Indian Journal of Science and Technology*, v. 8, n. 16, 2015. Citado na página 21.
- MUNIA, T. T. K. et al. Automatic diagnosis of melanoma using linear and nonlinear features from digital image. In: IEEE. *Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE*. [S.l.], 2017. p. 4281–4284. Citado na página 13.
- NETTO, S. M. B. et al. Segmentação automática de nódulos pulmonares com growing neural gas e máquina de vetores de suporte. Universidade Federal do Maranhão, 2010. Citado na página 21.
- NUGROHO, H. A. et al. Segmentation of optic disc on retinal fundus images using morphological reconstruction enhancement and active contour. In: IEEE. *Science in Information Technology (ICSITech), 2016 2nd International Conference on*. [S.l.], 2016. p. 362–366. Citado na página 15.
- NUNES, F. L. Introdução ao processamento de imagens médicas para auxílio a diagnóstico—uma visão prática. *Livro das Jornadas de Atualizações em Informática*, p. 73–126, 2006. Citado na página 13.
- OLIVEIRA, F. S. S. d. et al. Classificação de tecidos da mama em massa e não-massa usando índice de diversidade taxonômico e máquina de vetores de suporte. Universidade Federal do Maranhão, 2013. Citado na página 21.
- PEDRINI, H.; SCHWARTZ, W. R. *Análise de imagens digitais: princípios, algoritmos e aplicações*. [S.l.]: Thomson Learning, 2008. Citado na página 18.
- PIANKA, E. *Evolutionary Ecology. 5a Edição*. [S.l.]: Editora Harper Collins College Publishers, 1994. Citado na página 18.
- PLATT, J. Sequential minimal optimization: A fast algorithm for training support vector machines. 1998. Citado na página 21.
- QUIGLEY, H. A.; BROMAN, A. T. The number of people with glaucoma worldwide in 2010 and 2020. *British journal of ophthalmology*, BMJ Publishing Group Ltd, v. 90, n. 3, p. 262–267, 2006. Citado na página 13.
- RICOTTA, C. A parametric diversity measure combining the relative abundances and taxonomic distinctiveness of species. *Diversity and Distributions*, Wiley Online Library, v. 10, n. 2, p. 143–146, 2004. Citado na página 19.
- SILVA, I. A. D.; BATALHA, M. A. Taxonomic distinctness and diversity of a hyperseasonal savanna in central brazil. *Diversity and distributions*, Wiley Online Library, v. 12, n. 6, p. 725–730, 2006. Citado na página 19.

- SIVASWAMY, J. et al. Drishti-gs: Retinal image dataset for optic nerve head (onh) segmentation. In: IEEE. *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*. [S.l.], 2014. p. 53–56. Citado 2 vezes nas páginas 26 e 27.
- SOUSA, U. S. et al. Classificação de massas na mama a partir de imagens mamográficas usando índice de diversidade de shannon-wiener. Universidade Federal do Maranhão, 2011. Citado 2 vezes nas páginas 17 e 19.
- TEAM, O. D. *OpenCv: Open Source Computer Vision Library*. [S.l.]: Erişim Tarihi, 2017. Citado na página 26.
- TRUCCO, E. et al. Validating retinal fundus image analysis algorithms: Issues and a proposal validating retinal fundus image analysis algorithms. *Investigative ophthalmology & visual science*, The Association for Research in Vision and Ophthalmology, v. 54, n. 5, p. 3546–3559, 2013. Citado 2 vezes nas páginas 26 e 27.
- UMBAUGH, S. E. *Computer vision and image processing: a practical approach using cviptools with cdrom*. [S.l.]: Prentice Hall PTR, 1997. Citado na página 16.
- VANDAMME, P. et al. Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiological reviews*, Am Soc Microbiol, v. 60, n. 2, p. 407–438, 1996. Citado na página 19.
- YADAV, D.; SARATHI, M. P.; DUTTA, M. K. Classification of glaucoma based on texture features using neural networks. In: IEEE. *Contemporary Computing (IC3), 2014 Seventh International Conference on*. [S.l.], 2014. p. 109–112. Citado 2 vezes nas páginas 24 e 31.
- YU, J. et al. Automated optic nerve analysis for diagnostic support in glaucoma. In: IEEE. *Computer-Based Medical Systems, 2005. Proceedings. 18th IEEE Symposium on*. [S.l.], 2005. p. 97–102. Citado 2 vezes nas páginas 24 e 31.
- ZILLY, J.; BUHMANN, J. M.; MAHAPATRA, D. Glaucoma detection using entropy sampling and ensemble learning for automatic optic cup and disc segmentation. *Computerized Medical Imaging and Graphics*, Elsevier, v. 55, p. 28–41, 2017. Citado na página 15.



**TERMO DE AUTORIZAÇÃO PARA PUBLICAÇÃO DIGITAL NA BIBLIOTECA
“JOSÉ ALBANO DE MACEDO”**

Identificação do Tipo de Documento

- () Tese
- () Dissertação
- (X) Monografia
- () Artigo

Eu, **Tomaz Ribeiro Viana Bisneto**, autorizo com base na Lei Federal nº 9.610 de 19 de Fevereiro de 1998 e na Lei nº 10.973 de 02 de dezembro de 2004, a biblioteca da Universidade Federal do Piauí a divulgar, gratuitamente, sem ressarcimento de direitos autorais, o texto integral da publicação **Determinação automática do glaucoma utilizando índices de diversidade taxonômica** de minha autoria, em formato PDF, para fins de leitura e/ou impressão, pela internet a título de divulgação da produção científica gerada pela Universidade.

Picos-PI 11 de Dezembro de 2018.

Tomaz Ribeiro Viana Bisneto

Assinatura