

Universidade Estadual Paulista “Júlio de Mesquita Filho”
Pós-Graduação em Ciência da Computação

Bárbara Ribeiro da Silva

Segmentação Interativa de Imagens
utilizando Competição e Cooperação entre Partículas

UNESP
2015

Bárbara Ribeiro da Silva

Segmentação Interativa de Imagens
utilizando Competição e Cooperação entre Partículas

Orientador: Prof. Dr. Fabricio Aparecido Breve

Dissertação de Mestrado elaborada junto ao
Programa de Pós-Graduação em Ciência da
Computação - Área de Concentração em
Matemática e Inteligência Artificial,
como parte dos requisitos para a obtenção do
título de Mestre em Ciência da Computação

UNESP
2015

Bárbara Ribeiro da Silva

Segmentação Interativa de Imagens
utilizando Competição e Cooperação entre Partículas

Dissertação de Mestrado elaborada junto ao
Programa de Pós-Graduação em Ciência da
Computação - Área de Concentração em
Matemática e Inteligência Artificial,
como parte dos requisitos para a obtenção do
título de Mestre em Ciência da Computação

BANCA EXAMINADORA

Prof(a). Dr(a). Fabricio Aparecido Breve - Orientador
Universidade Estadual Paulista Júlio de Mesquita Filho

Prof(a). Dr(a). Denis Henrique Pinheiro Salvadeo
Universidade Estadual Paulista Júlio de Mesquita Filho

Prof(a). Dr(a). Marcos Gonçalves Quiles
Universidade Federal de São Paulo

UNESP
2015

Agradecimentos

A Deus, primeiramente por me proporcionar a vida, a família e amigos, dando todo o auxílio e amparo durante esta jornada.

A meus pais, Inez e Ronaldo, e a minha irmã Gabriela, pelo amor, carinho, exemplo e por me incentivarem e apoiarem todas as minhas decisões, por me darem toda a base necessária para que eu chegasse até aqui, me auxiliando a superar todos os obstáculos.

Ao meu namorado, Bruno, pelo amor, paciência, companheirismo, atenção e cuidado, juntamente com seus pais, por todo o suporte, auxílio, carinho e acolhimento.

Aos meus tios e primos, de sangue ou não, pelo carinho e compreensão nos momentos de ausência e por me incentivarem e apoiarem.

Aos amigos que me auxiliaram em todos estes anos presentes na universidade, pela amizade, pelo conforto, auxílio, pelas trocas de aprendizado, suporte e companheirismo. Em especial para aqueles que dividiram república: Tati, Helen, Mayra, July, Carol, Nath, Nati, Elisa, Marina, Lara, Mariana, Mariana, David e Bruno, que tiveram paciência e me ajudaram a chegar até aqui.

Ao meu orientador Fabrício, pela dedicação, paciência e incentivo que permitiram que eu pudesse finalizar este projeto.

A todos que direta e indiretamente me auxiliaram, superando todas as dificuldades, e compartilhando as alegrias conquistadas até aqui.

Resumo

Para estudar e classificar imagens é necessário entender as subdivisões que elas apresentam. Para esta finalidade existem técnicas de segmentação de imagens, uma das tarefas de visão computacional, a qual consiste na repartição de uma imagem digital em múltiplos segmentos, regiões ou objetos, com o intuito de facilitar a análise e compreensão. Métodos de segmentação de imagens são frequentemente utilizados para localizar objetos e bordas em imagens. Para tal tarefa, frequentemente são empregados algoritmos de aprendizado de máquina; disciplina científica que trata do projeto e desenvolvimento de algoritmos que melhoram automaticamente com a experiência, imitando o comportamento de aprendizado humano. O processo de aprendizado de máquina pode ser utilizado de diferentes formas, no entanto a categoria de aprendizado semi-supervisionado foi a que melhor se adaptou à proposta deste projeto. Esta categoria representa uma mescla entre o aprendizado supervisionado e o aprendizado não-supervisionado e trabalha com ambos os dados rotulados e não rotulados em sua fase de treinamento, proporcionando o uso de características individuais e/ou por conjunto. Diante da necessidade de combinação do aprendizado semi-supervisionado com a segmentação de imagem, foram estudadas formas de se utilizar informação provenientes de imagens para alimentar o aprendizado de máquina semi-supervisionado. Como resultado deste estudo, o algoritmo de aprendizado semi-supervisionado conhecido como competição e cooperação entre partículas foi estendido para a aplicações na tarefa de segmentação interativa de imagens. Na imagem de entrada, alguns pixels são classificados pelo usuário e servem como pixels de partida para o algoritmo. Estes pontos pré-rotulados são utilizados para criar as partículas no modelo utilizado. Também foram alterados alguns parâmetros de execução e partes estruturais do algoritmo de forma a otimizar a classificação e formas de expressar resultados com o objetivo de torná-lo mais adequado para a tarefa de segmentação de imagens.

Palavras- chave: Segmentação de imagens; Aprendizado de maquina; Aprendizado de máquina Semi-Supervisionado; Competição e cooperação entre partículas.

Abstract

To study and classify images it is necessary to understand the subdivisions they present. For this purpose there are image segmentation techniques, one of the computer vision tasks, which consists in the distribution of a digital image into multiple segments, regions or objects, in order to facilitate analysis and understanding. Image segmentation methods are often used for locating objects in images and borders. For this task, machine learning algorithms are often employed; scientific discipline that deals with the design and development of algorithms that automatically improve with experience, imitating the human learning behavior. The machine learning process can be used in different ways, however the category that best adapted to the project proposal, which represents a blend of supervised learning and unsupervised learning, called semi-supervised learning. Such category of algorithms works with both labelled and unlabelled data in their training phase, allowing the use of individual characteristics and or per set. Given the necessity of combining semi-supervised learning with image segmentation, some means of using information delivered from image to feed the semi-supervised learning machine were studied. As a result of this study, the semi-supervised learning algorithm known as competition and cooperation among particles was extended to be applied in the interactive image segmentation task. In the input image, some pixels are classified by the user and serve as a starting pixels to the algorithm. These pre-labeled points are used to create particles in the proposed model. Also, some parameters and algorithm structures were changed to optimize the classification and means of expressing the results in order to make it more suitable for the image segmentation task.

Keywords: Image segmentation; Machine learning; Semi-supervised learning; Particle Competition and Cooperation.

Lista de Figuras

2.1	Estilos de bordas (Gonzalez and Woods, 2010)	15
2.2	Deteção de bordas (Ma and Manjunath, 2000)	16
2.3	Deteção de bordas em imagens coloridas (Gonzalez and Woods, 2010)	17
2.4	Técnica de Limiarização (Glasbey and Horgan, 1994)	17
2.5	Crescimento de regiões (Garcia-Ugarriza et al., 2008)	19
2.6	Aplicação técnica de bacias hidrográficas (Navon et al., 2005)	21
2.7	Limiares locais (Navon et al., 2005)	22
2.8	Método de corte de grafo(Bhosale et al., 2013)	23
3.1	Classificação no conjunto de dados com dois objetos (Zhou et al., 2004)	27
4.1	Dinâmica de partículas e nós (Breve et al., 2012)	31
5.1	Grafo montado pelo algoritmo de competição entre partículas	40
5.2	Imagem inicial 50X50 pixels	41
5.3	Imagem de Satélite 64X80 pixels	42
5.4	Imagem de Satélite 75X51 pixels	43
5.5	Imagem de Satélite 100X72 pixels	44
5.6	Imagem de Satélite 100X100 pixels	45
5.7	Imagem de Satélite 160X120 pixels	46
5.8	Modelo comparativo de segmentação Blake et al. (2004)	46
5.9	Imagem comparativa Blake et al. (2004)	47
5.10	Utilizando a segmentação em mosaico, obtido utilizando a média RGB em cada região segmentada. O gráfico está representado associado a segmentação em verde claro.	47
5.11	Imagem comparativa Harchaoui and Bach (2007)	48
5.12	Imagem comparativa Irony et al. (2005)	49
5.13	Imagem de saída para imagem de entrada da Figura 5.12a) usando o modelo proposto em Irony et al. (2005)	50
5.14	Imagem comparativa Gómez-Chova et al. (2008)	51
5.15	Imagem de saída para imagem de entrada da Figura 5.14a) usando o modelo proposto em Gómez-Chova et al. (2008)	52
5.16	Imagem comparativa Gómez-Chova et al. (2008) utilizando Filtro mediana aplicado na imagem de entrada	53
5.17	Imagem comparativa Gómez-Chova et al. (2008) utilizando Filtro mediana aplicado na imagem de saída	54

Lista de Tabelas

4.1	Tabela de parâmetros	37
4.2	Algoritmo	38

Sumário

1	Introdução	9
1.1	Objetivos e Motivações	10
1.2	Organização do Trabalho	11
2	Segmentação de Imagens	12
2.1	Detectores de borda	14
2.2	Técnicas de Limiarização	16
2.3	Crescimento de regiões	18
2.4	Técnica <i>Watersheds</i>	20
2.5	Técnicas Baseadas em Grafos	21
3	Aprendizado de Máquina	24
3.1	Aprendizado Semi-Supervisionado	25
3.1.1	Métodos Generativos	26
3.1.2	Clusterizar e Rotular	27
3.1.3	Co-Training	27
3.1.4	Separação de baixa densidade	28
3.1.5	Métodos Baseados em Grafos	28
4	Modelo de Movimentação de Partículas	30
4.1	Metodologia	34
4.1.1	Alteração do Modelo de Competição e Cooperação entre partículas	35
4.1.2	Inserção de Imagens ao algoritmo	36
4.1.3	Algoritmo de Competição e Cooperação entre Partículas aplicado a imagens . . .	37
5	Resultados	39
6	Conclusão	55
	Referências	57

Capítulo 1

Introdução

A segmentação de imagens é utilizada por diversas áreas, e exerce a função de dividir uma imagem digital em múltiplos segmentos, regiões ou objetos, com o objetivo de facilitar a análise e compreensão (Shapiro and Stockman, 2001). Métodos de segmentação de imagens são frequentemente utilizados para localizar objetos e bordas em imagens. Neste processo, um rótulo é atribuído para cada pixel da imagem, de modo que pixels com o mesmo rótulo compartilhem características visuais, tais como cor, intensidade ou textura.

Existem centenas de técnicas de segmentação de imagens na literatura (Zaart et al., 2002), mas não há um único método que possa ser considerado adequado para todos os tipos de imagem. Além disso, nem todos os métodos tem o mesmo desempenho em um tipo particular de imagens (Pal and Pal, 1993), de modo que este ainda é um campo de pesquisa que oferece muitos desafios. O problema de segmentação de imagens também está fortemente relacionado com o campo de aprendizado de máquina, onde o interesse é projetar e desenvolver algoritmos e técnicas que permitam aos computadores “aprender”, isto é, melhorar seu desempenho através da experiência (Mitchell, 1997).

Dentre as principais categorias de aprendizado de máquina, podemos citar duas das mais clássicas: *aprendizado supervisionado* e *aprendizado não supervisionado*. Uma das formas de aprendizado não supervisionado é o agrupamento de dados. Métodos de agrupamento de dados vêm sendo utilizados com sucesso em muitas aplicações, incluindo as de tomada de decisão, mineração de dados, recuperação de documentos, reconhecimento de padrões e também segmentação de imagens (Jain et al., 1999; Schaeffer, 2007; Duda et al., 2000; Kaufman and Rousseeuw, 1990; Jain and Dubes, 1988). Quando o agrupamento de dados é utilizado em segmentação, os pixels da imagem de entrada são agrupados de acordo com sua similaridade, com base na hipótese de que pixels que pertencem ao mesmo objetos/segmentos são mais similares do que pixels que pertencem a objetos/segmentos diferentes.

O aprendizado supervisionado, apesar de utilizar, também, como forma de aprendizado o agrupamento de dados, este tipo de aprendizado de máquina é aplicado quando uma quantidade significativa de elementos de dados está rotulada e disponível para o processo de aprendizado. O objetivo é extrair um modelo desses dados rotulados, que, quando aplicado a novos elementos, seja capaz de rotulá-los. No caso de segmentação de imagens, é necessário rotular uma quantidade significativa de pixels ou imagens de mesmo tipo para o processo de treinamento, e então o classificador gerado deverá ser capaz de classificar os demais pixels ou segmentar imagens semelhantes.

No entanto muitas das situações apresenta apenas um pequeno subconjunto de pixels de uma imagem efetivamente e rapidamente rotulado, pois em alguns tipos de imagem, a rotulação de dados manualmente

em processos de segmentação costuma ser trabalhoso e demorado, especialmente em regiões de borda. Em contra partida, muitas imagens, devido as suas características mescladas entre uma classe e outra, apresenta uma classificação pouco precisa ou com maior complexidade de implementação, quando utilizado o aprendizado não supervisionado. Uma nova categoria de algoritmos de aprendizado de máquina, chamada de *aprendizado semi-supervisionado* vem se mostrando bastante eficaz em tais situações. Estes métodos combinam uma pequena quantidade de itens de dados rotulados com uma grande quantidade de dados não rotulados, com o objetivo de construir classificadores melhores, ao mesmo tempo em que requerem menor esforço humano (Zhu, 2005; Chapelle et al., 2006; Abney, 2008). No caso específico da segmentação de imagens, um especialista humano pode rotular manualmente uma pequena quantidade de pixels em cada segmento, e o algoritmo de aprendizado semi-supervisionado pode ser capaz de classificar os demais pixels. (Wang et al., 2006)

O modelo aqui proposto teve como objetivo integrar a segmentação de imagens ao modelo de movimentação de partículas. Tal modelo foi inicialmente desenvolvido por Quiles et al. (2008), e caracteriza-se como um modelo de detecção de comunidades baseado em competição de partículas em redes complexas. Nele várias partículas caminham pela rede, competindo umas com as outras pela posse de nós, rejeitando partículas intrusas. Desta forma, após uma quantidade de iterações, cada partícula estará confinada dentro de uma comunidade, de forma que as comunidades possam ser divididas analisando os nós que cada partícula possui. Breve et al. (2012) estenderam esta abordagem de competição de partículas, gerando um modelo capaz de realizar classificação de dados genéricos através de aprendizado semi-supervisionado. Neste novo modelo existem múltiplas partículas, organizadas em times, para representar cada classe. As partículas do mesmo time caminham pela rede de maneira cooperativa para propagar seus rótulos, enquanto partículas de times diferentes competem umas com as outras para determinar as bordas das classes. Em resultados de simulações o modelo de competição e cooperação entre partículas foi capaz de classificar dados não linearmente separáveis, executando mais rápido que outros métodos tradicionais devido à menor ordem de complexidade computacional, o que o torna capaz de classificar bases de dados maiores. Além disso, o método é capaz de obter resultados melhores que o de vários algoritmos do estado da arte quando é utilizado apenas um pequeno conjunto de nós rotulados para o treinamento.

1.1 Objetivos e Motivações

O uso da segmentação de imagens representa uma etapa primordial em diversos problemas que dedicam-se a trabalhar com imagens digitais. Ao aplicar a segmentação iterativa de imagens é possível trabalhar com aprendizado de máquina sobre imagens. No caso deste projeto, é utilizado o aprendizado semi-supervisionado. Especificamente, estendeu-se o modelo de competição e cooperação entre partículas (Breve et al., 2012) para realizar segmentação semi-supervisionada dos pixels das imagens. Tal extensão não foi trivial devido a necessidade de extração de atributos utilizados na classificação diretamente da imagem a ser segmentada.

Nesta dissertação é apresentado um levantamento de informações sobre segmentação de imagem e aprendizado de máquina, incluindo as técnicas mais utilizadas e suas variações. Com base nos conhecimentos adquiridos, o modelo de aprendizado semi-supervisionado que utiliza competição e cooperação entre partículas foi estendido para realizar segmentação de imagens.

As principais motivações deste desenvolvimento foram:

1. No aprendizado semi-supervisionado, os algoritmos são capazes de classificar grandes bases de dados onde apenas um pequeno subconjunto de dados estão rotulados. Este tipo de configuração é comumente encontrada em imagens, onde podem haver grandes quantidades de pixels, mas rotulá-los manualmente é um processo bastante trabalhoso, principalmente nas regiões de borda entre os segmentos.
2. A abordagem de movimentação de partículas, que será apresentada no Capítulo 4, é diferente da maioria dos métodos que utilizam o aprendizado semi-supervisionado baseados em grafos já propostos, pois a grande maioria deles é equivalente a um *framework* de regularização, diferindo apenas na escolha dos regularizadores e funções de perda (Zhu, 2005). Além disso, este método apresenta ordem de complexidade computacional linear, $O(n)$, enquanto que a maioria dos métodos baseados em grafo tem complexidade computacional cúbica $O(n^3)$. Portanto, o desempenho da aplicação do modelo de partículas em segmentação interativa pode ter custo computacional inferior e, conseqüentemente, ser mais rápido que outros métodos tradicionais.
3. O modelo de partículas é bastante flexível, já foi adaptado com sucesso para diversas outras tarefas de aprendizado de máquina, incluindo detecção de comunidades sobrepostas (Breve and Zhao, 2013; Breve et al., 2011), aprendizado com dados imperfeitos (Breve and Zhao, 2012b, 2011; Breve et al., 2015), classificação de fluxos de dados com mudança de conceito (Breve and Zhao, 2012a), e aprendizado ativo (Breve, 2013). Esta flexibilidade faz com que a adaptação do modelo para a segmentação interativa de imagens seja uma abordagem bastante promissora.

1.2 Organização do Trabalho

Este trabalho é organizado da seguinte forma, no Capítulo 2, é apresentada a área de segmentação de imagens e suas principais categorias de algoritmos. No Capítulo 3 são expostas as principais técnicas de aprendizado de máquina, analisando os tipos de dados e comportamento de alguns métodos, com destaque para o aprendizado semi-supervisionado, que é utilizado neste trabalho.

No Capítulo 4 são expostos o conceito e raciocínio do algoritmo que será utilizado para a manipulação dos dados que serão extraídos utilizando técnicas de segmentação de imagem, bem como classificados usando técnicas de aprendizado de máquina. No capítulo Capítulo 5 são expostos os resultados obtidos com a execução dos modelo citado no capítulo anterior. Finalmente, no Capítulo 6 estão expostas as observações principais, pontos positivos e negativos encontrados durante a execução da pesquisa, analisando o projeto, bem como os trabalhos futuros possíveis com esta aplicação.

Capítulo 2

Segmentação de Imagens

Como citado anteriormente no Capítulo 1, o conceito de segmentação de imagens se dá através do processo de identificação e separação de objetos e estruturas que sejam considerados relevantes para a imagem (Malmberg, 2011). Tais separações e identificações são utilizadas como formas de extrair informações de uma determinada imagem.

Na definição matemática para segmentação de imagens vista em (Qi, 2005), R_i representa uma segmentação da imagem R onde:

- 1 : $R = \bigcup_{i=1}^n R_i$, a união de todas as regiões resultam em todo R ;
- 2 : $R_i \cap R_j = \emptyset$, para todo i e j , onde $i \neq j$ pois não há sobreposição de regiões;
- 3 : $P(R_i) = Verdadeiro$ para $i = 1, 2, \dots, n$ e P representa o predicado de uniformidade lógica nos pontos de R_i ;
- 4 : $P(R_i \cap R_j) = Falso$ para $i \neq j$ e R_i e R_j são regiões vizinhas;
- 5 : R_i representa uma região sendo $i = 1, 2, \dots, n$;

A segmentação de uma imagem pode ser dividida em duas partes, a primeira é a parte de reconhecimento, a qual identifica a região do objeto demarcado; e a segunda parte é a de delimitação, a qual consiste na busca pela extensão exata do objeto. Quando a tarefa de segmentação corresponde a junção da capacidade humana com a interação em máquina, ou seja a busca pela região ocorre em mais de uma interação, é denominada segmentação interativa de imagens, a qual corresponde a combinação da capacidade humana, para a tarefa de reconhecimento, com a capacidade de melhor delineamento fornecida pela máquina. (Malmberg, 2011)

O método de segmentação de imagem interativa ou Semi-automática consiste na combinação de informações vindas do usuário e de informações extraídas pela máquina, utilizando as marcações feitas pelo usuário para reconhecer as classes e divisões de regiões da imagem, enquanto a máquina executa o delineamento das regiões ou classes. Essa técnica consiste na união de conhecimentos, aumentando as possibilidades de uma melhor classificação final. (Malmberg, 2011)

Atuando de maneira oposta à segmentação interativa de imagens, há a segmentação autônoma, na qual a máquina fica responsável por todo o processo de segmentação. Este tipo de segmentação é uma das mais complexas (Gonzalez and Woods, 2006). Por este motivo, neste projeto utiliza-se uma técnica de aprendizado semi-supervisionado para realizar segmentação interativa de imagens. Deste modo, um

especialista pode auxiliar no processo de segmentação, rotulando manualmente alguns pixels de cada objeto. O algoritmo então se encarregará de completar a tarefa, usando para isso tanto os pixels rotulados quanto os não rotulados como fonte de aprendizado. Tais passos serão vistos mais à frente.

Voltando as características de segmentação de imagem, são encontradas diversas formas de se realizá-la, dentre elas a transformação de imagem natural em clusters de pixels, e assim é possível estudar as diversas partes das imagens separadamente. No entanto ao se obter uma imagem, antes de submetê-la ao processo de segmentação é necessário analisá-la e verificar sua qualidade diante do meio em que a imagem foi capturada ou a forma/detalhamento em que era esperada a captura.

Uma análise de forma pontual neste tipo de imagem pode acarretar perdas de informações. Tais perdas podem ser evitadas quando é realizado um prévio processamento sobre a imagem, de forma a melhorar o contraste, o brilho e a nitidez, e assim garantir uma melhor segmentação (Saldanha and FREITAS, 2009). Para a escolha do tipo de processamento é importante levar em conta o nível de detalhamento necessário para a imagem. Analisa-se os efeitos de ruídos e necessidade de correções para tornar a imagem mais próxima o possível da imagem ideal vista sem interferências. (Gonzalez and Woods, 2010)

Após trabalhar a imagem para assim obter uma melhor segmentação, é necessário analisar qual a necessidade de segmentação e consequentemente de resposta, atentando-se para que as estruturas presentes na imagem apresentem um nível de detalhamento intermediário, ou seja, nem muito delicado, nem muito grosseiro (Pantofaru, 2008). Devido ao fato de existirem diversas formas de analisar e extrair informações de uma determinada imagem, há a dificuldade de se escolher o melhor algoritmo para realizar a segmentação. Com base nos estudos de Shi and Malik (2000), a dificuldade consiste na análise do detalhamento necessário e a divisão hierárquica de aspectos dentro da imagem. Torna-se necessária a análise de como o algoritmo exerce a separação dos grupos e os aspectos gerais relevantes ao algoritmo, em conjunto com o tempo e critério de execução por linha de pixels, com a finalidade de obter um método prático e eficiente.(Felzenszwalb and Huttenloche, 2004)

"...não deve ter como objetivo produzir um final 'correto' da completa segmentação. O objetivo deve ser usar a coerência de baixo nível, de brilho, cor, textura, movimento ou atributos para vir sequencialmente com partições hierárquicas. ... o conhecimentos de alto nível pode ser usado tanto para confirmar estes grupos ou selecionar alguns para mais atenção. Esta atenção pode resultar em mais reparticionamento ou agrupamento. ... como se fosse um pintor de primeira marcando as principais áreas e, em seguida, preencher os detalhes."(Shi and Malik, 2000)

Estas características possibilitam uma garantia na resposta obtida através da segmentação, pois todas as imagens serão analisadas sob o mesmo padrão de parâmetros. No entanto é necessário que ocorram testes com diversos tipos de algoritmos de segmentação, para que assim seja possível analisar os resultados obtidos entre eles, e seguir com o algoritmo de melhor resultado. Essa escolha envolve a análise dos algoritmos que melhor trazem a realidade da imagem original, se há um padrão na análise das imagens. (Pantofaru, 2008)

Dentre os tipos de imagens que podemos encontrar, as coloridas são as mais exigentes quando analisadas, em consequência da quantidade de informações que podem ser extraídas e detalhes presentes. Segundo Skarbek and Koschan (1994), para este tipo de imagem é utilizado um processo de extração de

regiões que possuam homogeneidade. Esta obtenção pode ser auxiliada com propriedades geométricas, em união com análises ópticas de especialistas, ou ainda analisando a diferença na intensidade dos tons. Para tal extração é necessário observar a imagem para definir a quantidade de regiões distintas que podem ser encontradas e escolher a melhor opção para análise das cores.

Existem algumas formas de representar os pixels da imagens frente a segmentação, dentre elas há a representação em RGB (Vermelho, Verde e Azul), a qual ocorre através de três vetores, representando cada uma das componentes de cores, o modelo de cores CMY (Ciano, Magenta e Amarelo) o qual também identifica as cores através de três vetores, porém utilizando componentes de cores diferentes e ainda o modelo HSI (Matiz, saturação e intensidade), onde a matiz representa a cor dominante; a saturação representa a pureza da cor, ou o quanto de luz branca está presente; e a intensidade, a qual representa a luminosidade.

Diferente dos modelos de cores RGB e CMY, o modelo HSI não é representado utilizando um vetor de cores. Segundo Gonzalez and Woods (2010) o sistema de cores RGB não é totalmente adequado para se descrever uma cor levando em consideração a percepção humana. No exemplo citado, Gonzalez and Woods (2010) fazem a menção das cores de carros para exemplificar seu pensamento, na qual normalmente não é referida a porcentagem de vermelho, verde ou azul. Dizemos apenas a cor equivalente a combinação correspondente a estes tons.

Nas próximas subseções serão descritas técnicas de segmentação conhecidas da literatura que permitem trabalhar com imagens coloridas.

2.1 Detectores de borda

Como vimos no início do capítulo, existem diversas formas de realizar divisões dentro da imagem. Seguindo este preceito há a possibilidade de separação de regiões ou objetos de algumas imagens utilizando as bordas. Em algumas figuras, é comum encontrar uma borda bem definida entre os objetos da imagem. Esta borda pode ser definida por uma linha inserida entre os dois objetos ou diferença abrupta entre os pixels vizinhos, porém de objetos ou regiões distintas. As bordas também podem ser identificadas pela destaque em torno dos limites da figura. (Gonzalez and Woods, 2010)

No decorrer da interação a região delimitada pelas bordas crescem, fortalecendo cada vez mais as bordas na tentativa de detectar o final do objeto ou região em questão, separando por completo as regiões internas da figura, detectando a parte de margem do objeto e a direção da parte central. (Navon et al., 2005)

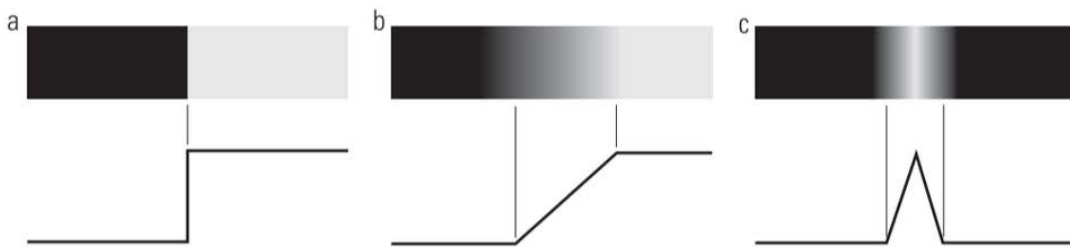
A aplicação destas técnicas em imagens que possuem apenas tons de cinza, possibilita uma classificação de acordo com as bordas formadas pelo contraste entre tons de cinza, utilizados em mudanças de tons dentro de um objeto e/ou borda entre os objetos (Dass et al., 2012). Os métodos de detecção de borda necessitam de um equilíbrio entre a precisão e a imunidade ao ruído. Quando existe um grau alto de precisão, o ruído de classificação pode interferir nos contornos na imagem. No entanto a filtragem de ruídos pode causar perda de detalhes na imagem. Deste ponto de vista, o ideal para este tipo de algoritmo é a escolha de imagens simples com baixo ruído, ou seja bordas precisas entre um objeto e outro.

Existem alguns modelos de borda representados em Gonzalez and Woods (2010). Tais modelos, representados na Figura 2.1, são: *bordas em degrau*, ou denominadas ideais, exemplificadas na Figura 2.1a), idealmente esta borda ocorre utilizando a transição de um pixel entre as cores; *bordas em rampa*

exemplificada na Figura 2.1b), diferentemente da borda em degrau, esta borda ocorre utilizando mais de um pixel de transição, tendo a inclinação da rampa como medida de nitidez da imagem; ou ainda *bordas em forma de telhado* exemplificada na Figura 2.1c) este modelo de borda utiliza uma rampa crescente seguida e outra decrescente, exercendo a função de uma linha divisória, onde a nitidez é medida segundo a espessura de tal linha.

Na pratica é extremamente difícil ter uma imagem com uma borda em degrau, sendo as bordas determinadas pelo mecanismo de focalização. Normalmente encontramos bordas em forma de telhado em digitalizações de desenhos, imagens de satélite quando mostram características finas como escadas.

Figura 2.1: Estilos de bordas: a) Borda em degrau; b) Borda em Rampa; c) Borda em forma de telhado



fonte: (Gonzalez and Woods, 2010)

Uma forma simplificada para se realizar a delimitação dos objetos é utilizando marcações nas bordas, por exemplo, desenhando uma linha branca em torno do objeto, de forma a identificar estas bordas. Outra forma de se aplicar a delimitação de objetos é utilizando *Laplaciano da Gaussian* (LoB), o qual possui características importantes para detecção de bordas, tal filtro permite a possibilidade de detecção de bordas, em união como a correções de foco e nitidez presentes na imagem. (Gonzalez and Woods, 2010)

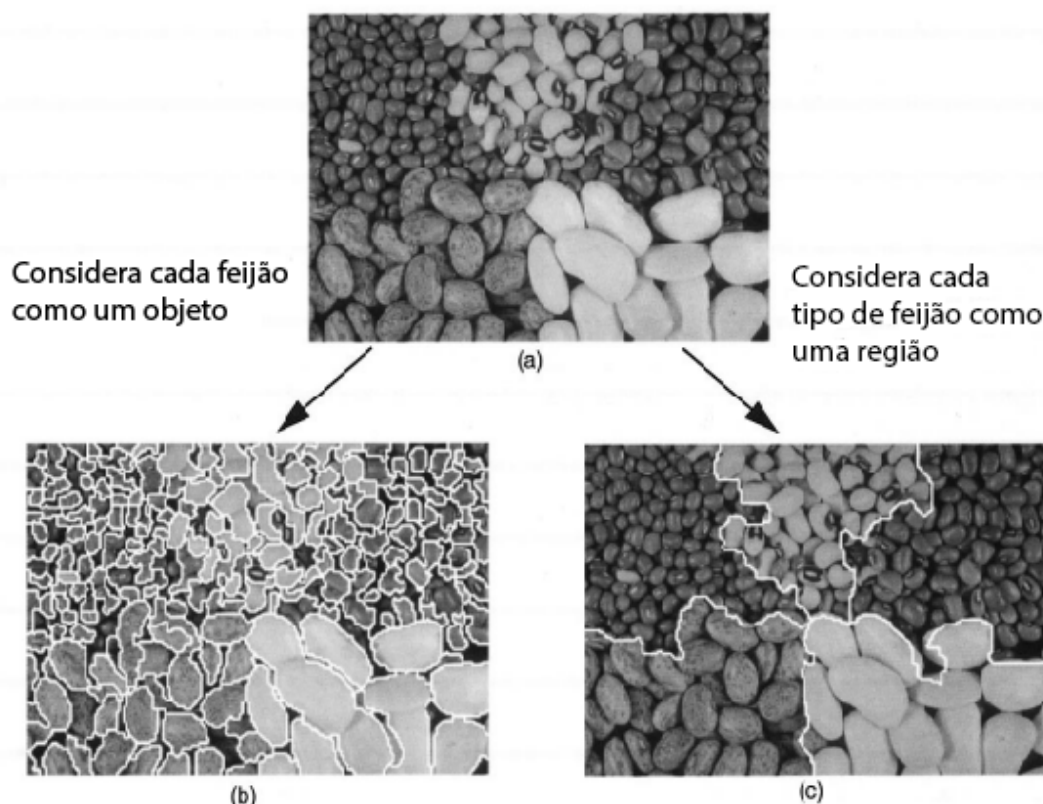
Mostrado na Equação 2.1, onde o σ corresponde ao desvio padrão, e (x, y) a posição analisada; este filtro tem como objetivo mapear toda a imagem utilizando um plano cartesiano. Porém, cada objeto presente na figura utiliza uma malha com um ponto central. Assim cada objeto possui um ponto 'zero', independentemente do número de objetos que seja identificado na figura, afim de especificar o ponto de borda individual dos objetos, identificando assim a melhor região em que se deve inserir o vetor de separação de regiões. (Glasbey and Horgan, 1994).

$$Log(x, y) = -\frac{1}{\pi\sigma^4} \left[1 - \frac{x^2 + y^2}{2\sigma^2} \right] e^{-\frac{x^2 + y^2}{2\sigma^2}} \quad (2.1)$$

Na Figura 2.2 é possível analisar melhor o funcionamento do algoritmo. Primeiramente ocorre a detecção das bordas de cada feijão, e posteriormente é utilizado uma escala maior detectando as bordas entre os tipos de feijão.

Ao realizar a detecção de bordas em uma imagem colorida, é utilizada a detecção através de cada componente da imagem seguindo o método descrito para imagens em preto e branco, de tal forma que cada componente da imagem seja visto como uma imagem única e ao ser segmentada utilizando a técnica de determinação de bordas, estas imagens são unidas novamente, combinando os resultados de forma a achar a detecção de borda correspondente para a imagem colorida. (Gonzalez and Woods, 2010) As Figuras 2.3a), b) e c) representam uma classificação, onde os 3 componentes estão nas mesmas posições x e y , e neste caso ao se somar os valor da presença das três cores em uma única parte, resulta em preto e na parte de ausência de cor, corresponde a uma parte branca, como na Figura 2.3d). Considerando uma

Figura 2.2: Detecção de bordas: (a) Apresenta a imagem com cinco tipos de feijões. (b) Resultado da segmentação utilizando uma escala pequena. (c) Resultado de segmentação utilizando uma escala maior.



Fonte: (Ma and Manjunath, 2000)

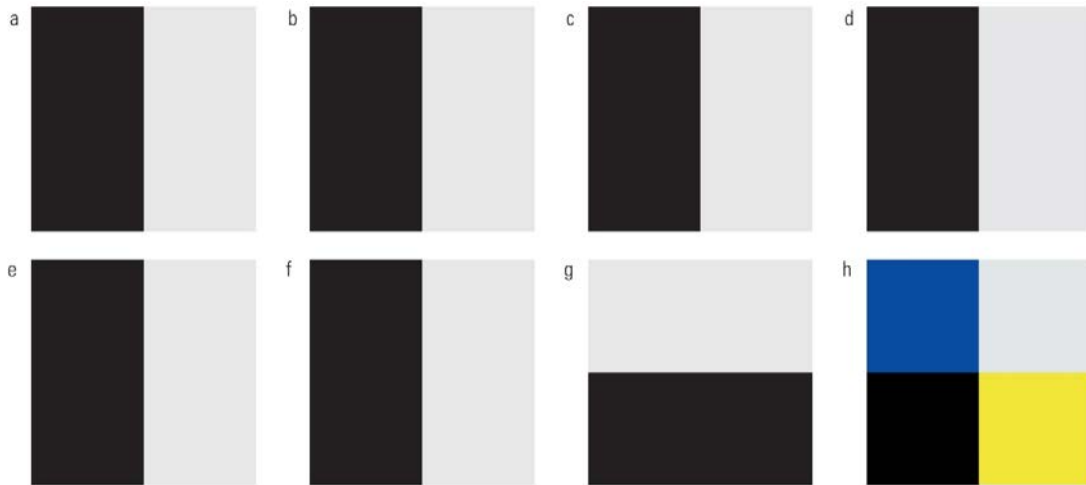
outra situação de posição entre casa divisão, temos as Figuras 2.3e),f) g) as quais não estão totalmente sobrepostas e por isso ao serem unidas na Figura 2.3h) tem-se como retorno cores diferentes, desta forma há uma parte em que há sobreposição das três cores, ocorre a exibição do preto, onde todas se sobrepõem, uma parte em que nenhuma se sobrepõe, resultado em branco, e diferentemente do primeiro exemplo, a parte em que g e f se sobrepõem representado por azul, e a parte em que apenas g é expressa, representando a parte amarela.

2.2 Técnicas de Limiarização

Uma segunda opção de separação de regiões ou objetos de uma determinada imagem é verificar o limiar da imagem. Esta técnica é utilizada principalmente em imagens escuras com fundo claro ou vice-versa. Neste caso, utiliza-se a representação de um histograma com mais de um pico, separado através de vales, para identificar os segmentos.

As técnicas de limiarização utilizam a posição de classificação no histograma e um valor T como limiar de comparação. Quando T representa um valor constante durante toda a limiarização da imagem é dito que há uma limiarização global. No entanto quando o valor de T oscila no decorrer da imagem é dito que há uma limiarização variável. Atribuído um valor e um comportamento de T , ao finalizar a classificação, os pixels são separados em duas classes. Os que se encontram abaixo do limiar T e os que se encontram acima, para encontrar o valor do limiar T é utilizado como base a Equação 2.4

Figura 2.3: Detecção de bordas em imagens coloridas: a) b) e c) Representam componentes R, G e B e d) o resultado das cores RGB, e), f) e g) Representam componentes R,G e B e h) Imagem colorida resultante das cores RGB



Fonte: (Gonzalez and Woods, 2010)

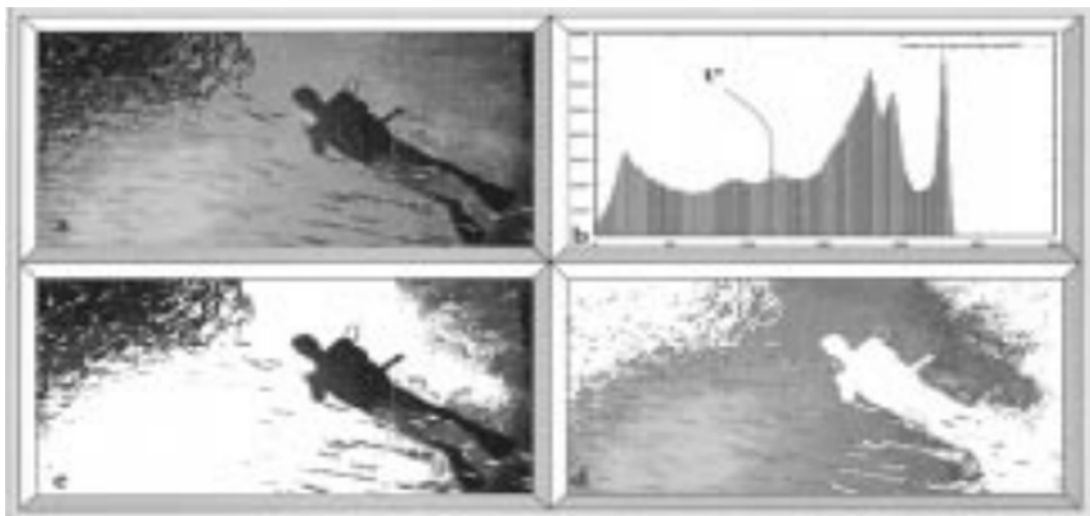
(Glasbey and Horgan, 1994). A Figura 2.4 apresenta os passos da técnica de limiarização.

$$g(x, y) = 1 \text{ se } f(x, y) > T \quad (2.2)$$

$$g(x, y) = 0 \text{ se } f(x, y) \leq T \quad (2.3)$$

$$(2.4)$$

Figura 2.4: Técnica de Limiarização: A primeira parte da imagem é a imagem original a ser segmentada, a segunda parte superior representa o histograma gerado. A ultima parte é o resultado final seguido do fundo limiarizado.



Fonte: (Glasbey and Horgan, 1994)

As técnicas de limiarização podem classificar uma série de pixels em uma única categoria por terem valores semelhantes, no entanto nem sempre eles representaram a mesma forma dentro da figura.

Devido a isto mais de uma técnica de limiarização pode ser utilizada simultaneamente e então serão produzidas mais de duas categorias, com o objetivo de evitar este tipo de falha. A junção de técnicas de limiarização apresenta uma melhora quando há pouca sobreposição da distribuição de valores dentre as diferentes categorias. Segundo Dass et al. (2012) as técnicas de limiarização, apesar de serem simples, representam uma poderosa abordagem em segmentação de imagens claras em fundos escuros.

Para o uso de imagens coloridas pode-se utilizar algumas definições vistas em Skarbek and Koschan (1994), onde regiões com poucas características tendem a pertencer ao fundo da imagem. Torna-se necessário normalizar os valores RGB de cada pixel e assim extrair regiões conectadas, objetos da imagem e plano de fundo da imagem.

No entanto, a sobreposição de categorias pode causar ruídos, ou seja produzir pixels com valores intermediários nas bordas, havendo a possibilidade de utilizar filtros para afinar os ruídos, quando estes são mais grosseiros.

2.3 Crescimento de regiões

Seguindo o objetivo da segmentação de imagens, há outra forma de se separar regiões ou objetos dentro de uma imagem, sem que seja necessário escolher suas bordas com base na diferença entre pixels próximos ou limiar. A técnica apresentada nesta subseção consiste no agrupamento dos pixels, utilizando para isto critérios pré-estabelecidos para o crescimento.

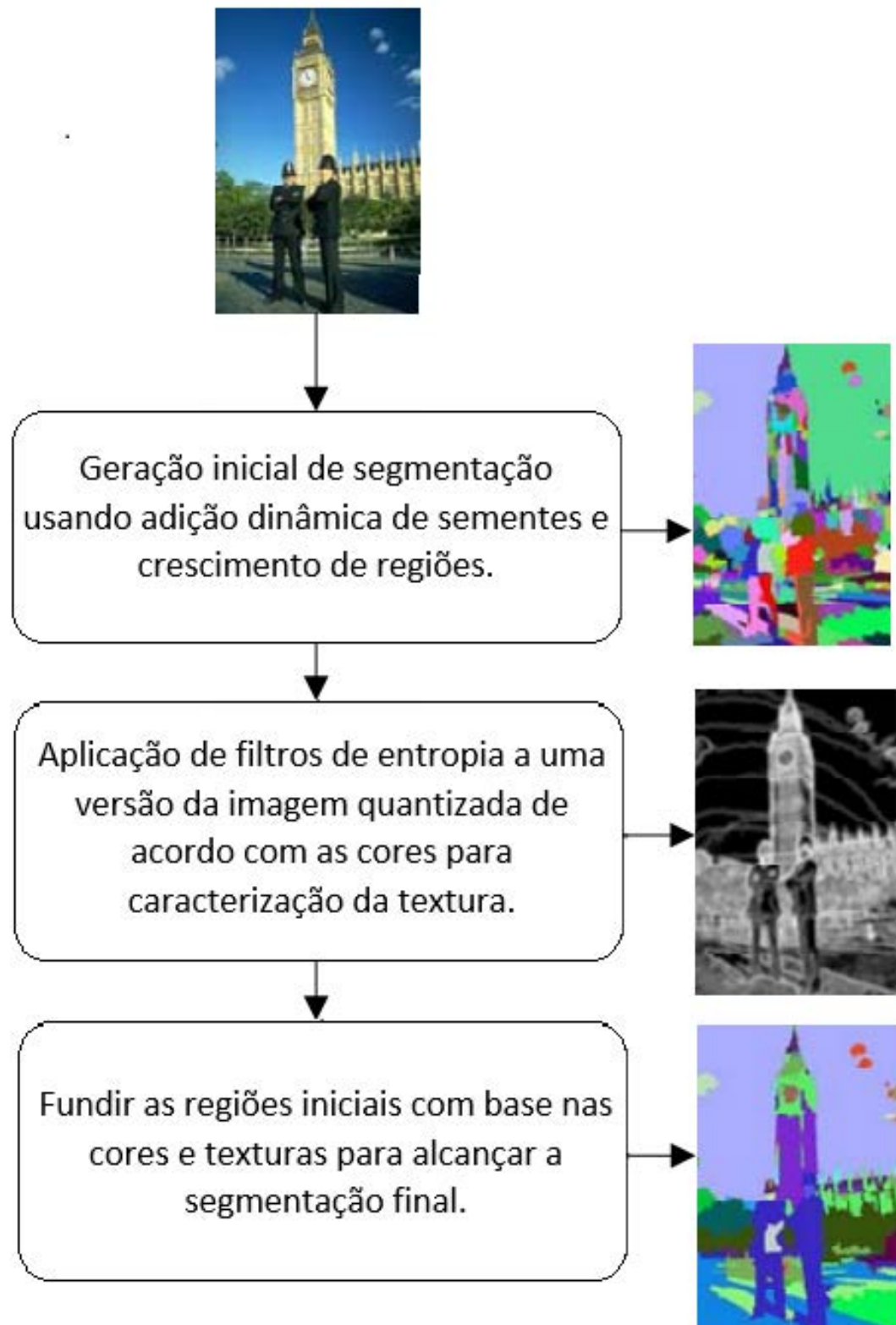
O método de crescimento de regiões consiste na determinação de pontos isolados dentro da figura em questão, seguida da procura de pixels vizinhos com atributos similares. Este crescimento é realizado de forma interativa até que todos os pixels sejam processados e diferentes formas de fronteiras sejam definidas. As similaridades são baseadas em critérios pré-definidos.

Este método, segundo Dass et al. (2012), é relativamente mais simples que o conjunto de algoritmos que utilizam detecção de bordas. Também é mais imune a ruído, pois consiste na separação da imagem em regiões semelhantes de acordo com critérios previamente definidos.

Para a implementação dos métodos de crescimento de regiões é necessário selecionar um grupo de pixels da imagem original, sementes, e selecionar um conjunto de critérios para estabelecer a afinidade entre regiões. Feitas estas duas etapas, o crescimento ocorre ao comparar cada pixel com os pixels sementes utilizando os critérios pré-estabelecidos. Outra característica, vista em Dass et al. (2012), é a divisão da imagem em um conjunto de regiões desconexas, moldadas arbitrariamente, com o intuito de se obter uma segmentação de imagem razoável. Este método aplica o conceito de divisão e junção entre as partes de acordo com as propriedades encontradas no conjunto, sendo executado até o momento que se possa obter diversas regiões, e dentro de cada uma encontrar partes que contenham propriedades idênticas. Estes passos podem ser melhores assimilados na Figura 2.5

No entanto, um dos problemas de tal técnica é a condição de parada do crescimento. Para melhorar a técnica é possível acrescentar o conceito de tamanho da região, semelhança entre os pixels e os pixels selecionados até o momento, além dos demais itens como por exemplo intensidade, textura e cor. (Gonzalez and Woods, 2010) A desvantagem do método é sua necessidade de pós-processamento, com a intenção de unir regiões próximas que possuam características semelhantes.

Figura 2.5: Crescimento de regiões: Primeiramente é apresentada a imagem original, após a primeira interação é gerado a segmentação utilizando sementes e crescimento de regiões. Na segunda interação é aplicado o filtro de cores para depois realizar a junção de regiões semelhantes.



Fonte:Imagem traduzida de (Garcia-Ugarriza et al., 2008)

2.4 Técnica *Watersheds*

Desde o início deste capítulo vimos algumas técnicas muito específicas para determinados tipos de imagens ou com algumas restrições de processamento. Com o intuito de mesclar as vantagens apresentadas por tais técnicas de segmentação foi criada uma nova técnica, denominada *watersheds* ou bacias hidrográficas. Esta técnica permite a incorporação de uma segmentação estável e com fronteiras de segmentação conectadas. (Gonzalez and Woods, 2010)

O objetivo da técnica de bacias hidrográficas é preencher os espaços de cada região até que sejam encontradas as bordas entre as regiões impedindo a junção das mesmas, tornando a imagem visível de acordo com as bordas inseridas. Existem três tipos de pontos considerados na técnica de bacias hidrográficas: os pontos pertencentes ao mínimo local, os pontos em que a gota de água quando inserida cairia nestes pontos mínimos, denominados de bacia hidrográfica, e os pontos em que a água quando inserida tem a probabilidade de ficar em mais de um ponto mínimo, denominado de linha de divisão. (Gonzalez and Woods, 2010)

O funcionamento da técnica pode ser exemplificado imaginando que em cada mínimo local da imagem exista um furo, e que em cada ponto máximo, ou divisão entre as bacias, haja uma barreira, de tamanho igual a todas as barreiras construídas. Ao imergir esta estrutura em água, esta começará a entrar e encher até chegar no nível da barreira. (Salman and Chong-Qing, 2001)

A imagem de entrada, para o caso das imagens coloridas, deve passar por uma conversão de tons e ser reescrita em tons de cinza, para auxiliar na identificação de contorno de objetos, deixando-os mais nítidos. Utiliza-se também, para tal etapa de análise, a camada da matiz da imagem. Assim torna-se possível aplicar o gradiente e detectar as bacias e barragens da imagem. (Navon et al., 2005)

O uso desta técnica consiste primeiramente em adicionar à imagem marcações as quais definem objetos em plano de fundo, considerando a posição da coordenada (x, y) e intensidade, garantindo que cada região marcada não pertença a mais de uma região ou objeto. Posteriormente é inserido um gradiente para que selecione apenas os contornos relevantes da imagem.

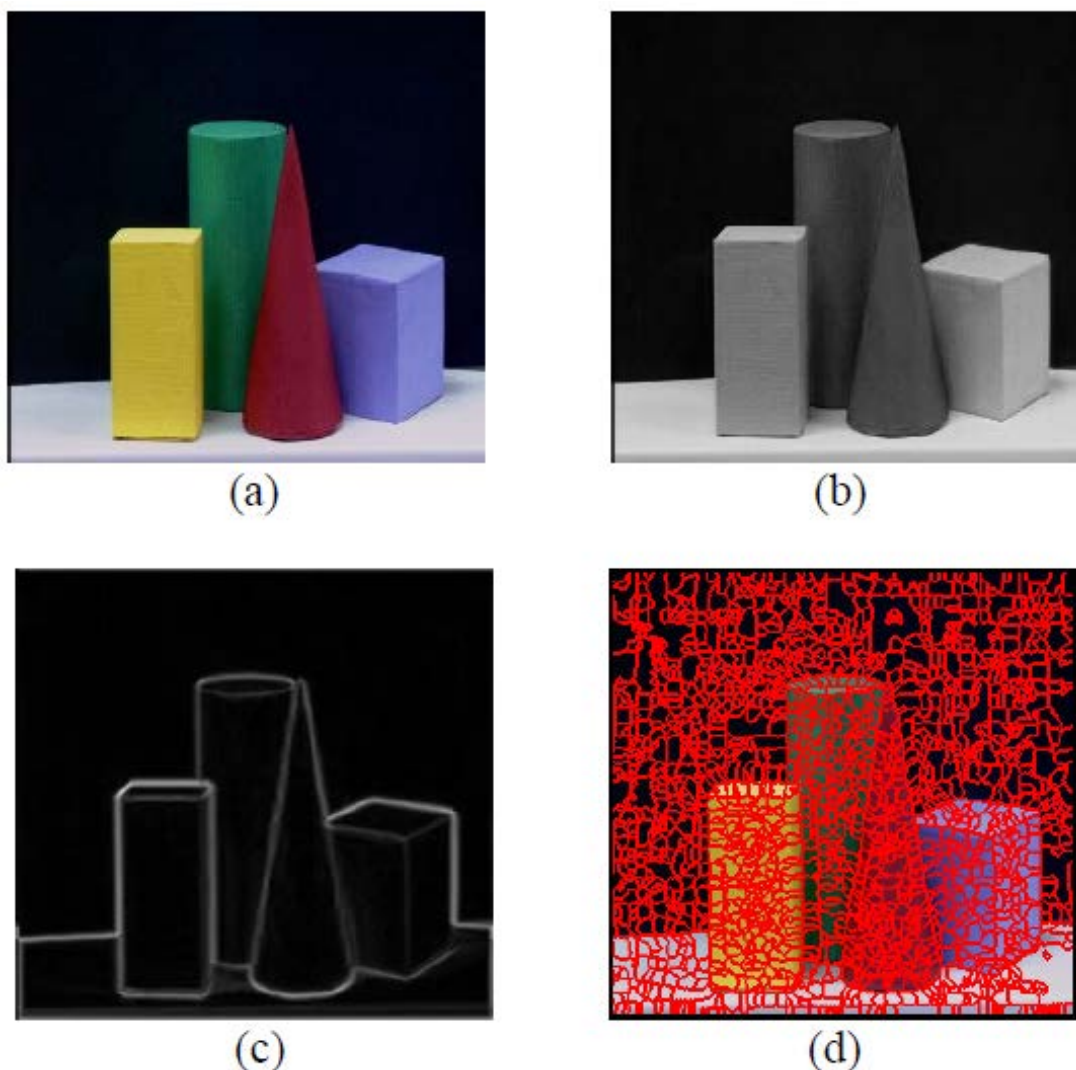
O algoritmo de bacias hidrográficas consiste em duas etapas, a de triagem e a de inundações. A etapa de triagem consiste na classificação dos pixels em ordem crescente de acordo com a intensidade de cada um. A etapa de inundações consiste na verificação dos pixels pela ordem de classificação para a construção de bacias, onde cada bacia hidrográfica recebe um rótulo, construindo ao final do processo um mosaico da imagem representada em bacias.

Após a aplicação destas etapas do algoritmo de bacias hidrográficas é verificado a similaridade entre as regiões vizinhas. Para isso é tomado como referência as cores e bordas de cada região analisada, influenciados pela iluminação e sombra. O subgrafo é gerado com as mínimas regiões semelhantes encontradas, podendo atribuir pesos aos nós que interligam esta região. Caso encontre nós interligados com pesos menores, estes nós são adicionados ao mesmo ponto do grafo, unido as regiões. Para um melhor estendimento, o funcionamento deste grafo é descrito na subseção seguinte.

Muitas figuras apresentam partes onde a utilização de um limite global seria suficientemente válido. Tendo em vista a separação detalhada do algoritmo de bacias hidrográficas, podem existir partes do mesmo objeto a volta de um determinado ponto, porém com uma pequena alteração de luminosidade ou intensidade da cor. Contudo continuam representando o mesmo objeto. Contudo há imagens que não se enquadram nestas características.

Para o meio termo destas soluções são utilizadas os limiares locais, identificando partes homogêneas

Figura 2.6: Aplicação técnica de bacias hidrográficas: (a) Imagem original colorida. (b) Imagem em escala de cinza. (c) Imagem após a aplicação do detector de bordas. (d) Segmentação resultante do algoritmo de bacias hidrográficas



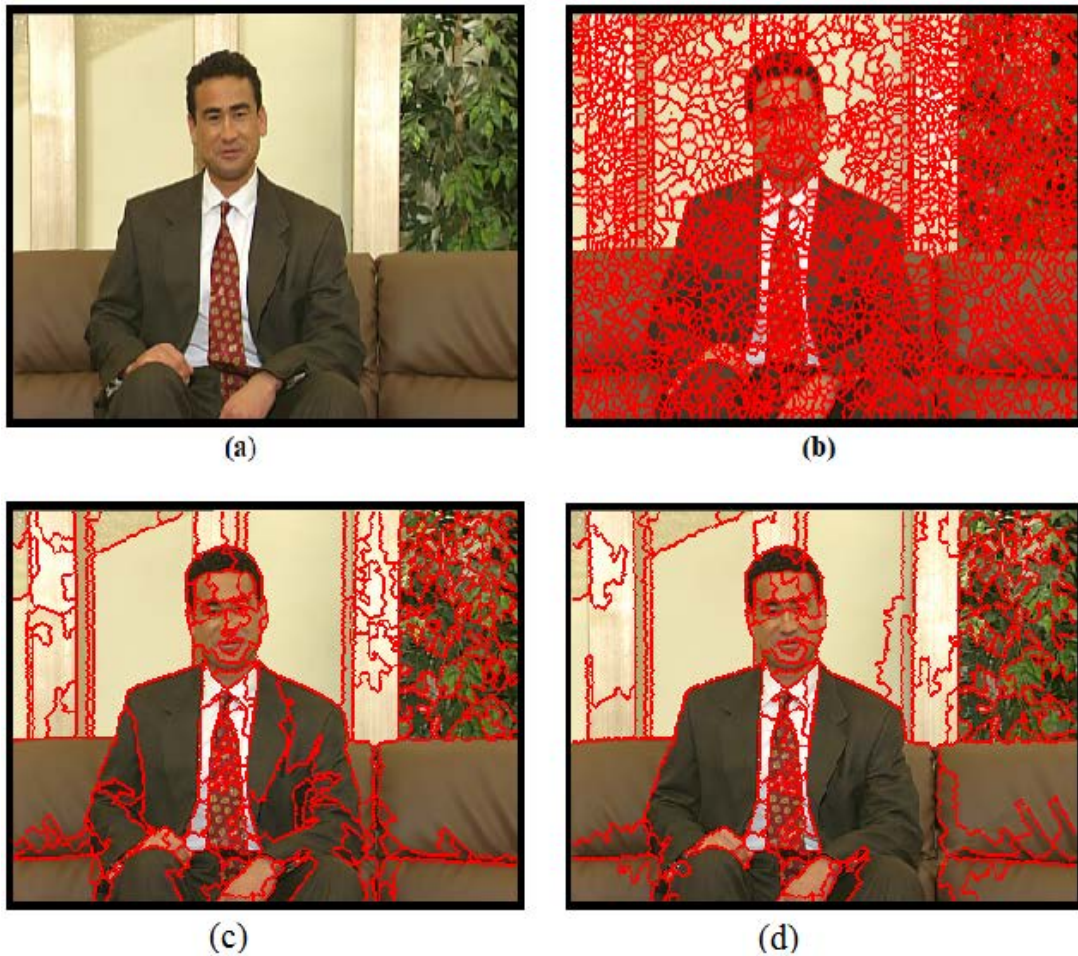
Fonte: (Navon et al., 2005)

na imagem e unindo as pequenas partes, com a intenção de diminuir o numero de regiões que dividem a imagem como ilustrado nas Figuras 2.6 e 2.7.

2.5 Técnicas Baseadas em Grafos

Mencionamos na subseção acima o uso do modelo de grafo como sendo um próximo passo da técnica de bacias hidrográficas, no entanto esta técnica pode ser utilizada em outras situações. As técnicas de construção de grafos possuem como objetivo a representação dos pixels através de nós, onde a presença, ou falta, de vizinhança determina a afinidade entre eles, formando arcos a fim de representar pontos com maiores afinidades. Na montagem do grafo, fica a critério da necessidade da aplicação a atribuição de pesos para a ligação entre os nós, podendo ser estabelecido um peso padrão, neutro, ou um peso referente à afinidade dos nós das pontas da aresta. (Navon et al., 2005)

Figura 2.7: Limiares locais: (a) Imagem original colorida. (b) Segmentação resultante do algoritmo de bacias hidrográficas. (c) e (d) Após a união de regiões.



Fonte: (Navon et al., 2005)

A imagem é transformada em uma matriz de valores e posteriormente em um grafo, o qual é montado seguindo a proximidade entre características de cada pixel, e cada nó corresponde a um pixel da imagem, formando um grafo $G = (V, E)$ não direcionado com vértices $v_i \in V$, os quais possuem uma lista de vizinhos $(v_i, v_j) \in E$ com um peso não negativo de $w((v_i, v_j))$ utilizando a distância entre nós para classificar à similaridade entre v_i e v_j .

O objetivo da segmentação é definir fronteiras e distinção da intensidade entre os pixels, dentro de uma determinada região. Algumas características podem ser extraídas com o conhecimento da distância entre nós e características comuns em nós próximos, sendo possível identificar dentro da imagem diferença da intensidade do brilho, cor, entre outros. É necessário estabelecer critérios os quais tornem possível distanciar pixels que representem diferentes objetos dentro da imagem, e simultaneamente aproximar pixels que representem o mesmo objeto ou semelhante, mesmo que este objeto seja representado em áreas separadas por demais objetos.

Ao trabalhar com imagens coloridas é necessário incorporar alterações ao método, devido a semelhança entre cores sobrepostas. A solução para tal problema é baseado na montagem de um grafo 3D, o qual representa a construção de um grafo para cada vetor de cor. A técnica consiste na divisão das cores de cada ponto em R, G e B e a divisão da imagem em conjuntos para cada uma destas camadas de cores.

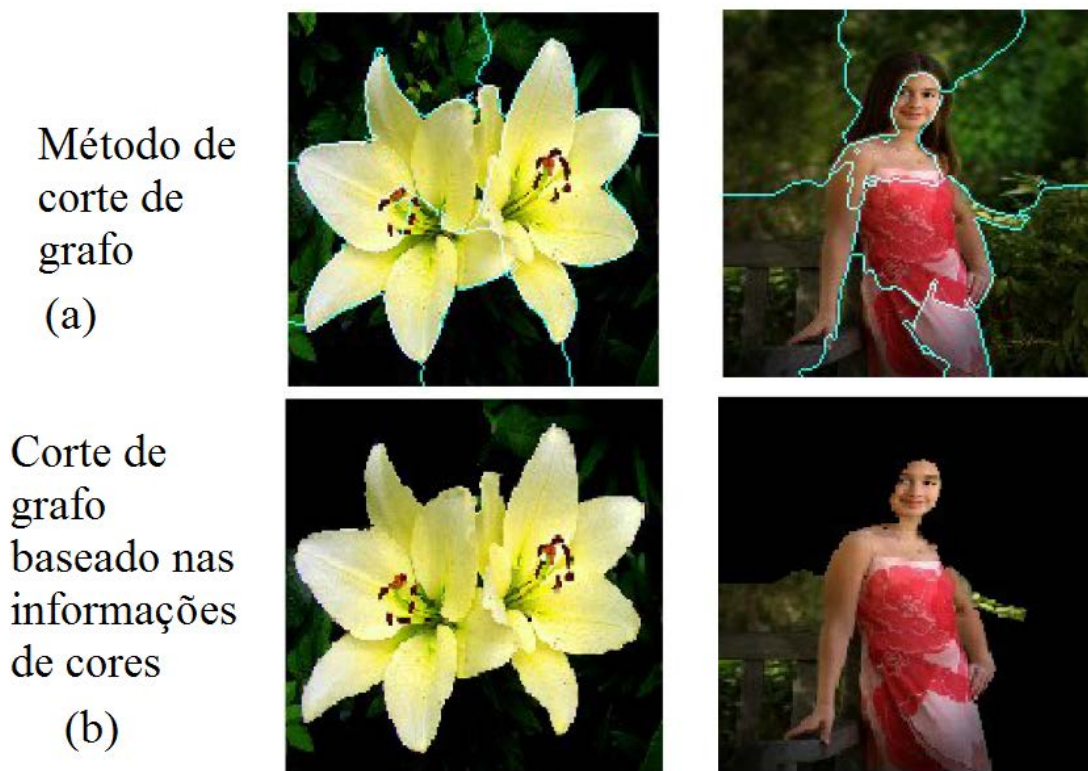
Também é possível utilizar o modelo HSI de cores, seguindo o mesmo critério.

Uma forma de construção do grafo em 3D exibida por Quan et al. (2006), faz o uso de uma estrutura semelhante ao método descrito a cima, porém no grafo 3D é utilizado $K = 3$ vizinhos apenas por nó, onde o nó mais próximo é calculado pela distância Euclidiana. No entanto, o peso sobre as arestas é calculado de acordo com a possibilidade dos nós pertencerem ao mesmo objeto na imagem.

Para a aplicação do método baseado em grafo em imagens coloridas, é necessário realizar sobre a imagem uma pré-segmentação. Aplicando, segundo Navon et al. (2005), o algoritmo baseado em bacias hidrográficas, apresentado na seção anterior, é possível gerar a chamada super-segmentação. Em seguida aplica-se os processos de dissimilaridade entre regiões, gerando um grafo com os mínimos e cálculo de limiares locais.

Ao realizar as separações entre objetos, são encontradas arestas que ligam objetos ou classes, diferentes. Sendo assim, essas arestas se tornam desnecessárias no grafo, visto que o principal objetivo é isolar os objetos do grafo. Ao se retirar estes vértices é possível que ocorra uma quebra dos grafos, sendo impedida a junção novamente ou comunicação entre estes nós. Uma das adversidades é escolher os melhores vértices a serem retirados para que a separação do grafo seja realizada em apenas algumas partes, onde é aconselhável que se retire o mínimo de vértices respeitando a melhor identificação entre partes. A Figura 2.8 exemplifica esta separação do grafo em corte, onde na Figura 2.8 (a) são consideradas as regiões da imagem, e na Figura 2.8 (b) são executados os cortes com base nas regiões separadas em (a), cortando partes que não fazem parte do objeto principal da imagem.

Figura 2.8: Método de corte de grafo: a) Método de corte de grafo. b) Corte no grafo baseado nas informações das cores.



Fonte: (Bhosale et al., 2013)

Capítulo 3

Aprendizado de Máquina

Este capítulo tem como objetivo um estudo e análise em torno do conceito do aprendizado de máquina. Expondo as principais técnicas de aprendizado de máquina, analisando os tipos de dados e comportamento de alguns métodos, com destaque para o aprendizado semi-supervisionado, devido a sua utilização neste projeto.

Esta abordagem conceitual visa entender e analisar algoritmos que melhoram automaticamente com a experiência, imitando o comportamento de aprendizado de humanos. Um dos principais focos da pesquisa em aprendizado de máquinas é aprender automaticamente a reconhecer padrões complexos e tomar decisões inteligentes com base em dados, no entanto o impasse encontra-se na análise e compreensão do conjunto de todos os possíveis comportamentos, considerando as possíveis entradas, tornando-se complexo para ser descrito em forma de linguagens de programação. Muitas aplicações de aprendizado de máquina já foram desenvolvidas, variando de programas de mineração de dados que detectam transações fraudulentas em cartões de crédito a sistemas de filtragem de informações que aprendem os hábitos de leitura do usuário. Dentre os diversos campos relacionados podemos citar: inteligência artificial, teoria de probabilidade, mineração de dados, reconhecimento de padrões, controle adaptativo e teoria da computação (Mitchell, 1997; Alpaydin, 2004; Natarajan, 2014).

Alguns algoritmos de aprendizado de máquina dedicam-se a eliminar a necessidade de especialistas humanos na análise de dados. Outros adotam uma estratégia de colaboração entre humanos e máquinas. Considerando que a percepção humana não pode ser inteiramente desconsiderada, devido ao fato de que o projetista do sistema precisa especificar como os dados serão representados e quais mecanismos serão usados para caracterizar os dados.

O aprendizado de máquina é utilizado similarmente em casos onde o conhecimento humano não existe ou onde especialistas não conseguem transmitir o raciocínio para a máquina. A título de exemplo, há tarefas como o reconhecimento de objetos, recorrentes no dia-a-dia, que são complexas quando há a tentativa de explicar e detalhar como tal reconhecimento foi realizado no cérebro. O Objetivo do algoritmo de aprendizado de máquina é analisar uma grande coleção de amostras diferentes e aprender a mapeá-las, alguns problemas de aprendizado também mudam ao longo do tempo, ou dependem de um ambiente em particular. O objetivo é construir sistemas que possam se adaptar a circunstâncias específicas, sem que seja preciso escrever um novo programa para cada nova situação (Alpaydin, 2004).

Em virtude do aumento do uso de dados, decorrente do crescimento tecnológico, é necessário alterar e melhorar a manipulação destes dados. Para assimilar tal transformação são analisados certos padrões nos dados que podem ser extraídos, e este é o papel do aprendizado de máquina.

Dentre as principais categorias de aprendizado de máquina, podemos destacar:

Aprendizado Supervisionado : Os algoritmos desta categoria deduzem uma função a partir dos dados de treinamento. Tais dados consistem em pares de exemplos de entradas e saídas desejadas, apresentando um valor contínuo, ou um rótulo de classe para o objeto de entrada. O intuito deste método é obter uma função que seja capaz de prever a saída para qualquer entrada válida, após ter visto um número suficiente de exemplos de treinamento, apresentando uma previsão precisa. Para atingir este objetivo, o algoritmo deve ter capacidade de generalização para assim prever, de maneira aceitável, a saída para dados ainda não vistos.

Para o treinamento de modelos de aprendizado supervisionado há a necessidade de um supervisor, o qual tem a função de analisar e classificar previamente dados para a máquina, e assim a mesma consegue traçar um padrão que represente todos os dados fixados, formulando uma base de parâmetros com o objetivo de, ao alterar os dados de entrada, seja possível classificar os dados desconhecidos. No entanto este tipo de classificação acaba por exigir uma ampla base de dados previamente classificados. (Barber, 2012)

Aprendizado Não Supervisionado : Os algoritmos desta categoria buscam determinar como os dados estão organizados. Tais dados de treinamento consistem apenas de exemplos de entrada, sem valores de saída. O objetivo deste aprendizado é encontrar padrões no espaço de entradas, uma das formas de atingir este objetivo é dividir os dados em grupo que apresentem características mais próximas. Tal aprendizado exclui o uso de um supervisor. Para se conseguir a identificação dos dados, ocorre a procura de semelhanças e padrões de separação entre os dados. Este tipo de aprendizado, contudo, pode apresentar falhas devido ao fato de que algumas características podem ser diferentes entre dados de uma mesma classe, e serem semelhantes entre dados de outras classes, podendo ocorrer um aprendizado pouco correto dentro do conjunto de dados. (Barber, 2012)

Aprendizado Semi-Supervisionado : Os algoritmos desta categoria mesclam o uso tanto de dados rotulados quanto de dados não rotulados para o treinamento, trazendo benefícios de ambos o aprendizado supervisionado e o aprendizado não supervisionado. Frequentemente são utilizados poucos dados rotulados e bastante dados não rotulados. Em muitos casos, o uso de alguns dados rotulados em meio aos dados não rotulados melhora consideravelmente a precisão do aprendizado. No entanto, o bom funcionamento do aprendizado semi-supervisionado depende das características dos dados a serem classificados serem compatíveis com a(s) hipótese(s) adotada(s) pelo algoritmo utilizado.

Neste projeto de pesquisa, serão utilizados algoritmos de aprendizado semi-supervisionado, portanto esta categoria será descrita com maiores detalhes na próxima subseção.

3.1 Aprendizado Semi-Supervisionado

O Aprendizado Semi-Supervisionado é uma categoria de aprendizado de máquina que mantém-se entre as duas maiores categorias: *aprendizado supervisionado* e *aprendizado não supervisionado*. Como visto no início do capítulo, o aprendizado supervisionado, consiste em um algoritmo que permite o aprendizado da função de classificação a partir dos dados de treinamento conhecidos, traçando uma linha

de classificação mais precisa para identificar os dados futuros. Em contrapartida há o aprendizado não supervisionado, no qual todos os itens de dados não são rotulados e o objetivo é determinar a estrutura entre os dados para classificá-lo. (Mitchell, 1997; Bishop, 2006; Alpaydin, 2004; Witten and Frank, 2005; Hinton and Sejnowski, 1999; Cios et al., 2007; Aggarwal and Yu, 2009; Wolff et al., 2009).

A motivação inicial para o aprendizado semi-supervisionado foi a praticidade, desempenho e baixo custo, Zhu (2013). Com o enorme volume de dados e constante aumento devido ao crescimento tecnológico, é fácil encontrar dados de todos os tipos. No entanto estes dados normalmente não estão classificados. Tendo o conhecimento do custo de classificação destes dados manualmente, é possível escolher uma parte destes dados classificados ou classificá-los manualmente, e utilizá-los como exemplo para propagar a classificação aos demais. E, ainda assim, obter uma classificação com uma boa precisão. Este tipo de aprendizado se aproxima do modo que os seres humanos reconhecem e aprendem o que ocorre a sua volta, tendo em vista que grande parte de tudo o que existe é desconhecido durante um período, porém depois de analisados e comparados com outras situações acabam por se tornar comuns e simples.

Todavia há dados que não se adequam perfeitamente ao aprendizado de máquina semi-supervisionado, segundo Lu (2009) há tipos de dados em que o aprendizado supervisionado não pode ser substituído pelo aprendizado semi-supervisionado devido a distribuição e características presentes na base de dados. Outro ponto levantado por Lu (2009) é a possibilidade de que em algumas bases de dados a utilização do aprendizado semi-supervisionado, possa gerar um aprendizado errôneo quando as suposições criadas pela técnica são deturpadas.

O exemplo citado na Figura 3.1 apresenta uma distribuição de dados em que o aprendizado supervisionado não apresentaria bons resultados utilizando apenas os dados rotulados para treinamento.

O algoritmo trabalha com a suavidade, oscilação de características, presente entre dois dados, onde a classificação ocorre devido a proximidade e/ou distância entre os dados comparados. Permite-se também trabalhar com clusters, onde dados que estejam no mesmo grupo estão suscetíveis a estarem na mesma classificação. Nestes casos, as classes, apesar de separadas, não são fortemente espaçadas, podendo apresentar uma leve separação.(Barber, 2012)

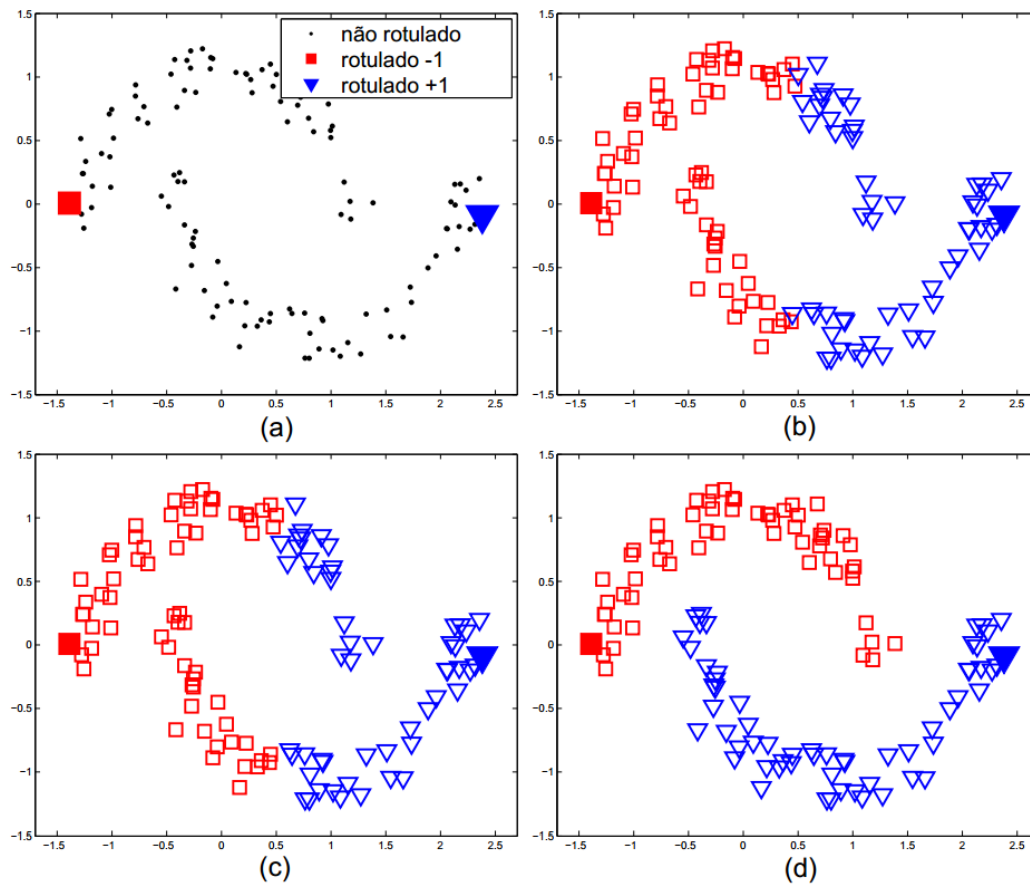
O aprendizado semi-supervisionado é uma mescla de características entre um aprendizado não supervisionado, onde apresenta uma parte orientada pelas características dos dados conhecidos, ao mesmo tempo que há a necessidade de reconhecer um padrão entre os dados desconhecidos.(Chapelle et al., 2006) Dentre as principais categorias de métodos de aprendizado semi-supervisionado podemos destacar: Generativos, Clusterização e Rotulação, Co-training, Separação de baixa intensidade e Baseado em grafos. Tais métodos serão discutidos nas subseções adiante.

3.1.1 Métodos Generativos

A categoria mais antiga no aprendizado semi-supervisionado, em que os métodos assumem um modelo $p(x, y) = p(y)p(x|y)$ onde $p(x|y)$ é uma distribuição com mistura identificável como, por exemplo, um modelo de mistura Gaussiano. Com uma grande quantidade de dados não rotulados os componentes da amostra podem ser identificados (Nigam et al., 2000; Fujino et al., 2005).

A partir desta amostra é selecionada a borda de vizinhos mais distantes, no entanto as quais ainda pertençam a classe, utilizando uma grande quantidade de dados não rotulados. É necessária a identificação de pelo menos um nó pertencente a cada classe em análise, e assim apresentar uma melhor classifica-

Figura 3.1: Classificação no conjunto de dados com dois objetos. (a) Dados de entrada;(b) resultado da classificação utilizando o Support Vector Machines (SVM);(c) Classificação utilizando k -vizinhos; (d) classificação ideal.



Fonte: (Zhou et al., 2004)

ção. Um fator que pode atrapalhar o algoritmo é encontrar classes diferentes misturadas.(Chapelle et al., 2006)

3.1.2 Clusterizar e Rotular

Este método tem como objetivo clusterizar a base de dados e em seguida rotular cada cluster de acordo com os dados rotulados. Esta técnica atua bem quando o algoritmo de clusterização é compatível com a estrutura dos dados, mas é mais difícil de analisar por sua natureza algorítmica (Demiriz et al., 1999; Dara et al., 2002)

3.1.3 Co-Training

Estas técnicas assumem que:

- a) os atributos podem ser divididos em dois sub-grupos;
- b) cada sub-grupo é suficiente para treinar um bom classificador;
- c) os dois conjuntos são condicionalmente independentes dada a classe.

Cada classificador é treinado no seu respectivo grupo. Com o objetivo de classificar os dados não rotulados e “ensinar” o outro classificador usando os rótulos preditos. Cada identificador é re-treinado com os dados do outro rotulador e o processo se repete até o final da execução (Blum and Mitchell, 1998; Mitchell, 1997, 1999; Zhou and Li, 2005, 2007). Os classificadores devem utilizar dados rotulados e não rotulados, assim de acordo com o treinamento dos dados é dada a confiabilidade da resposta.

A técnica de Co-training tem o objetivo de modelar a relação entre as variáveis dependentes e independentes, selecionar quais rótulos serão utilizados para um melhor treinamento, afim de abranger todas as classes e permitir uma boa classificação. Finalizando a etapa anterior são adicionados pesos às classificações obtidas para que haja uma classificação mediana, a qual será utilizada nos demais dados. (Hu et al., 2012)

Os classificadores não consideram a existência de redundância apresentada por características ou divisão entre as classes. Para esta identificação pode ser necessário o uso de algoritmos para seleção de características, com a finalidade de remover as repetições. O Co-Training estuda a divisão de recursos para ensinar e preparar os classificadores distribuídos entre os conjuntos, utilizando o fato de os dois classificadores deverem concordar com as rotulagem encontradas para se fazer uma boa aprendizagem. (Liu, 2007)

A intenção do Co-Training é ser iterativo, a cada interação é possível modelar um classificador, tendo como base um conjunto de dados, sendo marcado os recursos e assim rotular os nós não rotulados, terminando o processo quando todos os nós receberam uma classificação. (Liu, 2007)

3.1.4 Separação de baixa densidade

O método de separação de baixa densidade, assumem que a separação entre as classes se encontra em uma região de baixa densidade e, portanto, tentam empurrar a fronteira de decisão para longe das áreas de alta densidade. A abordagem mais comum é usar um algoritmo que maximiza a margem, como as máquinas de vetor de suporte¹. Tal método, quando usado para maximizar margens para dados rotulados e não rotulados, é chamado Transductive SVM (TSVM) (Vapnik, 1998).

Com o auxílio dos dados rotulados é estudado uma regra para ser utilizada como meio de classificação da base de dados. Este mecanismo difere do aprendizado supervisionado por utilizar um número finito de rótulos para o aprendizado (Barber, 2012). A intenção deste algoritmo é fortalecer a margem de dados rotulados, na tentativa de que, melhorando a classificação nas fronteiras entre objetos da imagem, distinguindo melhor as regiões.

3.1.5 Métodos Baseados em Grafos

A categoria dos métodos baseados em grafo tem sido mais ativa nos anos recentes. Esta categoria expoe os dados representados por nós de um grafo e as arestas as quais refletem as similaridades entre eles. Tais métodos assumem que existe suavidade dos rótulos no grafo. A maioria dos métodos podem ser vistos como um framework de regularização que estima uma função f que satisfaça tanto a suavidade dos rótulos quanto a proximidade dos rótulos estimados aos dados pré-rotulados (Zhu et al., 2003; Zhou et al., 2004; Wu and Schölkopf, 2007; Wang and Zhang, 2008; Szummer and Jaakkola, 2002; Grady, 2006; Zhu and Ghahramani, 2002; Wang et al., 2008).

¹ do inglês *support vector machines* ou SVM

Dentro de um grafo os nós consistem na marcação dos pontos de dados, onde as arestas representam as semelhanças entre pontos. As arestas são utilizadas como caminho para que os rótulos se propaguem através do grafo de maneira a tentar descobrir os rótulos desconhecidos. Os métodos que se baseiam em grafo permitem a conversão dos dados em uma matriz, onde o objetivo é a minimização de custo da classificação geral. O critério de custo é introduzido quando ocorrem mudanças suaves nos dados, onde os dados mais próximos de uma mesma classificação tendem a ter como vizinhos próximos dados da mesma classe (Liu and Lu, 2008). Quando um grafo está sendo analisado, e são encontrados dois nós cuja aresta que os ligam possui um peso baixo, a propensão é que estes sejam nós pertencentes a diferentes classes ou a mesma classe, mas com características distintas.

Ao finalizar a montagem do grafo, é possível extrair respostas previamente, como por exemplo que nós próximo uns aos outros no interior do grafo tendem a pertencer a mesma classificação. (Camps-Valls et al., 2007). Para a extração de conteúdos, ou melhor, encontrar formas de propagar os rótulos entre os nós, podemos utilizar diversas formas. Dentre elas, aplicar um passeio aleatório, campos randômicos gaussianos, nós globais e locais e diversos outros métodos, podendo ocorrer possíveis junções dentre estes.

A estrutura do grafo pode ser representada por $G = (V, E)$, onde V representa o conjunto de nós, e E representa os vértices, onde estes nós e vértices são posicionados de acordo com a distância calculada entre os nós, e de acordo com a semelhança possui o vértice interligando-os.

O modelo de competição e cooperação entre partículas que será utilizado neste projeto de pesquisa se enquadra na categoria dos métodos baseados em grafos. Ele será descrito com maiores detalhes no capítulo a seguir.

Capítulo 4

Modelo de Movimentação de Partículas

O modelo de movimentação de partículas foi desenvolvido recentemente com o objetivo de agrupar dados, inicialmente utilizando aprendizado não supervisionado (Quiles et al., 2008). Posteriormente foi estendido para realizar aprendizado semi-supervisionado utilizando as partículas para propagação de rótulos. O modelo é utilizado em bases de dados nas quais apenas uma parte dos dados são rotulados. Tal categoria de aprendizado de máquina foi descrita no Capítulo 3.

Como vimos na seção 3.1, os métodos baseados em grafos permitem uma boa identificação das classes, e apesar de apresentarem uma boa classificação, a maioria deles apresentam também uma complexidade computacional alta $O(n^3)$, tornando seu uso limitado a pequenas bases de dados. (Breve, 2010)

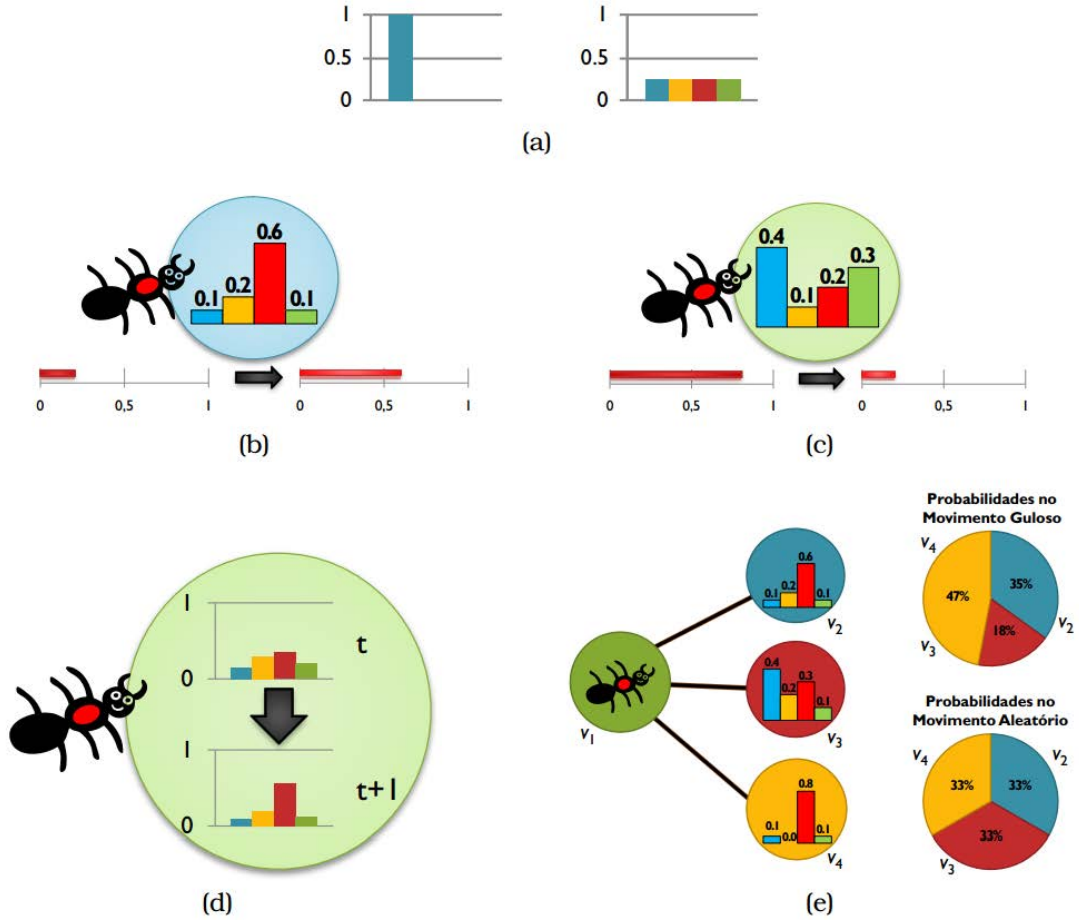
No modelo de movimentação de partículas, as partículas caminham em uma rede e competem umas com as outras de forma que cada uma tenta possuir a maior quantidade possível de nós. Ao mesmo tempo, cada partícula evita que outras partículas invadam seu território. Finalmente, cada partícula permanece confinada dentro de uma comunidade da rede.

O algoritmo proposto por Breve et al. (2012, 2009) é capaz de realizar classificação utilizando o aprendizado semi-supervisionado. De acordo com os trabalhos de Breve et al. (2012) este algoritmo produz resultados obtidos em simulações, onde este é capaz de classificar dados linearmente inseparáveis, em um tempo de execução menor que outros algoritmos, devido a menor complexidade computacional.

Na versão do modelo proposto por Breve et al. (2012) a partícula possui um nó casa, e mantém informações de distância entre a mesma e os demais nós do grafo. utilizada na forma de mover-se dentro do grafo, esta caminhada é determinada alternando a dinâmica de escolhas de nós a serem visitados pela partícula, levando em consideração a distância, entre a partícula e vizinhos do nós em que a mesma se encontra. A Figura 4.1 exemplifica as interações que ocorrem entre partículas.

Ao contrário dos modelos tradicionais, que utilizam o modelo de aprendizado de máquina semi-supervisionado baseado em grafo em que cada passo do algoritmo tem informação propagada de todos os nós para todos os demais nós levando em consideração os pesos das arestas, o novo modelo é baseado em cooperação e colaboração entre partículas, com uma propagação de rótulos de forma local, ou seja, a cada passo do algoritmo cada partícula escolhe apenas um dos vizinhos para propagação. De acordo com o método proposto por Breve et al. (2012), cada partícula visita nós que potencialmente pertençam ao seu time, evitando nós que já estejam dominados por outra partícula. Este modelo permite que cada partícula se concentre na região e proximidades da sua classe, dividindo a base de dados, e assim diminuindo a complexidade computacional, quando comparada com outros modelos baseados em grafos.

Figura 4.1: Dinâmica de partículas e nós. (a) Níveis de domínio inicial, (esquerda) correspondente a uma amostra rotulada, e (direita) um amostra não rotulada. (b) Representa o fortalecimento da partícula ao ter como objetivo um nó que esta sendo dominado pelo seu time. (c) o enfraquecimento de uma partícula ao receber como objetivo um nó dominado por outra partícula. (d) uma partícula aumenta o seu nível de domínio do seu time enquanto diminui o valor de domínio de outro time. (e) probabilidade de serem escolhidos o movimento guloso e o movimento aleatório por uma partícula, quanto todos os nós candidatos possui a mesma distância do nó-casa da partícula.



Fonte: (Breve et al., 2012)

Utiliza-se um conjunto de dados $X = x_1, x_2, \dots, x_n \subset R^m$ a ser analisado e transformado em uma rede sem pesos e não direcionada, e um conjunto de rótulos $L = 1, 2, \dots, c$, onde estes rótulos são escolhidos para que sejam anexadas marcações de cada um dos times.

Com tais informações é montada uma rede, representada através de um grafo $G = (V, E)$ com $V = v_1, v_2, \dots, v_n$ onde cada nó v_i corresponde a um elemento da rede. A matriz de adjacência W define quais nós da rede são interconectados. Para a representação em matriz é utilizada a Equação 4.1 onde W_{ij} especifica se há aresta entre os nós x_i e x_j .

$$W_{ij} = 1 \text{ se } \|x_i - x_j\|^2 \leq M \text{ e } i \neq j$$

$$W_{ij} = 0 \text{ se } \|x_i - x_j\|^2 > M \text{ e } i = j$$

$$W_{ii} = 0 \tag{4.1}$$

A Formula 4.1 determina quando haverá arestas entre um par de nós, onde M define a distância máxima entre x_i e x_j tal que x_i e x_j estarão conectados por uma aresta. No caso verdadeiro $W_{ij} = 1$ e $W_{ij} = 0$ caso contrário. Uma segunda forma de atribuir a existência da vizinhança é a comparação da distância entre cada pixel da imagem, dessa forma os nós que forem mais próximos do pixel analisado serão atribuídos à vizinhança do pixel. Se x_j esta entre os vizinhos mais próximos de x_i , mesmo que x_j não esteja entre os k vizinhos mais próximos de x_i , ele, é adicionado na lista de vizinhos de x_j devido a vizinhança mútua, pois o grafo formado é não direcionado..

Posteriormente é verificado o conjunto de dados rotulados, e estes serão anexados à rede como partículas, onde cada subconjunto de partículas tem o mesmo rótulo ou time, representada por $P = (p_1, p_2, \dots, p_c)$. Cada partícula p_i possui uma variável que representa o potencial da partícula representada pela Equação 4.2 e por padrão utilizamos a Equação 4.3 cuja posição inicial será este mesmo nó, que chamaremos de “nó casa” da partícula. Conforme as partículas mudam de posição, a distância do nó atual para o nó casa é registrada. Partículas geradas de amostras da mesma classe agem como um time, colaborando com as outras e competindo com partículas de outros times.

$$p_j^\omega(t) \in [\omega_{min} \ \omega_{max}] \quad (4.2)$$

$$\omega_{min} = 0 \text{ e } \omega_{max} = 1 \quad (4.3)$$

Cada nó da rede tem um vetor de tamanho L , $v_i^\omega(t) = v_i^{\omega_1}(t), v_i^{\omega_2}(t), \dots, v_i^{\omega_l}(t)$, em que cada elemento $v_i^{\omega_j}(t) \in [\omega_{min} \ \omega_{max}]$ representa o nível de dominação do time l sobre esse nó v_i . A soma desse vetor é sempre constante, $\sum_{l=1}^c v_i^{\omega_l} = 1$, devido ao fato de que quando a partícula aumenta o seu nível de domínio ela diminui o nível de domínio de outro time de partículas.

Outra variável que cada partícula recebe é um vetor de distâncias $p_j^d(t) = p_i^{d1}(t), p_i^{d2}(t), \dots, p_i^{dn}(t)$, onde cada elemento $p_j^{di}(t) \in [0, n - 1]$ equivale à distância medida entre o nó-casa da partícula p_j e o nó v_i .

Conforme o sistema executa, as partículas caminham pela rede e aumentam o nível de domínio de seu time sobre o nó, ao mesmo tempo em que diminuem o nível de domínio de outros times, sempre mantendo a soma constante. Além disso, cada partícula também tem um nível de força, que aumenta quando ela visita um nó dominado por seu time, e diminui quando ela visita um nó dominado por outro time. Essa força é importante porque a mudança que a partícula causa em um nó (aumento do nível de dominância do seu time) é proporcional à força que ela possui naquele momento.

Em cada interação, cada partícula seleciona um vizinho para visitar dentre os k vizinhos mais próximos, seguindo duas formas de seleção descritas mais adiante. Selecionado o nó e a ser visitado, a partícula eleva o domínio no nó da classe que ela representa, analisando em seguida qual a maior força dentro do vetor de forças do nó. Para o caso da classe da partícula ser a mesma classe de maior força do nós, este altera seu “nó atual” para o nó visitado, caso contrário ela permanece no nó atual, e atualiza-se a força da partícula, repetindo novamente o ciclo.

Tal mecanismo garante que a partícula fique mais forte quando está em sua vizinhança, protegendo-a; e fique mais fraca quando está tentando invadir outros territórios. Inicialmente, todas as partículas têm sua força configurada no nível máximo e todos os nós rotulados têm o nível de dominação configurado no máximo para o time correspondente, enquanto os demais nós tem nível de domínio distribuído

igualmente entre os times, seguindo aa Equação 4.4. As alterações de valores de força da partícula, sua inicialização e probabilidade de escolhas de método para selecionar o próximo vizinho são exemplificados na Figura 4.1.

$$\begin{aligned} v_i^{\omega l} &= 1 \text{ se } y_i = l \\ v_i^{\omega l} &= 0 \text{ se } y_i \neq l \text{ e } y_i \in L \\ v_i^{\omega l} &= \frac{1}{c} \text{ se } y_i = 0 \end{aligned} \quad (4.4)$$

A tabela de distância introduzida em cada partícula tem como objetivo mantê-las com um controle da distância entre o “nó atual” e o “nó casa”, evitando assim que as partículas se distanciem muito deixando suas vinhanças e ficando suscetíveis a ataques de partículas pertencentes a outras classes. Para evitar esse distanciamento dentro do grafo, as partículas analisam a distância juntamente com o domínio do nó, aumentando a probabilidade de escolher o nó que tem a maior chance de pertencer a mesma classe que a partícula.

Para se calcular a distância entre a partícula e cada nós inserido no grafo, assume-se inicialmente um conhecimento limitado da rede, a partícula sabe que no grafo há n nós. No entanto, não sabe exatamente a distância entre o seu “nó casa” e cada um dos demais nós, assim a partícula inicializa a distância com todos os nós como $n - 1$. A cada interação a partícula $p_{(j)}$ irá atualizar a distância $p_{(j)}^{d(k)}(t)$, a cada interação t , com base na Equação 4.5, a partícula irá atualizar a distância no vetor de distâncias com base na posição do nó anterior. Essa é uma tentativa de gravar a menor distância possível entre o “nó casa” da partícula e os demais nós do grafo.

$$\begin{aligned} p_{(j)}^{d(k)}(t+1) &= p_{(j)}^{d(k)}(t) + 1 \text{ se } p_{(j)}^{d(k)}(t) + 1 < p_{(j)}^{d(k)}(t) \\ p_{(j)}^{d(k)}(t+1) &= p_{(j)}^{d(k)}(t) \text{ caso contrário} \end{aligned} \quad (4.5)$$

As partículas escolhem os nós a serem visitados com base em uma de duas regras. A cada interação a partícula escolhe uma das regras aleatoriamente com probabilidades pré-definidas. As duas regras são descritas a seguir:

Regra aleatória: a partícula escolhe aleatoriamente, com iguais probabilidades, qualquer um dos nós vizinhos da posição em que a partícula se encontra. A regra aleatória não se preocupa com os níveis de domínio ou distância do nó-casa com os possíveis nós a serem visitados, sendo útil para exploração e aquisição de novos nós. A aplicação da caminhada aleatória é definida pela Equação 4.6, onde q se refere ao nó atual da partícula p_j .

$$p(v_i|p_j) = \frac{Wq^\mu}{\sum_{\mu=1}^n Wq^\mu} \quad (4.6)$$

Regra gulosa: a partícula escolhe aleatoriamente qualquer um dos nós vizinhos do nó em que ela se encontra, com probabilidades calculadas de forma diretamente proporcionais ao nível de dominância do time dessa partícula em cada nó, e inversamente proporcionais à distância de cada nó ao nó casa dessa partícula. O movimento guloso se torna útil para a defesa do território do time, sendo a aplicação definida pela Equação 4.7 onde q se refere ao nó atual da partícula p_j e $l = p_j^i$ e p_j^i

corresponde ao rótulo da classe da partícula p_j .

$$p(v_i|p_j) = \frac{W_{qi} v_i^{\omega_l} \frac{1}{(1+d_j^{d_i})^2}}{\sum_{\mu=1}^n W_{q\mu} \frac{1}{(1+d_j^{d_i})^2}} \quad (4.7)$$

A cada interação, cada partícula possui a probabilidade de p_{grd} de escolher o movimento guloso e a probabilidade de $1 - p_{grd}$ de escolher o movimento aleatório, com $0 \leq p_{grd} \leq 1$.

Dessa forma, ao escolher a regra gulosa, as partículas tendem a ficarem em territórios já dominados e preferem nós próximos de sua origem (em sua vizinhança), ou seja, elas assumem um comportamento mais conservador e defensivo. Por outro lado, ao escolher a regra aleatória as partículas ignoram níveis de dominância e distâncias e tem maior probabilidade de escolher nós que não dominam e nós distantes de sua origem, assumindo assim um comportamento de exploração.

Note que uma partícula só permanece no nó escolhido se conseguir que seu time seja o dominante naquele nó, caso contrário ela é expulsa e volta para o nó onde estava até o próximo instante de tempo. Essa regra evita que uma partícula deixe seu território ao tentar explorar inutilmente um território alheio onde ela rapidamente perderia sua força e não conseguiria provocar grandes mudanças. Esta regra também garante que um time só consegue dominar um nó se já tiver dominado ao menos um vizinho deste nó. Ao final das iterações o rótulo de cada nó não rotulado é definido com base nos níveis de dominância de cada time sobre esse nó, de forma que ele passará a ter o rótulo do time dominante.

O critério de parada segue o monitoramento dos domínios dos nós, o qual é feito analisando o comportamento das partículas. O monitoramento é feito sobre a média de conquista das partículas, quando as partículas não apresentarem melhoras no critério de conquista em um determinado intervalo de interações, é entendido pelo algoritmo que já houve uma separação entre as classe. Normalmente essa parada ocorre após um grande número de interações dentro da rede, na ordem de 10^5 , dependendo do tamanho da rede (quantidade de amostras) e da quantidade de partículas (quantidade de amostras rotuladas).

Estes modelos baseados em movimentação de partículas ainda oferecem muitas possibilidades de aperfeiçoamento e aplicação em outros problemas da computação. O objetivo deste projeto de pesquisa é estender o modelo para tratar o problema de segmentação de imagens de maneira interativa. Deste modo, partículas seriam criadas para pixels da imagem rotulados pelo usuário e, através dos mecanismos de competição e cooperação, tais rótulos poderiam se espalhar para os demais pixels.

4.1 Metodologia

Esta seção tem como objetivo explicar e detalhar as alterações necessárias para que o algoritmo pudesse reconhecer e classificar imagens coloridas, pois o algoritmo foi projetado para classificação de dados baseados em vetores de atributos, e ainda não havia sido aplicado em tarefas de segmentação de imagens.

Em virtude da mudança no formato dos dados de entrada foi necessário um estudo sobre as formas de se analisar uma imagem. Primeiramente foi realizado um estudo acerca de Segmentação de imagens, o qual foi exposto no Capítulo 2, para uma melhor compreensão do funcionamento dos algoritmos existentes e manipulação das imagens. Ao se trabalhar com segmentação de imagens faz-se necessário realizar uma análise detalhada, identificando oscilações nas cores, intensidades, brilho e tipo de imagem a ser

utilizada. Tal estudo é realizado para uma melhor percepção sobre a imagem com o objetivo de obter uma melhor classificação.

Devido ao fato do algoritmo de competição e cooperação entre partículas utilizar o aprendizado semi-supervisionado, também foi necessário um estudo a respeito deste tema. O estudo acerca do aprendizado de máquina, bem como suas principais categorias, diferenças e algoritmos, está exposto no Capítulo 3, aprofundando-se mais no aprendizado semi-supervisionado, onde o conceito, algoritmos, análise de aplicação e funcionamento, estão expostos na Seção 3.1.

Com a análise conceitual envolvida e necessária sobre aprendizado de máquina, passou-se para a próxima etapa de estudos, que consiste em analisar e entender mais profundamente o funcionamento do algoritmo de competição e cooperação entre partículas. Realizando uma análise da forma de recebimento e manipulação dos dados de entrada originais do modelo de competição e cooperação entre partículas, para posteriormente alterar as partes necessárias para se trabalhar com imagens com o modelo. O estudo realizado a respeito deste assunto está exposto no Capítulo 4.

O modelo de competição e cooperação entre partículas utiliza uma pequena porção de dados rotulados em meio aos dados não rotulados. Ao associar este modelo à tarefa de segmentação de imagens, é esperado que seja possível analisar e destacar partes ou objetos de uma imagem a partir de pequenas marcações feitas pelo usuário para direcionar o algoritmo.

4.1.1 Alteração do Modelo de Competição e Cooperação entre partículas

O modelo de competição e cooperação entre partículas é uma abordagem de aprendizado semi-supervisionado recente, que faz o uso de partículas que movem-se em um grafo, seguindo os mecanismos de competição e cooperação com intuito de demarcar territórios e propagar rótulos. Essa abordagem se destaca por apresentar uma propagação local de rótulos, feita através das partículas, em vez de uma propagação global como acontece na maioria dos métodos baseados em grafos.

No algoritmo de competição e cooperação entre partículas original, os dados vetoriais são transformados em uma rede complexa com base na distância Euclidiana entre os pares de amostras. No caso de imagens, as redes são geradas com base nas similaridades e diferenças entre características extraídas dos pixels, como por exemplo, intensidade de cor, média e desvio padrão quando considerado um raio de pixels e entre outros, alterando o valor e destaque de cada característica dependendo do tipo de imagem em questão, incluindo a distância entre os pixels analisados.

A princípio foi implementado o algoritmo original para a realização de testes com o uso de bases de dados simples, com o intuito de realizar um estudo sobre o funcionamento do algoritmo e auxiliar na familiarização com as funções de sua implementação. Após uma habituação com o algoritmo foram feitas análises sobre formas de se incorporar a segmentação de imagens ao mesmo.

Para a escolha do método de segmentação é necessário analisar o confinamento de grupos que o algoritmo proporciona e tipos de aspectos gerais que serão extraídos. Tais agrupamentos podem corresponder aos aspectos globais da imagem. No entanto é necessário analisar o tempo e critério de execução por linha de pixels, obtendo um método mais prático e eficiente.

As imagens mais complexas de segmentar são as coloridas. Para este tipo de imagem é necessário utilizar um processo de extração de regiões que possuam homogeneidade, podendo ser analisadas propriedades geométricas ou tons. É possível fazer uma análise com base nos valores RGB, por exemplo, estabelecendo uma medida de distância Euclidiana entre valores de pixels, e assim medir a distância

entre os tons.

A abordagem de segmentação de imagem utilizada nesta proposta, é baseada em grafos, devido ao fato desta abordagem apresentar melhor forma de manipulação junto ao algoritmo de competição e cooperação entre partículas. Neste tipo de abordagem cada pixel da imagem é representado como um nó do grafo, de forma que a presença ou falta de vizinhança determina a afinidade entre eles.

O objetivo da segmentação é definir a forma em que serão extraídas as informações e divisões presentes na imagem, resultando distinção da intensidade entre os pixels, dentro de uma determinada região. Algumas características podem ser extraídas com o conhecimento da distância entre nós e características comuns em nós próximos, possibilitando a identificação de diferenças na intensidade do brilho, cor, entre outros. Com a necessidade de estabelecer critérios que tornem possível distanciar pixels que representem diferentes objetos dentro da imagem, bem como aproximar pixels que represente o mesmo objeto ou semelhante, mesmo que este objeto seja representado em áreas separadas por demais objetos.

4.1.2 Inserção de Imagens ao algoritmo

Estudadas as formas de inserir imagens, foram realizados testes utilizando imagens sintéticas, criadas e manipuladas manualmente para simplificar o entendimento da máquina a respeito dos objetos/regiões presentes na imagem. Tais imagens foram utilizadas para analisar o comportamento do algoritmo e realizar as primeiras mudanças, como por exemplo, alterar o arquivo de entrada, onde primeiramente era utilizado um arquivo texto, com as informações separadas em linhas. Posteriormente o algoritmo foi alterado para que fosse recebido um arquivo de imagem, onde a informação dada consiste na distância entre os pixels e os valores extraídos de pixels para alguns pontos da imagem.

Assim que o algoritmo apresentou boas classificações neste tipo de imagem artificial foram inseridas imagens extraídas de bases de dados para testes. Para esta etapa de implementação, foram escolhidas as imagens de satélite (USGS, 2015), devido ao constante crescimento de uso, melhora da resolução, complexidade, qualidade e detalhes neste tipo de imagem.

Com a escolha do uso de imagens por satélite, foi preciso ajustar o cálculo da Distância Euclidiana do algoritmo de competição e cooperação entre partículas, ressaltando particularidades mais relevantes ao tipo de imagem. Para uma melhor segmentação de imagens de satélite faz-se necessário utilizar a textura, unida a informações adicionais e contextuais no processo de segmentação. Para melhorar a resposta da segmentação da imagem de satélite utilizando o algoritmo de competição e cooperação de imagens foi adicionado ao cálculo da distância Euclidiana os seguintes atributos:

- a) Posição do pixel na imagem;
- b) Posição dos pixels na imagem relativa a posição do pixel comparado;
- c) Comparação dos componentes RGB;
- d) Comparação dos componentes HSI;
- e) Média de vermelho em um raio de pixels;
- f) Média de azul em um raio de pixels;
- g) Média de verde em um raio de pixels;

- h) Média de saturação de cor em um raio de pixels;
- i) Média de intensidade em um raio de pixels;
- j) Média da matiz em um raio de pixels.

Considerando a análise de imagens de satélite os atributos que apresentam maior grau de relevância dentro da distância Euclidiana são os atributos referente as noções cromáticas presente nas regiões, como RGB e HSI. Informações como distância entre pixels acabam tendo pouca relevancia devido ao fato de que pode haver partes de uma mesma classe mais distântes do pixel analisado, porém é raro achar pixels com as mesmas características de tons presentes em mais de uma classe.

Assim como os atributos, os pesos atribuídos a cada parâmetro podem variar de acordo com o tipo de imagem a ser analisada. Em especial com o uso de imagens por satélite, por exemplo, é necessário levar em conta que podem ocorrer classes não contínuas, fazendo com que o peso da distância entre pixels seja menos relevante para o cálculo da Distância Euclidiana.

Para realizar a análise de qual parâmetro causa maior interferência na classificação e o melhor conjunto de parâmetros, foi utilizado o conhecimento extraído do estudo das técnicas de segmentação de imagem, possibilitando a identificação dos principais pontos de interferência na classificação. Feita esta análise prévia, realizaram-se testes referente a valores distintos em cada característica, analisando assim a resposta de segmentação, tipo de imagem classificada, tempo de execução e qualidade da resposta final.

Ao utilizar imagens de satélite como imagens de entrada, a melhor combinação de valores com os atributos analisados e testados, levando em consideração a qualidade de classificação e tempo de execução, esta exposta na Tabela 4.1.

Tabela 4.1: Tabela de parâmetros

Parâmetro	Valor
Comparação do HSI	2
Comparação RGB	8
Distância normalizada	8

Para avaliar os resultados e a qualidade de classificação do algoritmo de competição e cooperação entre partículas, usaremos a sobreposição de imagens, desta forma iremos sobrepor a imagem real e a imagem obtida com a análise, com o objetivo de melhorar a percepção referente a qualidade de classificação, balanceando as intensidades das imagens para que possam aparecer os detalhes presentes na classificação, destacando os acertos e erros.

4.1.3 Algoritmo de Competição e Cooperação entre Partículas aplicado a imagens

No decorrer deste capítulo foram expostas as alterações e modificações feitas no modelo de competição e cooperação entre partículas. Esta seção tem como propósito dar uma visão geral dos resultados obtidos pelo algoritmo.

Apoiando-se na representação da imagem em forma de grafo, com a aplicação dos parâmetros necessários ao algoritmo para análise do dados, são delimitadas as informações relevantes de cada nó seguindo

Tabela 4.2: Algoritmo

Algoritmo: Competição e cooperação entre partículas Ler a imagem; Identificar quais são os pixels demarcados como partículas; Construa a matriz de adjacência W usando as Equação 4.1 e ??, utilizando os atributos expostos na tabela 4.1; Atribua os níveis de dominância de cada nó utilizando a Equação 4.4; Determine as posições de cada partículas, e atribuindo a posição da mesma como nó casa, seguindo a matriz W; Atribua a cada partícula a distância máxima entre cada nó; Repita Para cada partícula faça Selecione entre a escolha randômica ou gulosa; Aplique a partícula a Equação 4.6 ou 4.7 de acordo com a seleção no passo anterior; Atualize a força e distância da partícula; Atualize a dominância do nó visitado; Final Até (Numero de execuções) Percorra cada nó para saber qual a classe dominante; Pinte o pixel com a cor da classe dominante; Grave no arquivo de saída a imagem gerada;
--

as necessidades do modelo de cooperação e competição entre partículas, seguindo com a teoria de representação das partículas competindo uma com as outras, quando pertencerem a classes diferentes, e cooperando, quando estas pertencerem as mesmas classes. A representação destas informações, assim como o passos de execução do algoritmo estão expressas na Tabela 4.2.

Capítulo 5

Resultados

Este capítulo tem como objetivo mostrar, através de diferentes imagens, os resultados obtidos em cada passo da alteração, além do uso e funcionamento do algoritmo de competição e cooperação entre partículas para a classificação de imagens.

O primeiro passo realizado foi a estruturação do algoritmo para o recebimento de imagens como dados de entrada. Tal alteração foi feita dividindo as imagens em pixels e analisando propriedades de cada um, para desta maneira compor um grafo obtendo ligações entre pontos que possuam maior tendência a pertencer a mesma classe dentro do espaço de n vizinhos.

A Figura 5.1c) representa o grafo montado pelo algoritmo, baseando-se na Figura 5.1a), onde cada pixel da imagem é representado por um nó do grafo e as marcações de rótulos, considerando cada pixel demarcado na imagem original, serão as partículas utilizadas no modelo de competição e cooperação entre partículas, partículas as quais estão apresentadas na Figura 5.1b).

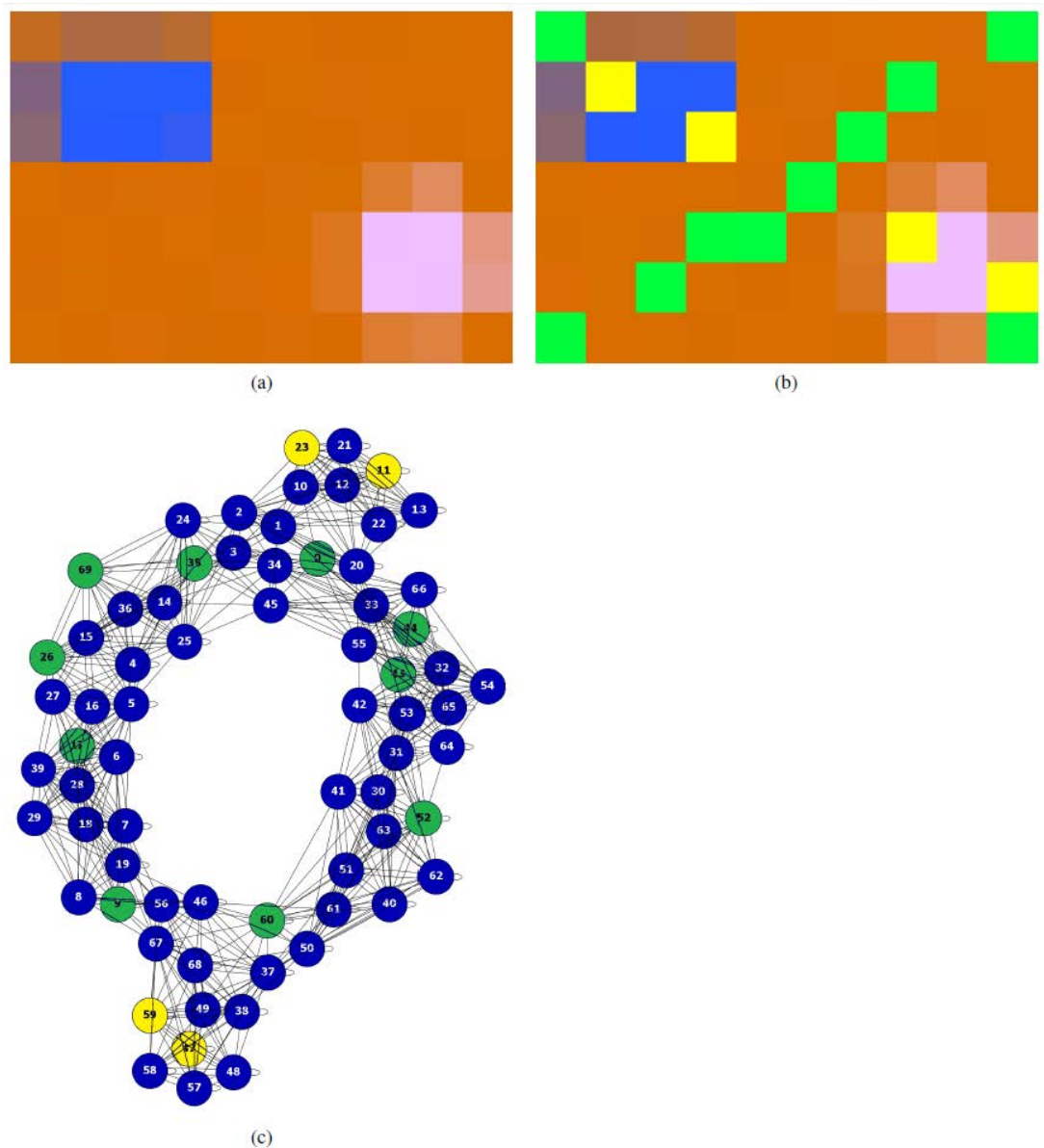
Para analisar o comportamento e modificações necessárias, foi utilizada a imagem 5.2a com as marcações de partículas expressas na Figura 5.2b. Utilizando variáveis que melhor influenciavam a Distância Euclidiana, obtendo como resposta a Figura 5.2c. Para uma melhor percepção de detalhes presente na classificação foi montada a Figura 5.2d).

Após a análise e inserção de valores extraídas dos componentes RGB dos pixels, foi necessário estabelecer um padrão de pesos para os atributos que melhor classificavam imagens de satélite, seguindo os valores expressos na Tabela 4.1 exposta na Seção 4.1. As Figuras 5.3, 5.4, 5.5, 5.6 e 5.7 apresentam algumas imagens de satélite submetidas a classificação utilizando o algoritmo de competição e cooperação entre partículas.

Ao observar e quantificar os acertos e erros encontrados na classificação da imagem de satélite utilizando o algoritmo de competição e cooperação entre partículas, foi encontrada uma proporção de acertos, nos casos estudados para esta dissertação, maior que 90%. Na Figura 5.2 a acurácia é em torno de 99%, com cerca de 24~30 pixels classificados incorretamente, e em imagens mais complexas como a 5.7 temos em torno de 96,99%, com 500~600 pixels classificados de forma incorreta, de um total de 19200 pixels. Porém quando consideramos imagens que apresentem cores semelhantes em mais de uma região observamos uma redução da qualidade de classificação, ficando com aproximadamente 96,12% de acurácia, na imagem 5.3.

Comparando a atuação do modelo de cooperação e competição entre partículas com outros algoritmos presentes na literatura para a segmentação de imagens, foram encontradas similaridade, melhoras e algumas restrições no nosso modelo. O modelo de segmentação semi-supervisionado GMMRF ("*Gaus-*

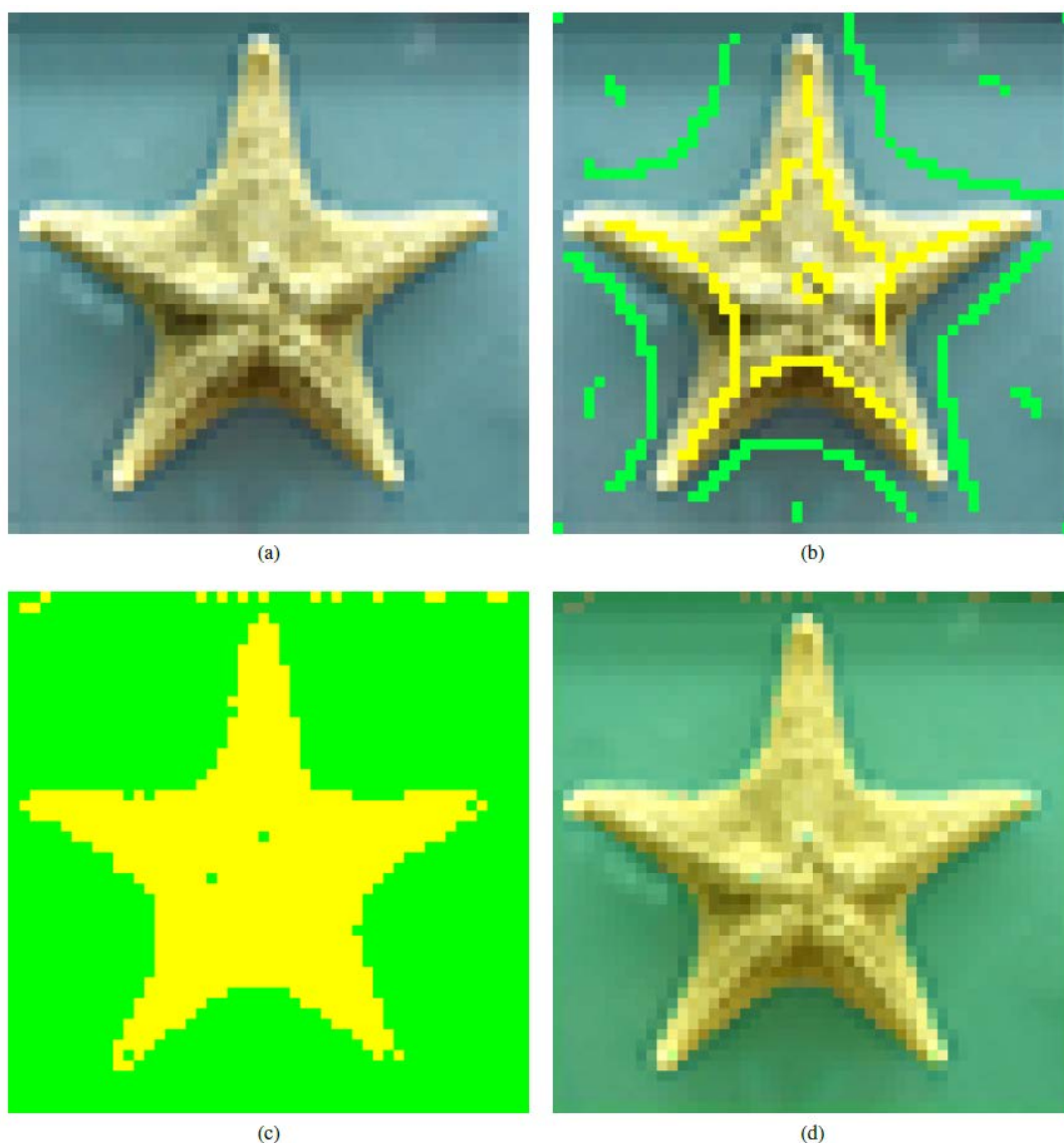
Figura 5.1: Grafo montado pelo algoritmo de competição entre partículas: a) Imagem de entrada; b) Imagem de entrada com as marcações de partículas; c) Grafo resultante das marcações realizadas através das partículas;



sian Mixture Markov Random Fiel") recebe uma imagem de entrada, 5.9a) adicionando uma prévia marcação de borda do objeto, no caso da Figura 5.8a), recebe a marcação em cinza, identificando assim qual a região do fundo (preta) e a região do objeto(branca).

Ao analisar as imagens de saída de ambos os algoritmos, 5.9c) 5.8b) é possível perceber alguns ruídos presentes nas imagens de classificação utilizando o nosso algoritmo. No entanto ao analisar apenas a classificação de borda, esta é muito similar entre os dois algoritmos. Ao observar a quantidade de dados previamente fornecidos, o modelo de competição e cooperação de partículas utiliza menos dados de entrada sem maiores prejuízos a classificação, apesar do algoritmo GMMRF não apresentar ruídos. A porcentagem de acerto encontrada na classificação quando utilizado o modelo de competição e cooperação entre partículas apresentou em torno de 92,47%, contra aproximadamente 95,17% do algoritmo proposto por Blake et al. (2004).

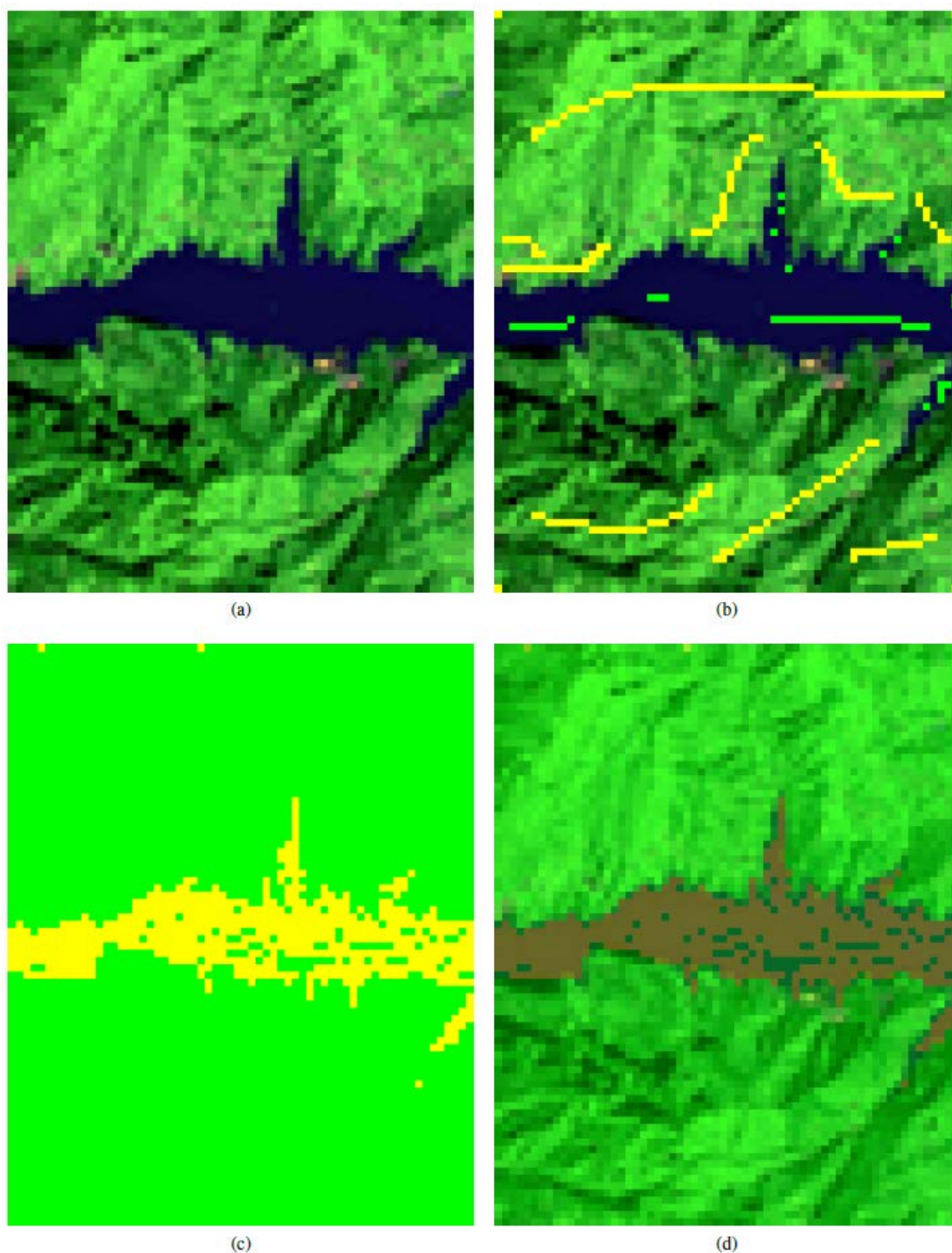
Figura 5.2: Imagem inicial 50X50 pixels: a) Imagem de entrada; b) Imagem de entrada com as marcações de partículas; c) Classificação da imagem usando o algoritmo de competição e cooperação entre partículas; d) Comparação entre a imagem de entrada e resposta de classificação;



O modelo de segmentação utilizando grafos Kernels, exposto no Harchaoui and Bach (2007), trabalha com o modelo de bacias hidrográficas, o qual recebe de entrada a imagem 5.10 a), com as demarcações de regiões expressas na imagem 5.11, em união com a resposta de segmentação adquirida com o modelo.

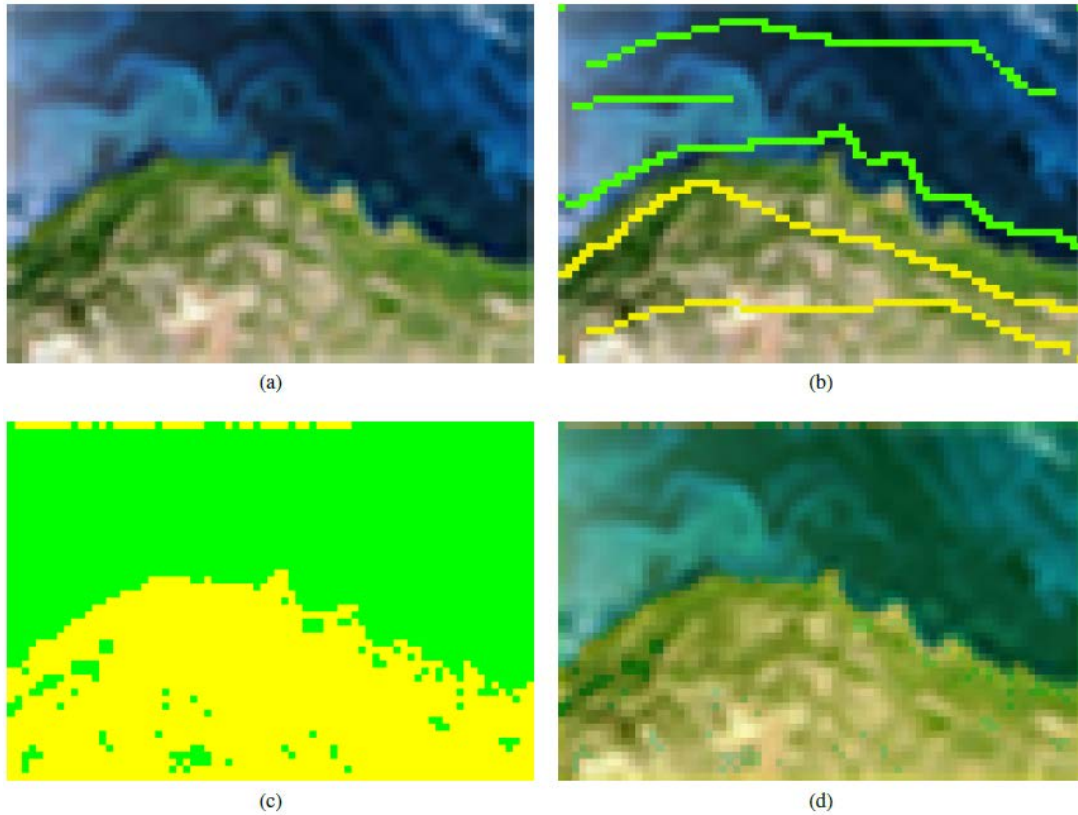
Examinando a imagem 5.10 a), com as demarcações expressas em 5.10 b), é obtida a resposta expressa na Figura 5.10 c). Comparando as duas imagens de saída, a Figura 5.10 c) utilizando o modelo proposto desta dissertação e a imagem 5.11, é possível perceber alguns detalhes ressaltados e destacados pelo modelo de competição e cooperação entre partículas que não foram percebidos pelo modelo de Harchaoui and Bach (2007), como bordas e detalhes presentes nos cavalos. No entanto há a presença de detalhes que foram melhor identificados pelo modelo comparado, como por exemplo a classificação das árvores como plano de fundo. A proporção de acerto para o uso do modelo de competição e cooperação entre partículas foi de 71,74% contra 85,40% do modelo de Harchaoui and Bach (2007).

Figura 5.3: Imagem de Satélite 64X80 pixels: a) Imagem de entrada; b) Imagem de entrada com as marcações de partículas; c) Classificação da imagem usando o algoritmo de competição e cooperação entre partículas; d) Comparação entre a imagem de entrada e resposta de classificação;



A Figura 5.12 mostra a atuação do modelo de competição e cooperação entre partículas a partir de uma imagem extraída de Irony et al. (2005). Analisando a classificação de regiões presente na Figura 5.12d) é possível encontrar bastante ruído, no entanto apresenta uma boa classificação quanto comparado com as bordas das classes. A Figura 5.13, esta permite uma melhor classificação das classes presentes na imagem, no entanto, para atingir esta classificação, utilizou-se um método baseado em aprendizado de máquina supervisionado. Comparando as respostas obtidas com os dois algoritmos, apesar de encon-

Figura 5.4: Imagem de Satélite 75X51 pixels: a) Imagem de entrada; b) Imagem de entrada com as marcações de partículas; c) Classificação da imagem usando o algoritmo de competição e cooperação entre partículas; d) Comparação entre a imagem de entrada e resposta de classificação;



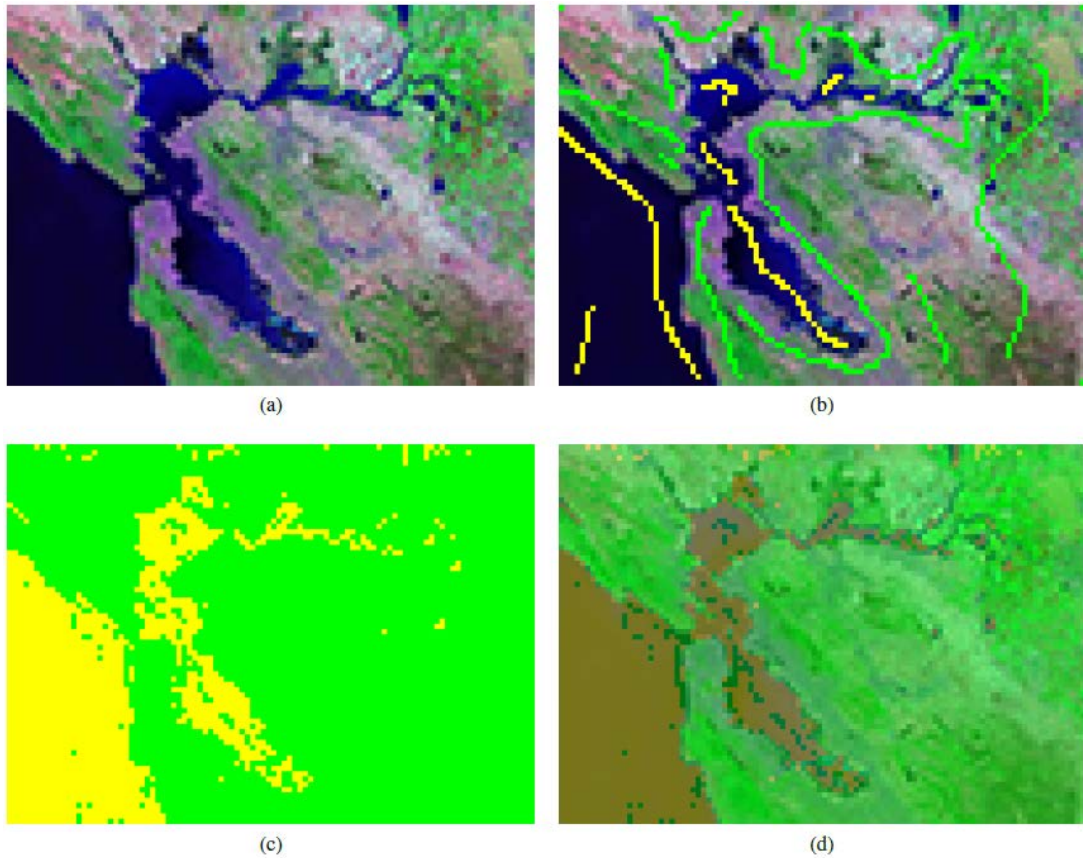
tra bastar ruído na classificação feito com o modelo aqui proposto, foram necessários menos dados de entrada, permitindo uma boa delimitação entre as classes.

A estrutura exposta no artigo Gómez-Chova et al. (2008) propõe o uso do SVM, *Support vector machine* aplicando o grafo não normalizado de Laplacian, ou Laplacian SVM. Com o intuito comparar o algoritmo Laplacian SVM e o modelo de competição e cooperação entre partículas, é utilizada a imagem 5.14 a). Após realizar a execução com o algoritmo de competição e cooperação entre partículas, é possível realizar a comparação entre a imagem 5.14c) com a imagem 5.15, tal análise permite notar pontos de melhor e inferior classificação entre os dois algoritmos. O algoritmo de competição e cooperação entre partículas apresentou uma boa delimitação entre as delimitações das regiões, no entanto os dois algoritmos apresentaram uma boa delimitação das bordas principais do objeto, mesmo que apresentando falha em ambos.

A Figura 5.16 mostra o uso da imagem extraída do artigo Gómez-Chova et al. (2008) processada com o uso do filtro mediana antes da execução do algoritmo de competição e cooperação entre partículas, seguindo as mesmas marcações feitas na imagem 5.14b). Em contrapartida na Figura 5.17 foi utilizado o filtro mediana após a execução do algoritmo proposto.

O filtro mediana, é utilizado para a redução de ruídos, o qual consiste em examinar e armazenar o valor RGB dos pixels pertencentes a um raio, r , do pixel analisado, x_i , este armazenamento ocorre em forma de lista com n posições. Após a criação da lista de pixel é feita uma ordenação da mesma e posteriormente é escolhido o valor central da lista para n ímpar ou $(n \div 2) + 1$ para n par, com o objetivo

Figura 5.5: Imagem de Satélite 100X72 pixels: a) Imagem de entrada; b) Imagem de entrada com as marcações de partículas; c) Classificação da imagem usando o algoritmo de competição e cooperação entre partículas; d) Comparação entre a imagem de entrada e resposta de classificação;



deste ser o valor que substituirá o valor inicial de x_i . No caso da aplicação do filtro sobre a Figura 5.16 foi utilizado um $r = 2$ pixels.

Comparando a resposta obtida para as Figuras 5.16 a 5.17 é possível perceber que ao processar a imagem utilizando o filtro mediana antes da execução, apareceram regiões em que houve mais erros do que quando a imagem é inserida sem correções, apresentando poucos pontos de melhora. Em compensação, ao processar a imagem de resposta, Figura 5.17, do algoritmo utilizando o filtro mediana, houve uma melhora considerável de acurácia, apresentando uma melhor classificação entre as classes.

O modelo de competição e cooperação entre partículas tem como resultado uma boa classificação das classes presentes na imagem, no entanto é necessário que a distribuição das partículas na imagem ocorra de forma a cobrir o máximo de cores e detalhes presentes. Desta forma as partículas terão maior probabilidade de acertar na classificação.

Figura 5.6: Imagem de Satélite 100X100 pixels: a) Imagem de entrada; b) Imagem de entrada com as marcações de partículas; c) Classificação da imagem usando o algoritmo de competição e cooperação entre partículas; d) Comparação entre a imagem de entrada e resposta de classificação;

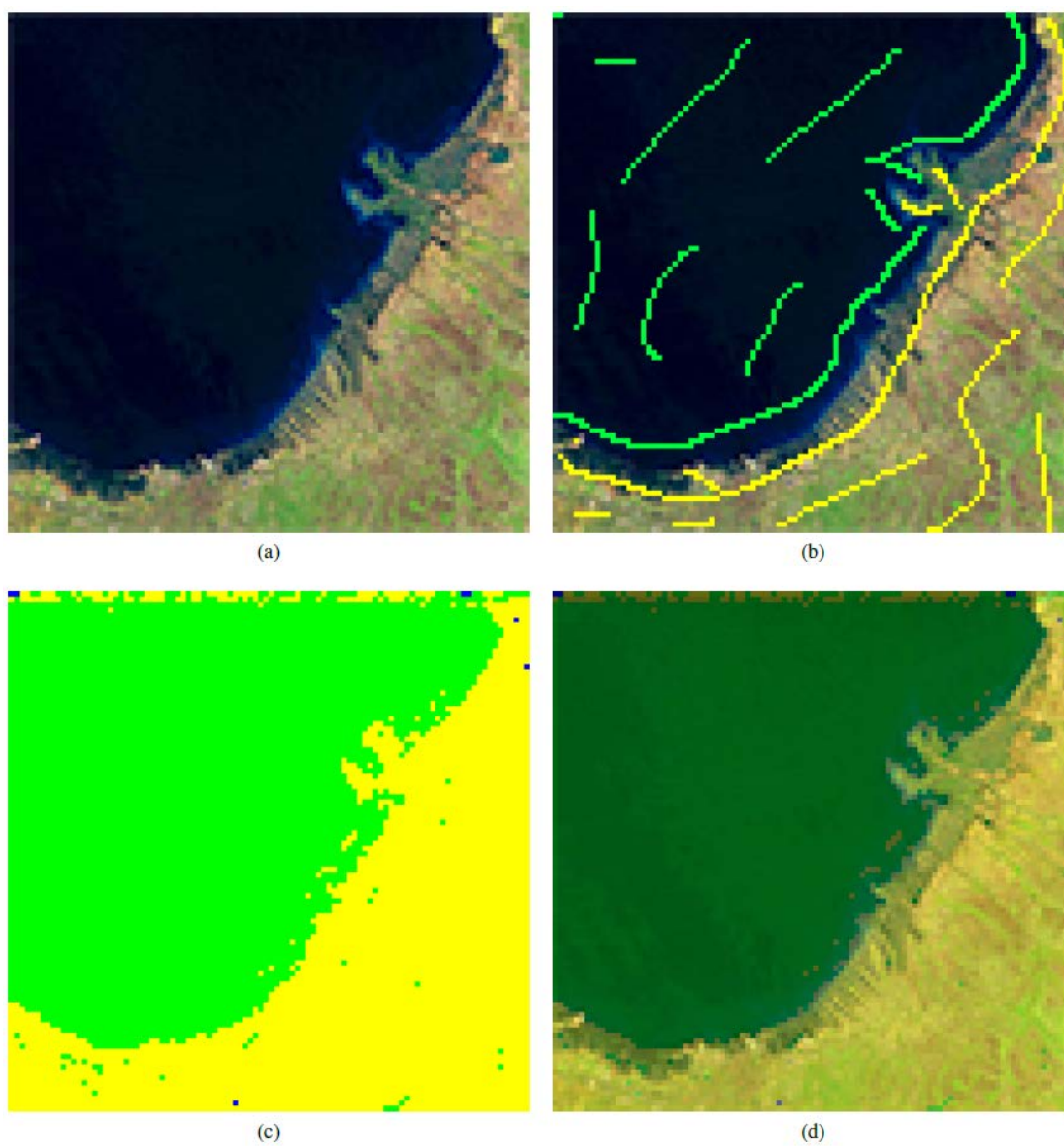


Figura 5.7: Imagem de Satélite 160X120 pixels: a) Imagem de entrada; b) Imagem de entrada com as marcações de partículas; c) Classificação da imagem usando o algoritmo de competição e cooperação entre partículas; d) Comparação entre a imagem de entrada e resposta de classificação;

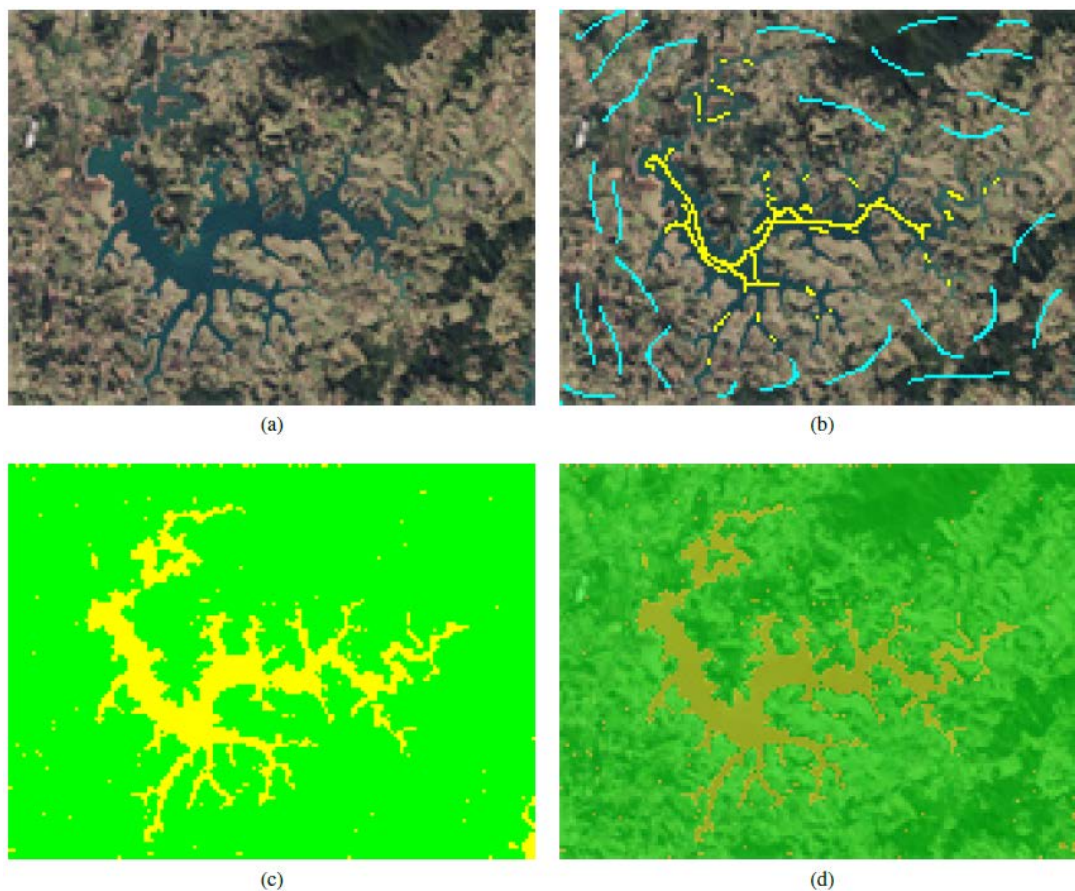
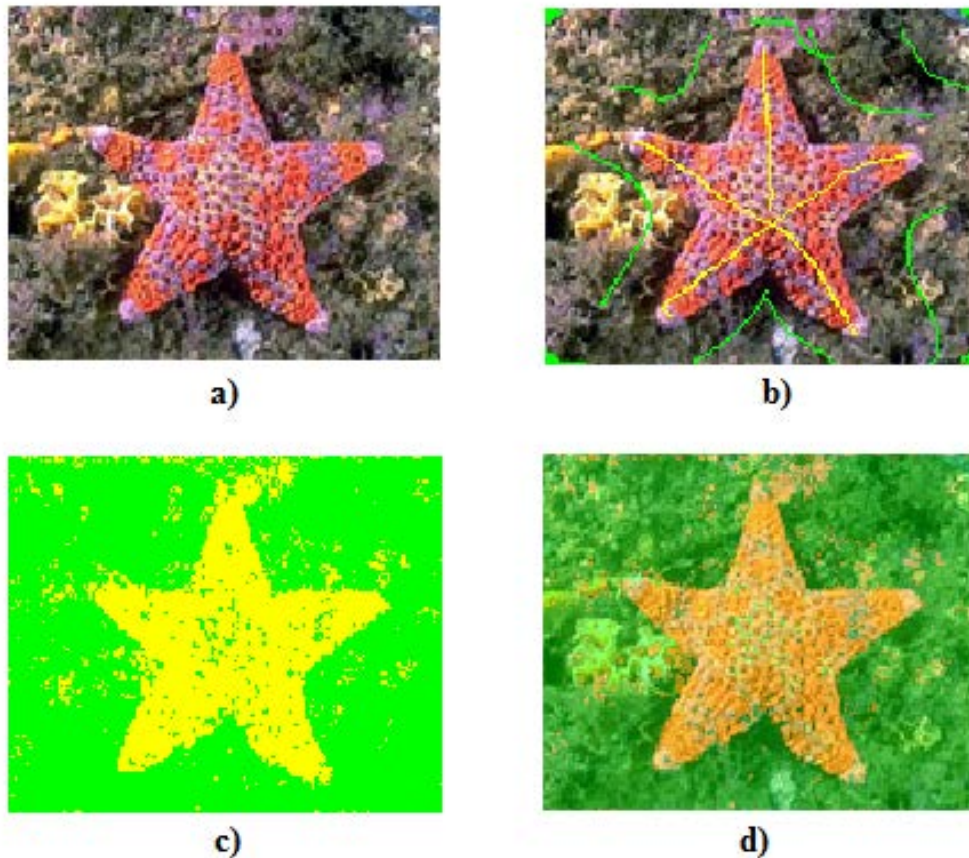


Figura 5.8: Modelo comparativo de segmentação: a) Imagem de entrada, com marcação de borda, cor Cinza; b) Imagem de saída;



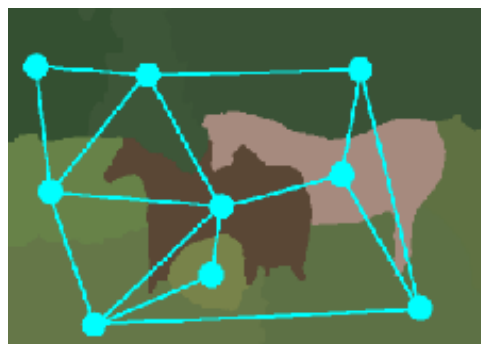
Fonte: Blake et al. (2004)

Figura 5.9: Imagem comparativa Blake et al. (2004): a) Imagem de entrada; b) Imagem de entrada com as marcações de partículas; c) Classificação da imagem usando o algoritmo de competição e cooperação entre partículas; d) Comparação entre a imagem de entrada e resposta de classificação;



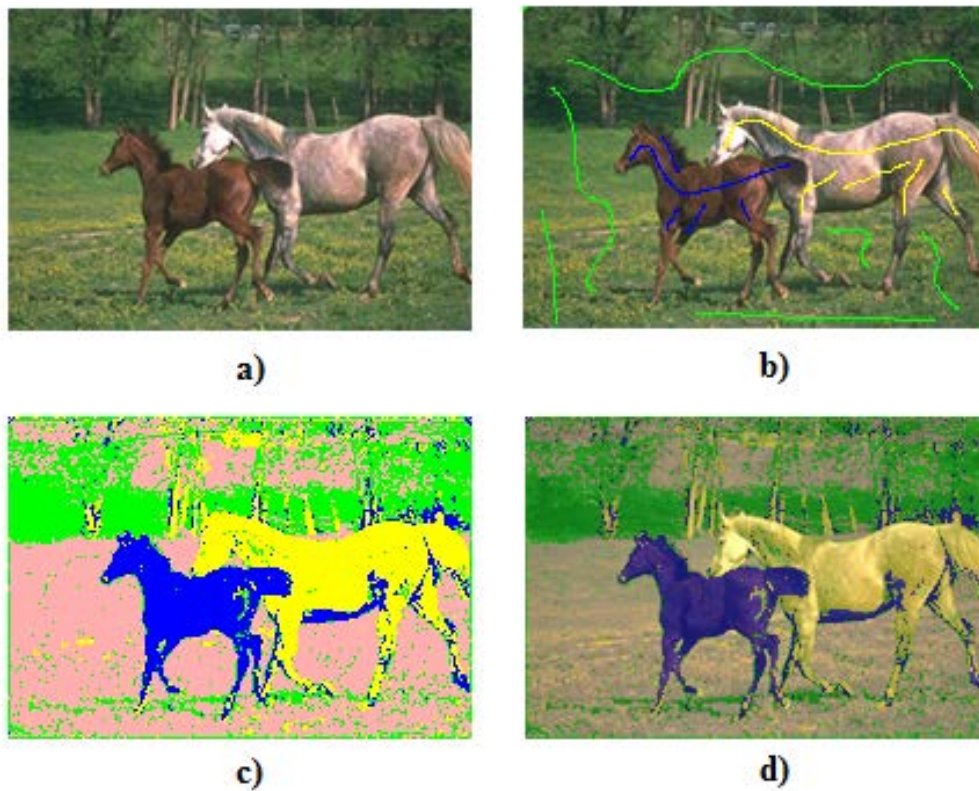
Fonte imagem a) Blake et al. (2004)

Figura 5.10: Utilizando a segmentação em mosaico, obtido utilizando a média RGB em cada região segmentada. O gráfico está representado associado a segmentação em verde claro.



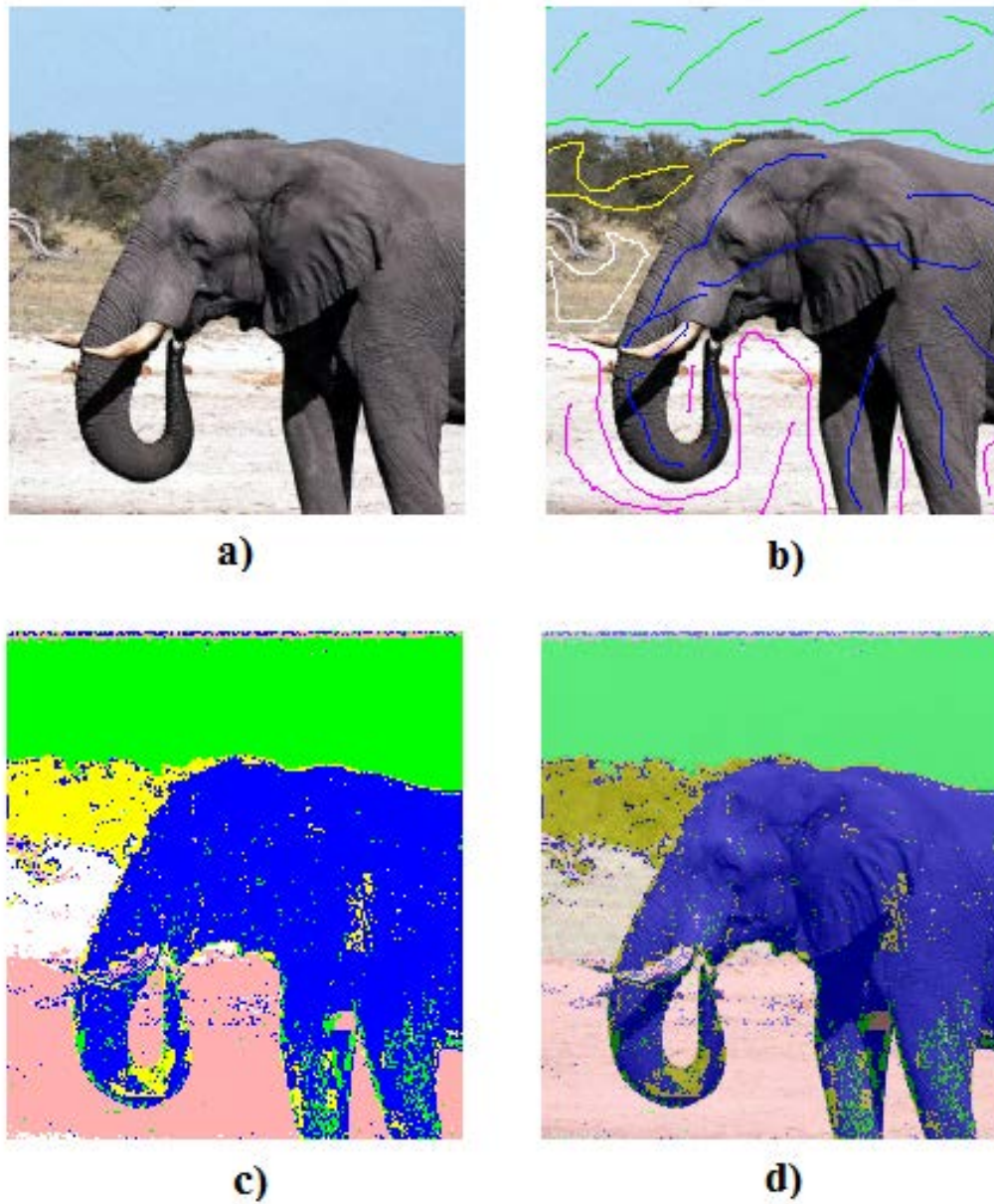
Fonte: Harchaoui and Bach (2007)

Figura 5.11: Imagem comparativa Harchaoui and Bach (2007): a) Imagem de entrada; b) Imagem de entrada com as marcações de partículas; c) Classificação da imagem usando o algoritmo de competição e cooperação entre partículas; d) Comparação entre a imagem de entrada e resposta de classificação;



Fonte imagem a) Harchaoui and Bach (2007)

Figura 5.12: Imagem comparativa Irony et al. (2005): a) Imagem de entrada; b) Imagem de entrada com as marcações de partículas; c) Classificação da imagem usando o algoritmo de competição e cooperação entre partículas; d) Comparação entre a imagem de entrada e resposta de classificação;



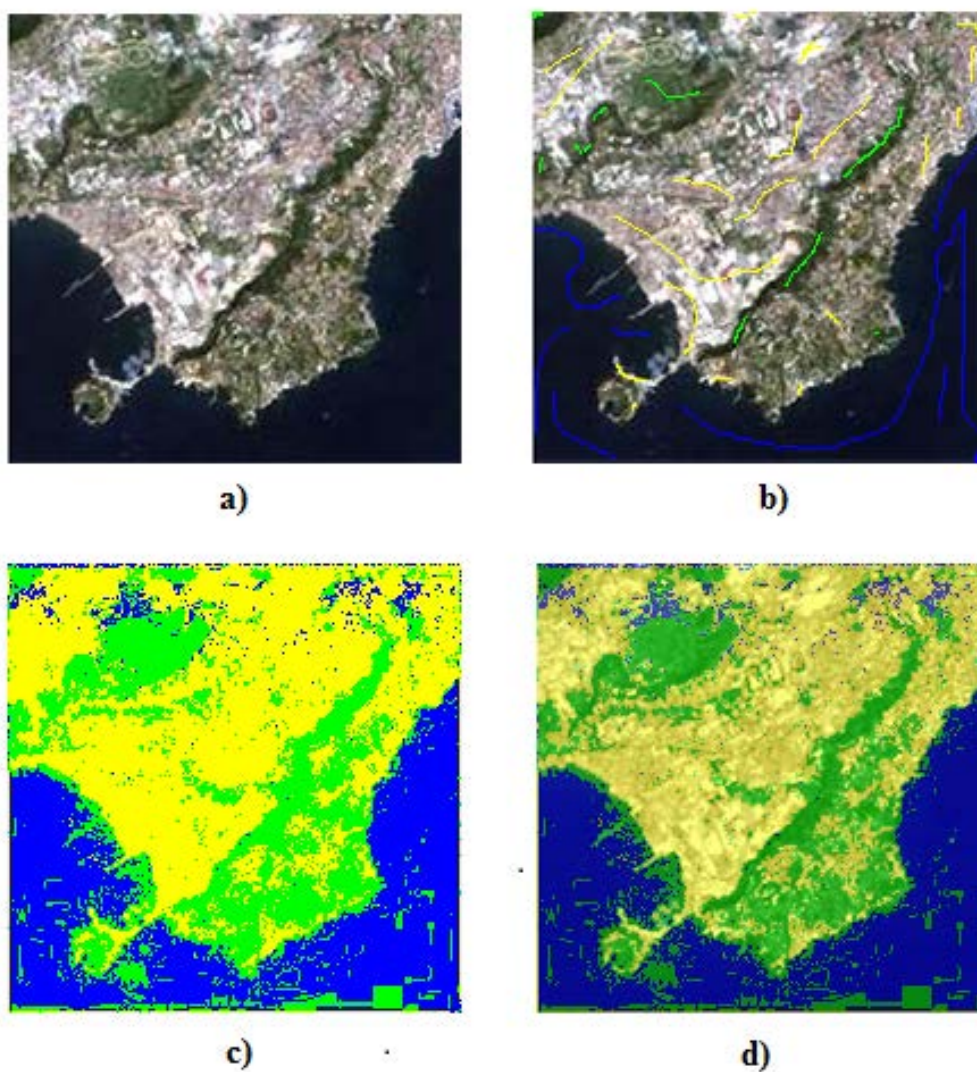
Fonte imagem a) Irony et al. (2005)

Figura 5.13: Imagem de saída para imagem de entrada da Figura 5.12a) usando o modelo proposto em Irony et al. (2005)



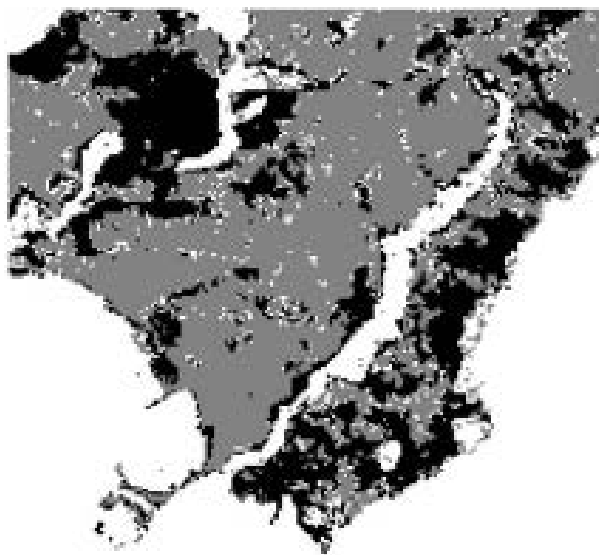
Fonte: Irony et al. (2005)

Figura 5.14: Imagem comparativa Gómez-Chova et al. (2008): a) Imagem de entrada; b) Imagem de entrada com as marcações de partículas; c) Classificação da imagem usando o algoritmo de competição e cooperação entre partículas; d) Comparação entre a imagem de entrada e resposta de classificação;



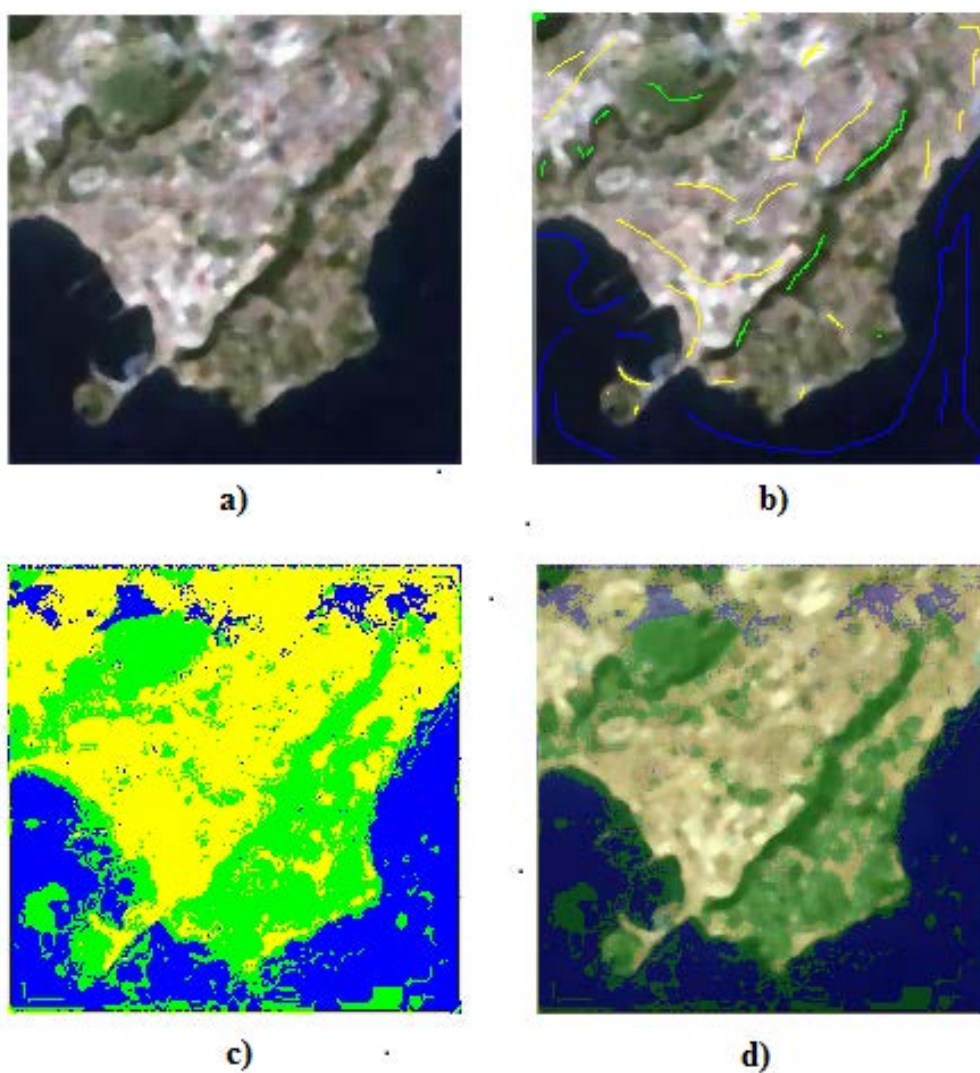
Fonte imagem a) Gómez-Chova et al. (2008)

Figura 5.15: Imagem de saída para imagem de entrada da Figura 5.14a) usando o modelo proposto em Gómez-Chova et al. (2008)



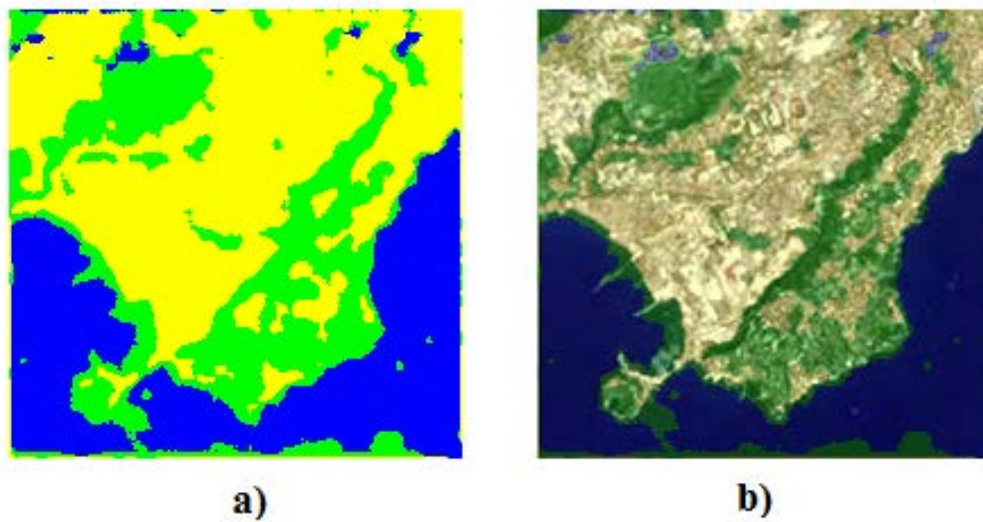
Fonte: Gómez-Chova et al. (2008)

Figura 5.16: Imagem comparativa Gómez-Chova et al. (2008) utilizando Filtro mediana aplicado na imagem de entrada: a) Imagem de entrada utilizando o Filtro Mediana; b) Imagem de entrada com as marcações de partículas; c) Classificação da imagem usando o algoritmo de competição e cooperação entre partículas; d) Comparação entre a imagem de entrada e resposta de classificação;



Fonte imagem a) Gómez-Chova et al. (2008)

Figura 5.17: Imagem comparativa Gómez-Chova et al. (2008) utilizando Filtro mediana aplicado na imagem de saída: a) Imagem de saída da Figura 5.14 processada pelo Filtro Mediana b) Sobreposição entre a) e 5.14 a)



Fonte imagem a) Gómez-Chova et al. (2008)

Capítulo 6

Conclusão

A abordagem de competição e cooperação entre partículas, apesar de recente, já foi utilizada em diversas tarefas como classificação de dados, detecção de comunidades sobrepostas, aprendizado com dados imperfeitos, aprendizado ativo, aprendizado com mudança de conceitos, etc. Porém sua aplicação sempre foi feita em dados baseados em vetores de atributos, nunca em imagens. A extensão do modelo para aplicação na tarefa de segmentação de imagens apesar de não ter sido trivial, demonstrou ser bastante promissora e apresentou uma boa resposta.

A segmentação autônoma representa uma das tarefas mais complexas quando trabalhamos com processamento de imagens. Para simplificar esta tarefa utilizamos nesta proposta o modelo de competição e cooperação entre partículas, que utiliza aprendizado semi-supervisionado. Desta forma, é possível fazer uso de um especialista para o auxílio na segmentação de imagens, rotulando manualmente alguns pixels de cada objeto que se deseja destacar. O algoritmo se encarrega de completar a tarefa, usando para isso ambos os pixels rotulados e não rotulados como fonte de aprendizado. Deste modo, partículas são criadas através dos pixels da imagem identificados pelo usuário e, assim tais rótulos podem se espalhar para os demais pixels.

Ainda que alcançado o objetivo inicial do projeto, foi necessário lidar com alguns problemas, como o custo computacional para o cálculo e comparação de distância Euclidiana, pixel a pixel. Devido ao aumento de resolução e detalhes de imagens como as de satélite, foi requerido um aumento na quantidade de pixels, principalmente para a percepção de detalhes. A solução encontrada para este problema foi a alteração da forma de armazenamento dos dados para o cálculo da Distância Euclidiana. Atualmente são armazenados apenas os dados que possuem maior probabilidade de pertencer a lista de vizinhos mais próximos referentes ao nó analisado, descartando os demais. Dessa forma, utiliza-se menos memória principal, e assim permitir o aumento de tamanho das imagens, sem que seja necessário aumentar drasticamente a quantidade de memória principal utilizada.

No entanto, apesar do controle do uso de memória computacional, existe ainda um obstáculo quanto ao tempo de execução. A comparação ocorrida pixel a pixel, causa uma elevação no tempo, representando um aumento exponencial em relação ao tamanho da imagem. Para contornar tal resistência foi implementada paralelização na etapa de montagem do grafo. Porém, apesar de apresentar uma boa melhoria, a execução ainda é um pouco demorada, pois as comparações e inserções no grafo de pixels são realizadas após o cálculo da distância Euclidiana ser feito em toda a imagem.

Devido ao fato de que o tempo de execução depende de diversos fatores, inclusive o desempenho do processador diante dos cálculos matemáticos exigidos pelo modelo de competição e cooperação entre

partículas. Com o intuito de contornar tal problema foi manipulado o algoritmo de forma a equilibrar a melhor classificação em menor tempo, para isto foram eliminados alguns atributos presentes na imagem. A redução de atributos ocorreu com a retirada do desvio padrão, calculado entre os valores de pixels em um determinado raio de pixels, e média destes valores. Na execução do modelo de competição e cooperação entre partículas, tais valores são utilizados para a redução do ruído, garantindo a presença de mais detalhes da imagem classificada com menores erros, no entanto esta melhora de classificação tem como desvantagem um maior tempo de execução.

Tais problemas referentes a tempo de execução e custo computacional também estão ligados à linguagem escolhida, no caso a linguagem Java. Criada para que os códigos sejam executados em uma máquina virtual, visando portabilidade, tem como desvantagem tornar a execução mais lenta em comparação com código compilado para executar diretamente no processador, o que os torna mais pesados e lentos. Apesar das máquinas virtuais de execução terem evoluído bastante, ainda é necessário um tempo para que o código seja carregado e otimizado (Real, 2015). Tal linguagem é vastamente utilizada devido a manutenção e compatibilidade com versões anteriores, além de possuir inúmeras bibliotecas, geralmente bem otimizadas, as quais auxiliam na programação, tornando-a mais simples e amigável. No entanto apresenta uma compilação mais lenta quando comparada a outros softwares matemáticos ao se utilizar equações matemáticas.

As principais vantagens observadas no desenvolvimento do projeto são:

- 1 : Apresenta uma boa classificação de objetos e regiões com o uso do modelo de competição e cooperação entre partículas, principalmente referente as bordas das regiões.
- 2 : Este método pode ser utilizado em aplicações direcionadas, devido ao fato de que é possível alterar os valores de importância de cada atributo para o cálculo da distância Euclidiana, permitindo uma melhor classificação diante de cada tipo de imagem.
- 3 : Apresenta uma boa classificação sem que sejam necessárias muitas intervenções de especialistas, precisando de poucos dados rotulados.
- 4 : O modelo de competição e cooperação entre partículas permite sua implementação em diferentes tipos de linguagens, se adaptando melhor às necessidades de cada aplicação.

Os estudos apresentados nesta dissertação possibilitam extensões futuras como alterar o modelo de competição e cooperação entre partículas para que este receba qualquer tipo de imagem, analisando as necessidade e regulando os valores dos atributos da distância Euclidiana em tempo de execução, representando um bom resultado independente das características de cada imagem de entrada, sem a necessidade de alteração manual dos parâmetros para cada entrada, de forma que o próprio algoritmo reconheça os melhores parâmetros a serem realçados. Uma possível alteração para este modelo é a inserção de uma interface para a manipulação e marcação de partículas na imagem em tempo de execução, com o objetivo de permitir que o usuário possa observar a imagem sendo classificada em tempo de execução e, eventualmente, classificar manualmente novos pixels com o intuito de ajudar o algoritmo a classificar corretamente regiões que de outra forma seriam classificadas incorretamente.

Referências Bibliográficas

- Abney, S. (2008). *Semisupervised Learning for Computational Linguistics*. CRC Press.
- Aggarwal, C. and P. Yu (2009). A survey of uncertain data algorithms and applications. *IEEE Transactions on Knowledge and Data Engineering* 21(5), 609–623.
- Alpaydin, E. (2004). *Introduction to machine learning*. MIT Press.
- Barber, D. (2012). *Bayesian Reasoning and Machine Learning*. Cambridge University Press.
- Bhosale, P., A. Gokhale, and Y. Motey (2013, April). Image segmentation using graph cut technique for outdoor scene images. In *Communications and Signal Processing (ICCSP), 2013 International Conference on*, pp. 280–282.
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.
- Blake, A., C. Rother, M. Brown, P. Perez, and P. Torr (2004). Interactive image segmentation using an adaptive gmmrf model. In *Computer Vision-ECCV 2004*, pp. 428–441. Springer.
- Blum, A. and T. Mitchell (1998). Combining labeled and unlabeled data with co-training. In *COLT: Proceedings of the Workshop on Computational Learning Theory*, pp. 92–100.
- Breve, F. (2013). Active semi-supervised learning using particle competition and cooperation in networks. In *Neural Networks (IJCNN), The 2012 International Joint Conference on*.
- Breve, F. and L. Zhao (2012a). Particle competition and cooperation in networks for semi-supervised learning with concept drift. In *Neural Networks (IJCNN), The 2012 International Joint Conference on*, pp. 1–6.
- Breve, F. and L. Zhao (2012b). Particle competition and cooperation to prevent error propagation from mislabeled data in semi-supervised learning. In *Neural Networks (SBRN), 2012 Brazilian Symposium on*, pp. 79–84.
- Breve, F. and L. Zhao (2013). Fuzzy community structure detection by particle competition and cooperation. *Soft Computing* 17(4), 659–673.
- Breve, F., L. Zhao, M. Quiles, W. Pedrycz, and J. Liu (2012). Particle competition and cooperation in networks for semi-supervised learning. *Knowledge and Data Engineering, IEEE Transactions on* 24(9), 1686–1698.
- Breve, F. A. (2010). Aprendizado de maquina utilizando dinamica espaco temporal em redes complexas.

- Breve, F. A. and L. Zhao (2011). Preventing error propagation in semi-supervised learning using teams of walking particles. In *Anais do X Congresso Brasileiro de Inteligência Computacional (CBIC 2011)*.
- Breve, F. A., L. Zhao, and M. G. Quiles (2009). Particle competition in complex networks for semi-supervised classification. In J. Zhou (Ed.), *Complex (1)*, Volume 4 of *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, pp. 163–174. Springer.
- Breve, F. A., L. Zhao, and M. G. Quiles (2015). Particle competition and cooperation for semi-supervised learning with label noise. *Neurocomputing 160*, 63–72.
- Breve, F. A., L. Zhao, M. G. Quiles, W. Pedrycz, and J. Liu (2011). Particle competition and cooperation for uncovering network overlap community structure. In D. Liu, H. Zhang, M. Polycarpou, C. Alippi, and H. He (Eds.), *Advances in Neural Networks - ISNN 2011*, Volume 6677 of *Lecture Notes in Computer Science*, pp. 426–433. Springer Berlin / Heidelberg. 10.1007/978-3-642-21111-9_48.
- Camps-Valls, G., T. V. Bandos, and D. Zhou (2007). Semi-supervised graph-based hyperspectral image classification. *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING XX*, 1–29.
- Chapelle, O., B. Schölkopf, and A. Zien (Eds.) (2006). *Semi-Supervised Learning*. Adaptive Computation and Machine Learning. Cambridge, MA: The MIT Press.
- Chapelle, O., B. Schölkopf, A. Zien, et al. (2006). Semi-supervised learning.
- Cios, K. J., W. Pedrycz, R. W. Swiniarski, and L. A. Kurgan (2007). *Data Mining: A Knowledge Discovery Approach*. Springer.
- Dara, R., S. Kremer, and D. Stacey (2002). Clustering unlabeled data with soms improves classification of labeled real-world data. In *Proceedings of the World Congress on Computational Intelligence (WCCI)*, pp. 2237–2242.
- Dass, R., Priyanka, and S. Devi (2012). Image segmentation techniques. *International Journal of electronics and communication technology 3*, 66–70.
- Demiriz, A., K. P. Bennett, and M. J. Embrechts (1999). Semi-supervised clustering using genetic algorithms. In *Proceedings of Artificial Neural Networks in Engineering (ANNIE-99)*, pp. 809–814. ASME Press.
- Duda, R. O., P. E. Hart, and D. G. Stork (2000). *Pattern Classification (2nd Edition)*. Wiley-Interscience.
- Felzenszwalb, P. F. and D. P. Huttenloche (2004, September). Efficient graph-based image segmentation. *International Journal of Computer Vision*, 167–181.
- Fujino, A., N. Ueda, and K. Saito (2005). A hybrid generative/discriminative approach to semi-supervised classifier design. In *AAAI-05, Proceedings of the Twentieth National Conference on Artificial Intelligence*, pp. 764–769.
- Garcia-Ugarriza, L., E. Saber, V. Amuso, M. Shaw, and R. Bhaskar (2008, March). Automatic color image segmentation by dynamic region growth and multimodal merging of color and texture information. In *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, pp. 961–964.

- Glasbey, C. and G. Horgan (1994). Image analysis for the biological sciences, chapter 4: Segmentation. pp. 1–31.
- Gómez-Chova, L., G. Camps-Valls, J. Munoz-Mari, and J. Calpe (2008). Semisupervised image classification with laplacian support vector machines. *Geoscience and Remote Sensing Letters, IEEE* 5(3), 336–340.
- Gonzalez, R. C. and R. E. Woods (2006). *Digital Image Processing* (3rd ed.). Upper Saddle River, NJ, USA: Prentice-Hall, Inc.
- Gonzalez, R. C. and R. E. Woods (2010). *Processamento Digital de Imagens*. São Paulo: Pearson Prentice Hall.
- Grady, L. (2006). Random walks for image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 28(11), 1768–1783.
- Harchaoui, Z. and F. Bach (2007). Image classification with segmentation graph kernels. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pp. 1–8. IEEE.
- Hinton, G. E. and T. J. Sejnowski (1999). *Unsupervised Learning: Foundations of Neural Computation*. MIT Press.
- Hu, C., B. Youn, and T. Kim (2012). Semi-supervised learning with co-training for data-driven prognostics. In *Prognostics and Health Management (PHM), 2012 IEEE Conference on*, pp. 1–10.
- Irony, R., D. Cohen-Or, and D. Lischinski (2005). Colorization by example. In *Eurographics Symp. on Rendering*, Volume 2. Citeseer.
- Jain, A. K. and R. C. Dubes (1988). *Algorithms for clustering data*. Publisher by Printice Hall.
- Jain, A. K., M. N. Murty, and P. J. Flynn (1999). Data clustering: a review. *ACM Computing Surveys* 31(3), 264–323.
- Kaufman, L. and P. J. Rousseeuw (1990). *Finding groups in data: An introduction to cluster analysis*. John Wiley and Sons.
- Liu, B. (2007). *Web data mining: exploring hyperlinks, contents, and usage data*. Springer Science & Business Media.
- Liu, B.-B. and Z.-M. Lu (2008). Image colourisation using graph-based semi-supervised learning. *IEEE-EXplore* 3, 115 – 120.
- Lu, T. T. (2009). Fundamental limitations of semi-supervised learning.
- Ma, W.-Y. and B. Manjunath (2000). Edgeflow: a technique for boundary detection and image segmentation. *Image Processing, IEEE Transactions on* 9(8), 1375–1388.
- Malmberg, F. (2011). *Graph-based Methods for Interactive Image Segmentation*. Ph. D. thesis, Uppsala University Uppsala University, Centre for Image Analysis, Computerized Image Analysis and Human-Computer Interaction.

- Mitchell, T. (1997). *Machine Learning*. McGraw Hill.
- Mitchell, T. M. (1999). The role of unlabeled data in supervised learning. In *Proceedings of the Sixth International Colloquium on Cognitive Science*.
- Natarajan, B. K. (2014). *Machine learning: a theoretical approach*. Morgan Kaufmann.
- Navon, E., O. Miller, and A. Averbuch (2005). Color image segmentation based on adaptive local thresholds. *Image and vision computing* 23(1), 69–85.
- Nigam, K., A. K. McCallum, S. Thrun, and T. Mitchell (2000). Text classification from labeled and unlabeled documents using em. In *Machine Learning*, Volume 39, pp. 103–134.
- Pal, N. R. and S. K. Pal (1993). A review on image segmentation techniques. *Pattern Recognition* 26(9), 1277 – 1294.
- Pantofaru, C. R. (2008). Studies in using image segmentation to improve object recognition.
- Qi, X. (2005). Image segmentation.
- Quan, L., P. Tan, G. Zeng, L. Yuan, J. Wang, and S. B. Kang (2006). Image-based plant modeling. In *ACM Transactions on Graphics (TOG)*, Volume 25, pp. 599–604. ACM.
- Quiles, M. G., L. Zhao, R. L. Alonso, and R. A. F. Romero (2008). Particle competition for complex network community detection. *Chaos* 18(3), 033107.
- Real, L. F. O. C. (2015). Linguagens de programação - java.
- Saldanha, M. F. and C. FREITAS (2009). Segmentação de imagens digitais: Uma revisão. *Divisão de Processamento de Imagens-Instituto Nacional de Pesquisas Espaciais (INPE), São Paulo*.
- Salman, N. H. and L. Chong-Qing (2001). Edge detection and image segmentation based on k-means and watershed techniques. In *Institute of Image Processing e Pattern Recognition Shanghai Jiao Tong University*, pp. 148–153.
- Schaeffer, S. E. (2007). Graph clustering. *Computer Science Review* 1(1), 27–64.
- Shapiro, L. G. and G. C. Stockman (2001). *Computer Vision*. New Jersey: Prentice-Hall.
- Shi, J. and J. Malik (2000, August). Normalized cuts and image segmentation. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE* 22, 888–905.
- Skarbek, W. and A. Koschan (1994). Colour image segmentation a survey.
- Szummer, M. and T. Jaakkola (2002). Partially labeled classification with markov random walks. In *Advances in Neural Information Processing Systems*, Volume 14.
- USGS (2015). Science for a changing world.
- Vapnik, V. N. (1998). *Statistical Learning Theory*. New York: Wiley-Interscience.

- Wang, F., T. Li, G. Wang, and C. Zhang (2008). Semi-supervised classification using local and global regularization. In *AAAI'08: Proceedings of the 23rd national conference on Artificial intelligence*, pp. 726–731. AAAI Press.
- Wang, F. and C. Zhang (2008). Label propagation through linear neighborhoods. *IEEE Transactions on Knowledge and Data Engineering* 20(1), 55–67.
- Wang, F., C. Zhang, H. C. Shen, and J. Wang (2006). Semi-supervised classification using linear neighborhood propagation. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on I*, 160–167.
- Witten, I. H. and E. Frank (2005). *Data Mining: Practical Machine Learning Tools and Techniques* (2 ed.). Morgan Kaufmann.
- Wolff, R., K. Bhaduri, and H. Kargupta (2009). A generic local algorithm for mining data streams in large distributed systems. *IEEE Transactions on Knowledge and Data Engineering* 21(4), 465–478.
- Wu, M. and B. Schölkopf (2007). Transductive classification via local learning regularization. In X. S. Meila, M. (Ed.), *11th International Conference on Artificial Intelligence and Statistics*, Brookline, MA, USA, pp. 628–635. Microtome.
- Zaart, A. E., D. Ziou, S. Wang, and Q. Jiang (2002). Segmentation of sar images. *Pattern Recognition* 35(3), 713–724.
- Zhou, D., O. Bousquet, T. N. Lal, J. Weston, and B. Schölkopf (2004). Learning with local and global consistency. In *Advances in Neural Information Processing Systems*, Volume 16, pp. 321–328. MIT Press.
- Zhou, Z.-H. and M. Li (2005). Tri-training: exploiting unlabeled data using three classifiers. *IEEE Transactions on Knowledge and Data Engineering* 17(11), 1529–1541.
- Zhou, Z.-H. and M. Li (2007). Semisupervised regression with cotraining-style algorithms. *IEEE Transactions on Knowledge and Data Engineering* 19(11), 1479–1493.
- Zhu, X. (2005). Semi-supervised learning literature survey. Technical Report 1530, Computer Sciences, University of Wisconsin-Madison.
- Zhu, X. (2013). Semi-supervised learning. *University of Wisconsin Madison*.
- Zhu, X. and Z. Ghahramani (2002). Learning from labeled and unlabeled data with label propagation. Technical Report CMU-CALD-02-107, Carnegie Mellon University, Pittsburgh.
- Zhu, X., Z. Ghahramani, and J. Lafferty (2003). Semi-supervised learning using gaussian fields and harmonic functions. In *Proceedings of the Twentieth International Conference on Machine Learning*, pp. 912–919.