

AVALIAÇÃO DO USO DE AGRUPAMENTO SEMISSUPERVISIONADO NA SEGMENTAÇÃO DO DISCO ÓPTICO PARA FINS DE DETECÇÃO DO GLAUCOMA

Luis Guilherme Teixeira dos Santos
Universidade Federal do Piauí/Brasil
luisguilherme.ufpi@gmail.com

Marcos Antonio de Sousa Frazão
Universidade Federal do Piauí/Brasil
frazaomarcos@yahoo.com.br

Rodrigo de Melos Souza Veras
Universidade Federal do Piauí/Brasil
rveras@ufpi.edu.br

Resumo: O glaucoma é uma doença lentamente progressiva, assintomática e que pode causar cegueira de forma irreversível. Dentre os principais exames utilizados por oftalmologistas para seu diagnóstico, a retinografia se destaca por ser o de menor custo e o menos invasivo. Esse exame consiste em registrar fotografias da retina e observá-las. O diagnóstico é realizado levando-se em consideração a razão entre os raios da parte externa do Disco Óptico (DO) e da parte interna, denominada escavação. Com o objetivo de segmentar essas regiões este trabalho avalia o desempenho de seis algoritmos de agrupamento semissupervisionado e dois supervisionado em 209 imagens de duas bases públicas. Os resultados demonstraram que o algoritmo $K\text{-means}_{ki}$ obteve, respectivamente, 99,59% e 97,71% de precisão na segmentação da escavação e contorno do DO.

Palavras-chaves: Imagens de retina; Glaucoma; Disco óptico; Agrupamento semissupervisionado; Segmentação.

Abstract: Glaucoma is a slowly progressive disease, asymptomatic and can cause irreversibly blindness. Among the main tests used by ophthalmologists for diagnosis the retinography is known for being the lowest cost and the least invasive. This exam consists of observe and record retinal photographs and the diagnosis is made taking into account the ratio between the radii outside the Optic Disc (OD) and the inner part, called cup. In order to segment these regions, this study evaluates the performance of six semi-supervised clustering algorithms and two supervised in 209 images of two public databases. The results demonstrated that the $K\text{-means}_{ki}$ algorithm obtained respectively 99.59% and 97.71% of precision in segmenting the contour of cup and OD.

Keywords: Retinal images; Glaucoma; Optic disc; Semi-supervised clustering; Segmentation.

I. INTRODUÇÃO

Recentes avanços na aquisição de imagens e poder computacional, tornaram possível usar informações presentes em imagens médicas de várias formas. Isso levou a um considerável interesse no desenvolvimento de sistemas automáticos de diagnóstico médico. Esses sistemas são chamados de sistemas de Diagnóstico Auxiliado por Computador (CAD - *Computer Aided Diagnosis*). Estudos sobre sistemas CAD mostram que eles podem auxiliar os médicos em diversas tarefas, como medir estruturas

anatômicas, monitorar as mudanças pela comparação de imagens sequenciais, diagnosticar e planejar o tratamento. Eles também aliviam o trabalho repetitivo, prevenindo erros por fadiga e aumentando a eficiência no trabalho. As aplicações de CAD mais estabelecidas em áreas médicas envolvem uso de sistemas automatizados em mamografia, tomografia computadorizada de tórax e radiografia [1].

A detecção automática da posição do disco óptico (DO) é um passo importante para a análise automática de imagens de retina. A correta detecção das coordenadas do centro e raio do DO, podem ser utilizadas para analisar a gravidade de algumas doenças, como o glaucoma. Aproximadamente 66,8 milhões de pessoas são portadoras de glaucoma primário de ângulo aberto e aproximadamente 10% desta população tornar-se-á cega de ambos os olhos devido a esta doença [2]. Por isso a necessidade de um diagnóstico mais rápido e preciso para a sua detecção em seus estágios iniciais [3].

Entre várias anormalidades da retina, glaucoma é uma das principais causas de perda de visão irreversível do mundo [4]. Ela envolve uma degeneração progressiva das fibras do nervo óptico. As mudanças estruturais na cabeça do nervo óptico ONH (*Optic nerve head*), na qual consiste em Disco Óptico (parte mais externa e que engloba a escavação) e Escavação (parte mais interna) da região, são uma maneira padrão para avaliar o dano ocasionado pelo glaucoma. Essas mudanças levam a um alargamento maior da escavação se comparado ao alargamento do disco óptico, e é capturado usando a relação do diâmetro da escavação e do disco (*cup-to-disk diameter ratio* - CDR). Um elevado valor do CDR indica a presença de glaucoma no paciente. A Figura 1 mostra a diferença marcada em um olho normal e com glaucoma, em que se reflete nos correspondentes valores do CDR.

Uma forma de auxiliar o diagnóstico do glaucoma, é a segmentação da imagem de retina e posteriormente o cálculo relacionado a proporção da Escavação e do contorno do DO, exemplificados na Figura 1. Segundo Póvoa et al. [5], quando essa proporção (Escavação/DO) chega a 0,7 suspeita-se que o paciente tenha glaucoma.

Segundo Muramatsu et al. [6] os três principais métodos utilizados para segmentação do DO são modelos de contorno ativos, redes neurais artificiais e algoritmos de agrupamento

(clustering). Partindo dessa ideia, analisamos algoritmos de agrupamento semissupervisionados, que são algoritmos de aprendizado que utilizam a combinação de dados rotulados e não rotulados. Assim, eles podem ser aplicados em domínios que se tem poucos exemplos rotulados (com classe definida) e uma vasta quantidade de exemplos não rotulados disponíveis. A grande motivação para se aplicar esse tipo de aprendizado neste problema, se dá pelo fato de exemplos não rotulados existirem em abundância e exemplos rotulados serem geralmente escassos [7]. Segundo Muramatsu et al. [6], o agrupamento não supervisionado pode ser vantajoso na minimização do efeito de *overtraining* (perda da capacidade de generalização).

II. MATERIAIS E MÉTODOS

A. Base de Imagens

O problema de segmentação do DO envolve dois desafios: a segmentação das bordas do Disco Óptico e da Escavação. Dentre as bases de imagens de retina públicas, apenas duas possuem os dados da verdade-terrestre: DRISHTI e RIM-ONE 3.

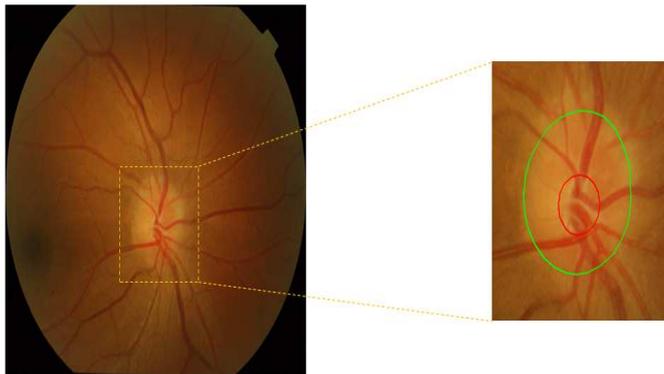
A base DRISHTI [8], é composta por 101 imagens de retina. Todas as imagens foram tiradas com centro no DO, com um "campo de visão" de 30° e dimensões 2896 x 1944 pixels. Apenas 50 desse total possuem informações sobre suas verdades terrestres, então somente elas foram utilizadas para testes.

A base RIM-ONE-3 [9] é um conjunto de dados que consiste em 159 imagens de retina, todas foram utilizadas, pois todas possuem informações sobre suas verdades terrestres. O disco óptico e a escavação foram marcados por 2 especialistas, a fim de criar a verdade terrestre. A segmentação média também está disponível para usar como referência.

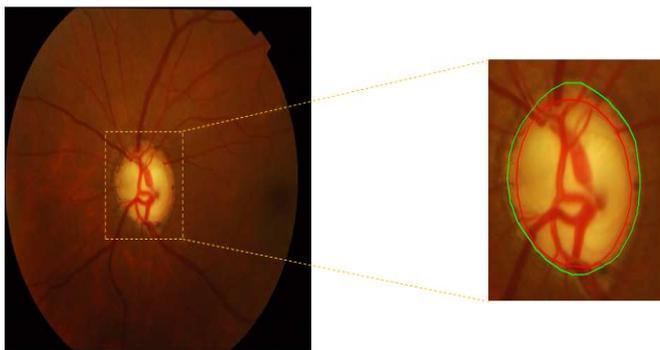
B. K-means

O algoritmo *K-means* é um método de agrupamento particional e supervisionado, primeiramente apresentado por [10]. O algoritmo *K-means* é largamente aceito como uma técnica "padrão" de detecção automática de classes pela análise de dados. Trata-se de um método baseado em heurística, que possui um grande poder de modelagem. O princípio de aglomeração consiste na determinação do modelo da distribuição dos exemplos no espaço atribuído. O *K-means* atinge esses objetivos de forma rápida e eficiente, utilizando um número fixo de classes iniciais, determinadas pelo usuário. Assim como todos os métodos particionais, o algoritmo *K-means* produz uma partição dos n objetos em K agrupamentos, geralmente otimizando uma função objetivo. Essa função objetivo geralmente é representada pelo somatório das distâncias entre os dados e o centro que o classifica, o que corresponde ao erro total absoluto. Existem algumas vantagens, como a possibilidade de mudanças de pertinência de objetos em relação a um agrupamento durante todo o processo de formação dos agrupamentos, e possibilidade de trabalhar com grandes bases de dados. Esse método requer tempo de ordem $O(n)$, onde n é o número de objetos no conjunto de dados [11].

O referido algoritmo é sensível à escolha das classes iniciais. O fato do número de classes ter que ser fixa e conhecida *a priori*, e de soluções diferentes poderem ser geradas dependendo do número K de classes escolhidas, constituem algumas desvantagens do *K-means*. Cada classe é representada por um ponto (centro) e, inicialmente, é formada somente por este. A escolha inicial correta do número de classes é um dos problemas mais fundamentais e ainda não solucionados. O problema de se escolher K de forma errada é que o método irá impor uma estrutura aos dados ao invés de buscar uma inerente a estes. A escolha dos centros deve ser feita dentro da base de dados, pelo usuário ou de forma aleatória.



(a) Exemplo de retina saudável.



(b) Exemplo de retina com glaucoma.

Fig 1. Exemplos de marcação do contorno do disco óptico (verde) e escavação (vermelho) feita por um especialista.

Na seção 2 temos os materiais e métodos utilizados nesse artigo. Na seção 3 apresentamos a metodologia de avaliação para cada método. Na seção 4 são apresentados e discutidos os resultados, bem como as bases de imagens utilizadas para avaliação do mesmo. Por fim, na seção 5 mostramos as conclusões e quais serão os trabalhos a serem feitos futuramente.

Após escolhidos os centros, cada dado é assinalado como pertencente à classe cujo centro está mais próximo. Após todos os dados serem processados, a posição média de cada classe é então calculada e usada como o novo centro, e os dados processados novamente. Esse processo é repetido até que não mais ocorram mudanças nos centros encontrados ou um número limite de iterações definido pelo usuário seja atingido. O algoritmo 1 mostra melhor o funcionamento do *K-means*.

Algoritmo 1 Algoritmo de aglomeração *K-means*

1. Inicialize K centros (w_1, \dots, w_k), representam as classes;
2. Cada classe C_j é associada a um conjunto w_j ;
3. Para cada dado de entrada x , assinale-o como pertencente à classe C_j caso w_j seja o centro mais próximo a este dado;
4. Para cada classe C_j , atualize os centros w_j , através da medida de todas as amostras de dados assinaladas como pertencentes a cada classe C_j ;
5. Calcular a função de erro E

$$E = \sum_{j=1}^k \sum_{x \in C_j} \|x - w_j\|^2;$$

6. Repetir passos 3, 4 e 5 até que o erro E não mude significativamente.
-

C. *K-means Fuzzy*

Como foi dito, o algoritmo *K-means* é um algoritmo de aprendizagem competitiva onde o vencedor é o único a ser atualizado a cada iteração. Este tipo de aprendizagem competitiva é denominada de competição dura (*hard competition*). Em outros algoritmos, como a rede SOM, não só neurónio vencedor tem seu protótipo atualizado, mas também os protótipos de seus vizinhos sofrem modificações. Este tipo de aprendizagem competitiva é denominado de competição suave (*soft competition*). Entretanto, independente do tipo de competição, um vetor de atributos só pode pertencer a um único grupo, que é aquele representado pelo protótipo mais próximo.

Já no caso das técnicas de agrupamentos nebulosos (*fuzzy clustering*), os dados podem possuir características que permitam que eles pertençam a diversos grupos com uma intensidade controlada por uma função de pertinência [11]. Dessa forma, a principal diferença do *K-means Fuzzy* para a sua versão clássica está no cálculo do erro, que é dado pela equação 1.

$$E = \sum_{j=1}^k \sum_{x \in C_j} \|x - w_j\|^2, \quad (1)$$

onde u_j é uma função que determina o grau de pertinência de um vetor x ao agrupamento representado pelo protótipo w_j e z representa o grau de nebulosidade da função.

D. *Seeded K-means*

Esse algoritmo, proposto por Basu et al. [7], é uma variação do *K-means*, e também particiona o conjunto de dados em k clusters. A diferença mais característica, é o fato do *Seeded K-means* utilizar exemplos inicialmente rotulados como os centroides iniciais dos clusters, isto é, as sementes (*Seed*, em inglês), e não os escolher aleatoriamente, tornando-o assim como um algoritmo semissupervisionado. Uma exigência do algoritmo é que para cada cluster seja atribuído, no mínimo uma semente.

Dados um conjunto de exemplos E , toma-se um subconjunto $S \subset E$ como sendo o conjunto de sementes. Na inicialização do algoritmo, o usuário é responsável por atribuir cada exemplo $x_i \in S$ a um dos k clusters a serem encontrados, dividindo o conjunto S em k subconjuntos S_i , de tal forma que $S = \bigcup_{i=1}^k S_i$.

E. *Constrained K-means*

Assim como o *Seeded K-means*, o *Constrained K-means* também tem na inicialização dos centroides a utilização de sementes e é um algoritmo semissupervisionado. A modificação principal feita por Wagstaff et al. [13] ao *Seeded K-means* foi no momento da atualização dos centroides, o algoritmo garante que nenhuma das restrições antes especificadas são violadas, ou seja, o elemento que faz parte do conjunto do cálculo dos centroides iniciais não pode ter a classe que lhe foi dada alterada, garantindo assim que tais elementos não serão rotulados erroneamente.

F. *K-means_{ki}*

O algoritmo *K-means_{ki}* é uma abordagem, também semissupervisionada, baseada no algoritmo *Seeded K-means* proposta em Sanches [14]. Assim como o *Seeded K-means*, o *K-means_{ki}* também se utiliza de exemplos rotulados, assim a sua diferença se dá no processo de *clustering* propriamente dito. Quando o *Seeded K-means* é utilizado, cada elemento é associado ao cluster (centroide) mais próximo. No caso do *K-means_{ki}* é estipulado a priori um *threshold* (limiar) t . Esse *threshold* será o responsável pela associação amostra/cluster. Ou seja, a amostra somente poderá ser associada a um dado cluster caso esteja a uma distância menor ou igual a t de seu respectivo centroide. Esse valor de t não é um valor absoluto, mas sim relativo. O valor absoluto do *threshold* é baseado em um vetor ordenado v contendo todas as distâncias entre todos os centroides (exemplos rotulados) e todos os exemplos não rotulados.

III. METODOLOGIA DE AVALIAÇÃO

Para avaliar o desempenho dos algoritmos foram utilizadas as métricas de Sobreposição, Sensibilidade, Especificidade, Acurácia, Precisão e *F-Score*. Todas essas métricas são calculadas com base nos valores Verdadeiro Positivo (VP), Falso Positivo (FP), Falso Negativo (FN) e Verdadeiro Negativo (VN) e são representados nas Equações de 1 a 6.

A taxa de VP refere-se aos pixels corretamente informados como positivos, FP refere-se aos pixels incorretamente informados como positivos, FN refere-se aos pixels

incorretamente informados como negativos, e por fim VN corresponde aos pixels corretamente informados como negativos.

A Sobreposição é a interseção da região segmentada com a real, a Sensibilidade representa a proporção de verdadeiros positivos, ou seja, avalia a capacidade de o algoritmo prever que um elemento pertence a um grupo quando ele realmente pertence a tal. A Especificidade reflete a proporção de verdadeiros negativos, isto é, avalia a capacidade de o algoritmo prever que um elemento não pertence a um grupo dado que ele realmente não pertence. A Acurácia é a proporção de predições corretas, sem considerar o que é positivo e o que é negativo e sim o acerto total. Por fim a *F-Score* é a média ponderada da Precisão e Sensibilidade.

$$\text{Sobreposição} = \frac{VP}{VP + FP + FN}, \quad (1)$$

$$\text{Sensibilidade} = \frac{VP}{VP + FN}, \quad (2)$$

$$\text{Especificidade} = \frac{VN}{VN + FP}, \quad (3)$$

$$\text{Acurácia} = \frac{VP + VN}{VP + FP + FN + VN}, \quad (4)$$

$$\text{Precisão} = \frac{VP}{VP + FP}, \quad (5)$$

$$\text{F-score} = \frac{\text{Precisão} * \text{Sensibilidade}}{\text{Precisão} + \text{Sensibilidade}}, \quad (6)$$

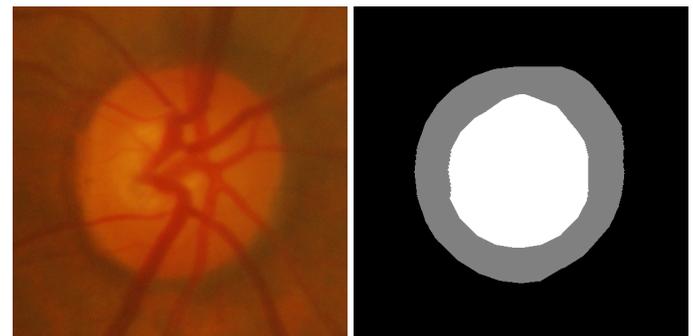
Cada algoritmo foi executado 5 vezes, e em cada inicialização as sementes utilizadas foram diferentes. No grupo dos rotulados foi usado 1% dos dados da imagem original como sementes iniciais, e 99% nos conjuntos dos não rotulados. Particularmente no *K-means_{ki}* e em sua versão *fuzzy*, utilizamos um limiar igual a 0,5%.

Em todas as execuções, foi utilizado um vetor de atributos como entrada para os algoritmos. Esse vetor contém o valor dos componentes: **R** (*red* - vermelho), **G** (*green* - verde), **B** (*blue* - azul), **H** (*hue* - matiz), **S** (*saturation* - saturação), **V** (*value* - valor do brilho), **I** (*intensity* - intensidade) e **Luv** de cada pixel da imagem.

IV. RESULTADOS E DISCUSSÃO

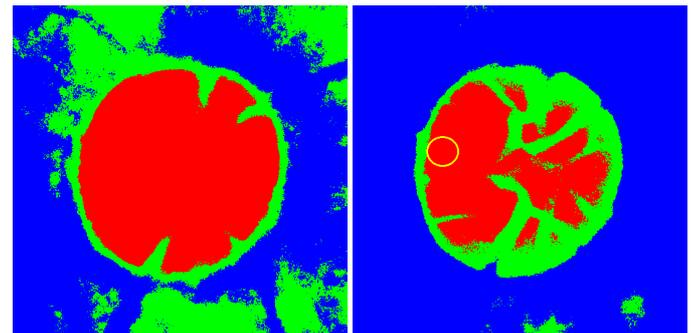
A Figura 2(a) mostra um exemplo de imagem contida na base DRISHTI, bem como sua verdade terrestre na Figura 2(b). As Figuras 2(c) a 2(f) apresentam os resultados das segmentações feitas em uma das imagens da base. Os pixels marcados de vermelho correspondem a região da Escavação, os marcados de verde correspondem ao Disco Óptico e os que foram marcados de azul correspondem ao fundo. Especialmente nas Figuras 2(f) e 2(j), a cor preta é a região em que não foi possível rotular, já que o *K-means_{ki}* e *K-means_{ki Fuzzy}* só agrupam um dado elemento quando tem um grau de certeza (definido previamente), ou seja, se ao tentarem agrupar algum elemento e o mesmo não estiver dentro do limiar, ficará sem rótulo.

Como já mencionado, o diagnóstico do glaucoma depende da segmentação de duas regiões: disco óptico e escavação. A Figura 2(b) apresenta uma segmentação ideal. Dessa forma, analisamos o resultado das aplicações dos 8 algoritmos para a detecção das duas regiões. Nas Figuras 2(d), 2(e), 2(h) e 2(i), podemos ver a diferença entre as imagens geradas após o agrupamento feito pelo *Seeded K-means*, *Constrained K-means*, *Seeded K-means Fuzzy* e *Constrained K-means Fuzzy*, que a olho nu fica imperceptível a sua distinção. Por isso, em cada uma das quatro imagens foi feita uma marcação (círculo) no mesmo ponto exemplificando as diferenças de pixels referentes aos centroides iniciais que não poderiam mudar de classe.



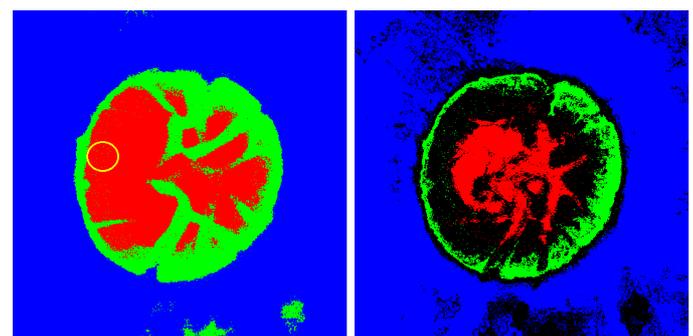
(a) Imagem de retina

(b) Verdade Terrestre



(c) Resultado do *K-means*

(d) Resultado do *Seeded K-means*



(e) Resultado do *Constrained K-means*

(f) Resultado do *K-means_{ki}*

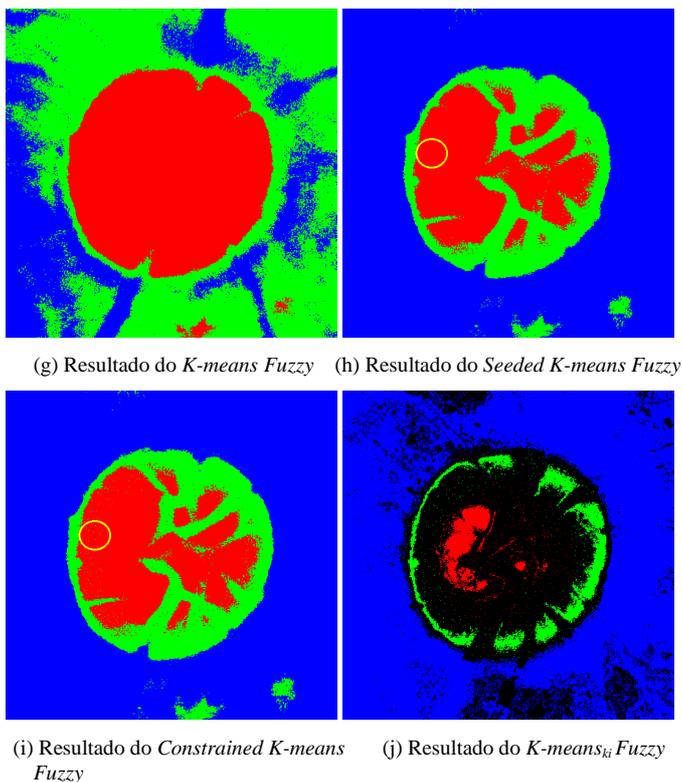


Fig 2. Resultado da segmentação: a) imagem original, b) verdade terrestre, c-j) resultados dos algoritmos avaliados.

As Tabelas I e II apresentam os resultados da avaliação dos algoritmos nas bases DRISHTI e RIM-ONE 3, respectivamente. Como pode ser visto, os algoritmos *K-means_{ki}* e *K-means_{ki} Fuzzy* se sobressaíram no cálculo das taxas de Especificidade, Acurácia e Precisão. Já nas taxas de Sobreposição, Sensibilidade e *F-Score*, os melhores resultados ficaram divididos entre o *K-means* e principalmente o *Constrained K-means*. Na base RIM-ONE 3, o *K-means_{ki}* se destacou na Especificidade, Acurácia e Precisão, justamente nos três parâmetros de desempenho que melhor demonstram quanto o algoritmo acertou, pois a Especificidade é basicamente o cálculo da taxa de acerto referente àqueles grupos, nesse caso são dois (Escavação e DO), a Acurácia é a proporção de acertos no agrupamento geral de todas as classes (classificados corretamente / total de exemplos), enquanto a Precisão é a porcentagem de acerto na classificação das regiões corretas sobre o total de amostras classificadas corretamente.

Já na base DRISHTI, esses valores ficaram divididos entre *K-means_{ki}*, que obteve melhor taxa de acurácia, e *K-means_{ki} Fuzzy*, que se sobressaiu nas taxas de especificidade e precisão.

A base RIM-ONE 3 forneceu a informação de quais imagens possuem glaucoma e quais não possuem. Dessa forma, as Tabelas III e IV apresentam o resultado da avaliação dos algoritmos apenas das imagens saudáveis e apenas em imagens com glaucoma. Novamente o algoritmo *K-means_{ki}* obteve destaque se comparado aos outros métodos. É mostrado esse domínio tanto nas amostras saudáveis como nas patológicas. Mais uma vez é enfatizado o seu acerto na segmentação de tais regiões, ratificando assim a sua alta capacidade em agrupar elementos corretamente.

Como vimos nas tabelas citadas anteriormente, os algoritmos obtiveram melhores resultados no agrupamento de imagens com glaucoma. Essa vantagem se dá justamente porque o contraste da escavação em relação ao restante do disco óptico é menor nas imagens saudáveis, enquanto nas imagens com glaucoma há mais contraste, o que colabora no processo de diferenciação das regiões.

Tabela I. Avaliação de desempenho dos métodos na base DRISHTI.

Métodos	Taxas	Escavação				Contorno do DO			
		Max (%)	Min (%)	Média (%)	Desvio (%)	Max (%)	Min (%)	Média (%)	Desvio (%)
<i>K-means</i>	Sobreposição	30,93	11,68	26,93	0,0853	26,49	19,17	24,62	0,0308
	Sensibilidade	77,91	32,92	68,57	0,1993	45,92	29,76	42,22	0,0698
	Especificidade	82,08	64,44	78,41	0,0781	69,1	62,18	67,2	0,0289
	Acurácia	81,54	62,12	77,53	0,0862	64,41	55,88	62,23	0,0359
	Precisão	35,56	14,01	30,93	0,0946	34,39	27,01	32,33	0,0302
	F-Score	47,18	19,02	41,33	0,1248	39,32	28,32	36,56	0,0463
Seeded <i>K-means</i>	Sobreposição	32,41	32,2	32,31	0,001	31,88	30,8	31,28	0,0046
	Sensibilidade	74,51	72,79	73,75	0,007	54,93	53,56	54,16	0,0056
	Especificidade	83,95	83,36	83,67	0,0029	74,93	74,18	74,55	0,0038
	Acurácia	83,08	82,5	82,81	0,0029	71,1	70,25	70,65	0,0038
	Precisão	38,07	37,83	37,96	0,0012	40,72	39,4	40,02	0,0063
	F-Score	49,02	48,48	48,75	0,0023	46,63	45,29	45,9	0,0059
Constrained K- <i>means</i>	Sobreposição	32,74	32,58	32,66	0,0007	32,55	31,72	32,1	0,0039
	Sensibilidade	73,95	72,8	73,46	0,006	32,55	31,72	32,1	0,0039
	Especificidade	84,16	83,87	84,03	0,0012	75,48	74,95	75,2	0,0023
	Acurácia	83,28	82,98	83,15	0,0013	71,7	71,08	71,37	0,0028
	Precisão	38,41	38,23	38,33	0,0007	41,44	40,57	40,98	0,0042
	F-Score	49,24	48,85	49,06	0,0018	47,43	46,44	46,89	0,0047
<i>K-means</i> _{ki}	Sobreposição	28,85	27,95	28,53	0,0037	21,94	21,17	21,57	0,0032
	Sensibilidade	28,92	28,01	28,59	0,0038	22,18	21,39	21,81	0,0033
	Especificidade	95,92	95,91	95,92	0	95,72	95,7	95,71	0,0001
	Acurácia	92,01	91,96	91,99	0,0002	82,41	82,28	82,36	0,0005
	Precisão	95,52	95,2	95,39	0,0013	91,35	90,8	91,08	0,0022
	F-Score	43,5	42,41	43,09	0,0045	35,47	34,43	34,97	0,0043
<i>K-means</i> Fuzzy	Sobreposição	57	50,99	52,8	0,0247	24,59	17,48	20,65	0,0275
	Sensibilidade	84,73	82,13	83,24	0,0116	45,27	36,7	40,72	0,0313
	Especificidade	88,13	86,17	86,85	0,0076	73,96	66,56	70,53	0,0263
	Acurácia	87,37	84,71	85,38	0,0112	68,49	61,2	65,03	0,0262
	Precisão	67,55	64,11	65,36	0,0138	33,61	23,59	27,75	0,0394
	F-Score	75,17	72,06	73,22	0,0129	38,58	28,72	32,97	0,0376
Seeded <i>K-means</i> Fuzzy	Sobreposição	61,08	60,69	60,81	0,0017	45,77	45,45	45,63	0,0012
	Sensibilidade	75,59	74,69	74,98	0,0039	68,71	68,17	68,48	0,0024
	Especificidade	94,92	94,77	94,85	0,0006	89,64	89,41	89,51	0,0008
	Acurácia	90,72	90,61	90,67	0,0005	85,69	85,57	85,64	0,0005
	Precisão	76,81	76,57	76,66	0,0009	58,48	58,08	58,25	0,0016
	F-Score	76,08	75,66	75,81	0,0018	63,09	62,79	62,95	0,0011
Constrained K- <i>means</i> Fuzzy	Sobreposição	61,15	61,03	61,09	0,0004	46,04	45,89	45,96	0,0006
	Sensibilidade	75,26	75,12	75,18	0,0006	68,92	68,64	68,81	0,0012
	Especificidade	94,96	94,89	94,93	0,0003	89,63	89,58	89,61	0,0002
	Acurácia	90,79	90,73	90,76	0,0002	85,81	85,76	85,78	0,0001
	Precisão	76,99	76,74	76,88	0,0011	58,64	58,52	58,57	0,0005
	F-Score	76,06	75,94	76,02	0,0005	63,36	63,21	63,28	0,0006
<i>K-means</i> _{ki} Fuzzy	Sobreposição	21,97	16,52	19,79	0,0223	11,48	9,78	10,35	0,0069
	Sensibilidade	22	16,54	19,81	0,0223	11,52	9,8	10,37	0,0069
	Especificidade	99,99	99,98	99,99	0	99,97	99,96	99,96	0
	Acurácia	84,81	83,54	84,04	0,0053	84,59	84,14	84,38	0,0016
	Precisão	99,73	99,23	99,59	0,0021	99,71	99,18	98,43	0,0023
	F-Score	36,05	28,36	33,01	0,0313	20,61	17,83	18,76	0,0112

Tabela II. Avaliação de desempenho dos métodos na base RIM-ONE 3.

Métodos	Taxas	Escavação				Contorno do DO			
		Max (%)	Min (%)	Média (%)	Desvio (%)	Max (%)	Min (%)	Média (%)	Desvio (%)
<i>K-means</i>	Sobreposição	55,20	53,97	54,57	0,0058	26,85	23,55	25,68	0,0125
	Sensibilidade	81,74	78,75	80,40	0,011	53,17	50,02	51,87	0,0114
	Especificidade	90,67	89,65	90,20	0,0036	73,75	71,67	72,93	0,008
	Acurácia	87,32	86,78	87,07	0,0025	69,83	67,46	68,63	0,0118
	Precisão	70,23	67,39	68,61	0,0103	33,32	29,86	32,4	0,0144
	F-Score	75,25	73,71	74,03	0,0023	40,97	37,4	39,88	0,0144
Seeded K-means	Sobreposição	53,27	53,27	53,27	0	33,55	33,14	33,29	0,0015
	Sensibilidade	72,81	72,54	72,69	0,0009	60,61	60,32	60,42	0,0011
	Especificidade	92,67	92,59	92,62	0,0003	78,77	78,35	78,48	0,0017
	Acurácia	87,53	87,49	87,51	0,0001	75,18	74,78	74,90	0,0016
	Precisão	73,04	73	73,02	0,0002	41,4	41,24	41,29	0,0007
	F-Score	72,87	72,83	72,86	0,0002	49,16	49,01	49,05	0,0006
Constrained K- means	Sobreposição	53,85	53,81	53,8	0,0001	33,91	33,83	33,87	0,0003
	Sensibilidade	73,11	73	73,1	0,0004	61,11	60,98	61,04	0,0005
	Especificidade	92,77	92,73	92,8	0,0001	79,03	78,94	78,98	0,0004
	Acurácia	87,74	87,73	87,7	0	75,47	75,41	75,43	0,0002
	Precisão	73,26	73,2	73,2	0,0002	42,11	42,01	42,06	0,0004
	F-Score	49,85	49,76	49,8	0,0004	73,16	73,13	73,23	0,0001
K-means _{ki}	Sobreposição	55,01	54,68	54,79	0,0013	27,27	26,98	27,09	0,0012
	Sensibilidade	55,09	54,77	54,88	0,0013	27,41	27,11	27,22	0,0012
	Especificidade	99,95	99,94	99,95	0	99,89	99,88	98,89	0
	Acurácia	92,05	91,98	92,03	0,0002	87,87	87,81	87,84	0,0002
	Precisão	99,71	99,48	99,59	0,0009	97,81	97,54	97,71	0,0011
	F-Score	70,95	70,68	70,77	0,0012	42,83	42,46	42,59	0,0016
<i>K-means Fuzzy</i>	Sobreposição	25,31	4,81	11,30	0,0825	15,20	2,45	6,26	0,0522
	Sensibilidade	83,00	16,46	38,23	0,2691	28,06	5,07	12,13	0,0932
	Especificidade	75,98	15,05	34,49	0,2463	64,92	11,98	28,87	0,2138
	Acurácia	76,27	15,15	34,69	0,2470	58,19	10,70	25,78	0,1918
	Precisão	28,96	5,41	12,69	0,0954	20,86	3,34	8,68	0,0707
	F-Score	42,94	8,14	19,05	0,1408	23,93	4,03	10,11	0,0804
Seeded K-means Fuzzy	Sobreposição	36,32	7,21	16,53	0,1180	41,06	8,35	18,83	0,1326
	Sensibilidade	67,20	13,30	30,63	0,2184	62,56	12,64	28,63	0,2024
	Especificidade	88,65	17,70	40,44	0,2876	83,80	16,80	38,30	0,2717
	Acurácia	87,23	17,42	39,79	0,2831	79,70	15,99	36,43	0,2583
	Precisão	43,43	8,62	19,78	0,1411	53,17	10,82	24,40	0,1717
	F-Score	52,76	10,46	24,04	0,1714	57,49	11,66	26,35	0,1858
Constrained K- means Fuzzy	Sobreposição	36,39	7,28	16,63	0,118	42,11	8,39	19,16	0,1369
	Sensibilidade	66,37	13,35	30,39	0,2148	63,85	12,68	29,09	0,2075
	Especificidade	88,71	17,73	40,53	0,2878	83,99	16,8	38,33	0,2725
	Acurácia	88,27	17,44	39,88	0,2831	80,09	16,01	36,54	0,2599
	Precisão	43,56	8,72	19,91	0,1412	54,26	10,82	24,68	0,1763
	F-Score	52,6	10,55	24,06	0,1704	58,67	11,68	26,7	0,1907
K-means _{ki} Fuzzy	Sobreposição	11,25	2,29	4,99	0,0373	8,44	1,61	3,91	0,0284
	Sensibilidade	11,27	2,3	5,03	0,0373	8,48	1,62	3,93	0,0285
	Especificidade	95,78	19,16	43,74	0,3107	95,7	19,14	43,7	0,3104
	Acurácia	90,05	18,06	41,15	0,292	79,74	15,92	36,41	0,2589
	Precisão	95,42	19,09	43,6	0,3095	91,7	18,39	41,91	0,2971
	F-Score	20,16	4,09	8,97	0,0666	15,53	2,98	7,18	0,052

Tabela III. Avaliação de desempenho dos métodos na base RIM-ONE 3 apenas imagem saudáveis.

Métodos	Taxas	Escavação				Contorno do DO			
		Max (%)	Min (%)	Média (%)	Desvio (%)	Max (%)	Min (%)	Média (%)	Desvio (%)
<i>K-means</i>	Sobreposição	20,45	8,18	17,93	0,0545	24,18	19,1	23,01	0,0221
	Sensibilidade	75,79	33,36	67,1	0,1886	41,96	29,06	39,06	0,0559
	Especificidade	76,56	61,71	73,5	0,0659	66,82	59,25	64,98	0,0322
	Acurácia	76,46	60,41	73,16	0,0713	61,76	53,17	59,75	0,037
	Precisão	22,28	9,04	19,43	0,0581	31,89	26,38	30,51	0,0234
	F-Score	34,4	14,23	30,14	0,089	36,24	27,67	34,22	0,0368
Seeded <i>K-means</i>	Sobreposição	22,15	21,92	22,04	0,001	30,33	28,87	29,46	0,006
	Sensibilidade	72,27	69,99	71,45	0,0087	49,32	47,53	48,08	0,0072
	Especificidade	78,55	77,96	78,24	0,0028	71,4	70,51	71	0,0042
	Acurácia	78,55	77,61	77,89	0,0028	66,96	65,88	66,38	0,0045
	Precisão	24,57	24,34	24,45	0,0011	39,48	37,84	38,61	0,0076
	F-Score	36,67	36,43	36,43	0,0022	43,86	42,14	42,83	0,0072
Constrained <i>K-means</i>	Sobreposição	22,38	22,16	22,28	0,0009	31,39	30,3	30,82	0,0049
	Sensibilidade	71,26	69,7	70,62	0,0081	50,52	49,12	49,74	0,0063
	Especificidade	78,73	78,26	78,53	0,002	72,52	71,63	72,04	0,0038
	Acurácia	78,35	77,86	78,14	0,002	68,04	67,08	67,54	0,0043
	Precisão	24,84	24,61	24,74	0,0009	40,9	39,78	40,34	0,0051
	F-Score	36,84	36,38	36,64	0,002	45,2	43,96	44,55	0,0057
<i>K-means</i> _{ki}	Sobreposição	18,81	17,99	53,97	0,0035	23,38	22,63	23,05	0,003
	Sensibilidade	18,85	18,02	78,75	0,0035	23,65	22,88	23,31	0,003
	Especificidade	92,39	92,38	92,39	0	92,14	92,13	92,14	0
	Acurácia	89,45	89,41	89,43	0,0001	78,42	78,3	78,37	0,0005
	Precisão	92,03	91,86	91,98	0,0007	88,02	87,65	87,86	0,0014
	F-Score	31,29	30,14	30,83	0,0048	37,27	36,32	36,84	0,0038
<i>K-means</i> Fuzzy	Sobreposição	18,14	3,62	8,08	0,0592	14,96	2,79	6,12	0,0507
	Sensibilidade	87,25	16,98	40,15	0,2841	26,22	5,27	11,27	0,0868
	Especificidade	77,53	15,57	35,23	0,2509	67,20	12,55	29,75	0,2215
	Acurácia	77,94	15,45	35,45	0,2522	59,08	11,11	26,05	0,1948
	Precisão	19,43	4,08	8,71	0,0628	21,30	3,76	8,78	0,0720
	F-Score	31,78	6,58	14,31	0,1028	23,50	4,39	9,86	0,0788
Seeded <i>K-means</i> Fuzzy	Sobreposição	28,61	5,65	13,00	0,0930	43,81	9,02	20,19	0,1410
	Sensibilidade	68,24	13,51	31,10	0,2217	63,41	12,85	29,05	0,2049
	Especificidade	91,71	18,25	41,77	0,2979	88,16	17,74	40,34	0,2855
	Acurácia	90,70	18,05	41,31	0,2946	83,12	16,75	38,04	0,2691
	Precisão	33,01	6,53	15,01	0,1074	58,32	12,05	26,95	0,1876
	F-Score	44,50	8,80	20,26	0,1447	60,76	12,44	27,96	0,1958
Constrained <i>K-means</i> Fuzzy	Sobreposição	28,71	5,74	13,13	0,093	45,03	9,06	20,54	0,1461
	Sensibilidade	67,52	13,52	30,92	0,2187	64,72	12,97	29,54	0,21
	Especificidade	91,7	18,33	41,89	0,2975	88,41	17,71	40,36	0,2868
	Acurácia	90,67	18,12	41,42	0,2941	83,59	16,74	38,16	0,2712
	Precisão	33,23	6,63	15,19	0,1078	59,55	12,03	27,17	0,1931
	F-Score	44,54	8,9	20,38	0,1444	62,03	12,48	28,3	0,2012
<i>K-means</i> _{ki} Fuzzy	Sobreposição	8,18	1,25	3,31	0,0286	10,33	1,79	4,59	0,0357
	Sensibilidade	8,2	1,25	3,32	0,0287	10,39	1,8	4,61	0,0359
	Especificidade	99,99	19,99	45,67	0,3244	99,89	19,98	45,62	0,324
	Acurácia	96,04	19,18	43,83	0,3116	81,96	16,32	37,39	0,2663
	Precisão	99,79	19,98	45,57	0,3237	95,23	19,14	43,64	0,3082
	F-Score	15,16	2,35	6,17	0,0528	18,73	3,28	8,33	0,0644

Tabela IV. Avaliação de desempenho dos métodos na base RIM-ONE 3 apenas com imagens glaucomatosas.

Métodos	Taxas	Escavação				Contorno do DO			
		Max (%)	Min (%)	Média (%)	Desvio (%)	Max (%)	Min (%)	Média (%)	Desvio (%)
<i>K-means</i>	Sobreposição	42,92	15,7	37,26	0,1206	29,14	19,26	24,47	0,0408
	Sensibilidade	80,34	32,41	70,26	0,2117	50,47	30,53	45,84	0,0858
	Especificidade	88,43	67,56	84,05	0,0922	71,72	65,54	69,76	0,025
	Acurácia	87,37	64,08	82,55	0,1033	67,46	59	65,09	0,0347
	Precisão	50,82	19,73	44,13	0,1365	37,25	27,74	34,42	0,0381
	F-Score	61,87	24,52	54,19	0,1659	42,86	28,07	39,24	0,0573
Seeded <i>K-means</i>	Sobreposição	44,18	44,01	44,1	0,0009	33,67	33,04	33,38	0,003
	Sensibilidade	77,08	76,01	76,4	0,005	61,37	60,5	61,14	0,0037
	Especificidade	90,14	89,55	89,91	0,003	78,98	78,39	78,62	0,0032
	Acurácia	88,69	88,12	88,47	0,0029	75,86	75,27	75,54	0,0031
	Precisão	53,58	53,33	53,48	0,0012	42,14	41,18	41,65	0,0049
	F-Score	63,19	62,68	62,9	0,0024	49,82	48,92	49,42	0,0044
Constrained <i>K-means</i>	Sobreposição	44,64	44,55	44,59	0,0004	33,88	33,35	33,56	0,0028
	Sensibilidade	77,04	76,36	76,74	0,0036	61,93	61	61,38	0,0049
	Especificidade	90,4	90,31	90,35	0,0004	75,9	75,66	75,77	0,0012
	Acurácia	88,94	88,85	88,9	0,0004	75,9	75,66	75,77	0,0012
	Precisão	54,01	53,94	53,94	0,0005	42,05	41,71	41,71	0,0031
	F-Score	63,48	63,17	63,34	0,0016	49,98	49,3	49,58	0,0037
<i>K-means</i> _{sk}	Sobreposição	40,37	39,39	40,06	0,004	20,28	19,5	19,88	0,0035
	Sensibilidade	40,49	39,48	40,17	0,0041	20,5	19,69	20,09	0,0036
	Especificidade	99,98	99,97	99,97	0	99,83	99,8	99,82	0,0001
	Acurácia	94,96	94,89	94,93	0,0003	86,99	86,85	86,93	0,0006
	Precisão	99,52	99,04	99,32	0,002	95,17	94,42	94,79	0,0031
	F-Score	57,52	56,51	57,18	0,0041	33,4	32,26	32,82	0,005
<i>K-means</i> Fuzzy	Sobreposição	35,93	6,62	16,07	0,117	17,15	2,29	7,02	0,059
	Sensibilidade	85,98	17,43	39,63	0,2774	32,83	5,32	14,27	0,1095
	Especificidade	81,38	15,89	36,91	0,2642	68,43	12,45	30,60	0,2252
	Acurácia	81,57	16,02	37,09	0,2645	62,66	11,26	27,91	0,2066
	Precisão	42,65	7,44	18,48	0,1418	22,34	3,17	9,38	0,076
	F-Score	57,02	10,43	25,18	0,1878	26,58	3,97	11,32	0,0898
Seeded <i>K-means</i> Fuzzy	Sobreposição	48,61	9,69	22,15	0,1578	41,78	8,36	19,04	0,1355
	Sensibilidade	72,35	14,32	32,92	0,2352	67,50	13,58	30,58	0,2187
	Especificidade	93,51	18,74	42,73	0,3031	86,72	17,31	39,58	0,2814
	Acurácia	91,49	18,34	41,8	0,2966	83,31	16,64	38,02	0,2703
	Precisão	59,50	11,84	27,13	0,1931	52,29	10,43	23,79	0,1698
	F-Score	65,30	12,96	29,75	0,2121	58,93	11,8	26,86	0,1912
Constrained <i>K-means</i> Fuzzy	Sobreposição	48,66	9,74	22,23	0,1578	45,03	9,06	20,54	0,1461
	Sensibilidade	71,31	14,43	32,66	0,2307	64,72	12,97	29,54	0,21
	Especificidade	93,68	18,73	42,81	0,304	88,41	17,71	40,36	0,2868
	Acurácia	91,62	18,32	41,87	0,2973	83,59	16,74	38,16	0,2712
	Precisão	59,54	11,94	27,21	0,193	59,55	12,03	27,17	0,1931
	F-Score	64,9	13,06	29,68	0,2102	62,03	12,48	28,3	0,2012
<i>K-means</i> _{sk} Fuzzy	Sobreposição	15,83	3,38	7,4	0,051	7,06	1,56	3,5	0,0227
	Sensibilidade	15,86	3,39	7,41	0,0511	7,09	1,57	3,52	0,0227
	Especificidade	99,99	19,99	45,66	0,3244	99,94	19,99	45,64	0,3242
	Acurácia	91,7	18,47	41,96	0,2971	84,72	16,97	38,73	0,2748
	Precisão	99,43	19,87	45,45	0,3224	96,32	19,27	43,88	0,3124
	F-Score	27,36	5,97	12,74	0,0882	13,21	2,91	6,51	0,0424

V. CONCLUSÃO E TRABALHOS FUTUROS

Com o objetivo de segmentar as regiões do disco óptico, implementamos oito algoritmos de agrupamento e utilizamos 209 imagens de retina como forma de avaliar seus desempenhos.

Verificamos que os algoritmos semissupervisionados avaliados obtiveram bons resultados. A grande razão para tal destaque é a utilização de sementes iniciais, fazendo com que os métodos usufruam da vantagem de começar o cálculo dos centroides com valores corretos em relação aos grupos. Isso trouxe precisão na rotulação dos elementos.

Em quase todos os algoritmos analisados o desvio padrão obtido foi baixo. Em alguns parâmetros, o $K\text{-means}_{ki}$ se sobressaiu e obteve desvio menor ou similar quando comparado ao restante, somente na Sobreposição, Sensibilidade e no $F\text{-Score}$ que ele não obteve bons resultados. Com isso, percebemos que esses métodos mantêm sempre uma saída semelhante, o que os torna algoritmos estáveis.

Já os outros algoritmos, que são as versões *Fuzzy*, apresentam um desvio padrão mais elevado. Esse efeito se dá principalmente pelo fato do cálculo do grau de pertinência depender das distâncias do elemento, então nesse caso o resultado bem mais dependente da inicialização dos centroides do que no caso dos algoritmos que não são versões *Fuzzy*.

Foi percebido que a presença dos vasos dificulta a diferenciação das regiões da Escavação e DO, visto que em algumas imagens esses vasos são segmentados ora como Escavação, ora com Disco Óptico. Os resultados satisfatórios do $K\text{-means}_{ki}$ e do $K\text{-means}_{ki}$ *Fuzzy* ocorreram justamente porque os dois algoritmos só agrupam o elemento quando eles estão com um grau de certeza dentro do limiar, ou seja, se as distâncias de um elemento que seja vaso estão muito confusas, tanto próximas aos centroides da Escavação como do DO, nenhum dos algoritmos rotula.

Em relação aos trabalhos futuros, pretende-se realizar um pré-processamento para todas as regiões segmentadas, tornando-as mais homogêneas, além de remover ruídos e os vasos presentes na imagem, e apresentar uma avaliação da localização das bordas das regiões obtidas.

REFERÊNCIAS

[1] Gonzalez, R. C. e Woods, R. E. Processamento Digital De Imagens. Pearson Education, 3 edição, 2011.

[2] Quigley, H. A. e Broman, A. T. (2006). The number of people with glaucoma worldwide in 2010 and 2020. British Journal of ophthalmology, 90:262–267.

[3] Dawn, S., Pearse, K., Adnan, T., Catherine, E., Lloyd, A., e Paolo, S. Automated retinal image analysis for diabetic retinopathy in telemedicine. Current Diabetes Reports, 2015.

[4] W.h.o. Prevention of blindness and deafness. global initiative for the elimination of avoidable blindness. In WHO document WHO/PBL/97.61 Rev2, 2000.

[5] Póvoa, C. A., Nicolela, M. T., Valle, A. L. S. L., de Siqueira Gomes, L. E., e Neustein, I. Prevalência de glaucoma identificada em campanha de detecção em são paulo. Arq Bras Oftalmol, 2001, 64:302–307.

[6] Muramatsu, C., Nakagawa, T., Sawada, A., Hatanaka, Y., Hara, T., Yamamoto, T., e Fujita, H. (2011). Automated segmentation of optic disc region on retinal fundus photographs: Comparison of contour modeling and pixel classification method. Computer Methods and Programs in Biomedicine, 101(1):23–32.

[7] Basu, S., Barnejee, A., e Mooney, R. Semi-supervised clustering by seeding. In Proceedings of the Nineteenth International Conference on Machine Learning, 2002, pp 19–26, Austrália

[8] Drishti-gs_retinal image dataset for optic nerve head (ONH) segmentation.readme. Disponível em: <http://www.midas3.kitware.com/midas/item/234748>. Acesso em 7 de julho de 2015.

[9] Pena-Betancor, C., Gonzalez-Hernandez, M., Fumero-Batista, F., Sigut, J., Medina-Mesa, E., Alayon, S., e de la Rosa, M. G. Estimation of the relative amount of hemoglobin in the cup and neuroretinal rim using stereoscopic color fundus images. Investigate Ophthalmology & Visual Science, 2015, 56(3):1562–1568.

[10] Macqueen, J. Some methods for classification and analysis of multivariate observations. In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, pp 17 – 33, University of California Press, 1967.

[11] J. A. F. Costa. “Classificação Automática e Análise de Dados por Redes Neurais Auto-Organizáveis”. Tese de doutorado, ‘Universidade de Campinas – UNICAMP’, 1999.

[12] X. Y. Wang, J. Garibaldi and T. Ozen. “Application of The Fuzzy C-Means Clustering Method on the Analysis of non Pre-processed FTIR Data for Cancer Diagnosis”. In in the Proceedings of the 8th Australian and New Zealand Conference on Intelligent Information Systems, pp. 233–238, 2003

[13] Wagstaff, K., Cardie, C., S, R., e Schroedl, S. Constrained k-means clustering with background knowledge. In Proceedings of the Eighteenth International Conference on Machine Learning, 2001, volume 18, pp 577–584.

[14] Sanches, M. K. Aprendizado de máquina semi-supervisionado: proposta de um algoritmo para rotular exemplos a partir de poucos exemplos rotulados. Dissertação de Mestrado, Universidade de São Paulo 2003.