

Mauro Medeiros Barbat

Classificação de sedimentos superficiais de leitos
oceânicos e de águas continentais: Uma
abordagem baseada em super segmentação e
sabedoria das massas



PETROBRAS



Rio Grande, RS
2016

Mauro Medeiros Barbat

Classificação de sedimentos superficiais de leitos
oceânicos e de águas continentais: Uma
abordagem baseada em super segmentação e
sabedoria das massas

Dissertação de Mestrado apresentada ao
Programa de Pós-Graduação em Com-
putação da Universidade Federal do Rio
Grande, como requisito parcial para a ob-
tenção do grau de Mestre em Engenharia
de Computação.

Orientador: Prof. Dr. Adriano Velasque
Werhli

Coorientador: Prof. Dr. Lauro Antônio
Saint Pastous Madureira

Rio Grande, RS
2016

Barbat, M. M. Classificação de sedimentos superficiais de leitos oceânicos e de águas continentais: Uma abordagem baseada em super segmentação e sabedoria das massas [115](#) páginas Dissertação (Mestrado) - Universidade Federal do Rio Grande. Centro de Ciências Computacionais, C3.

1. Aprendizado de Máquina
2. Classificação de sedimentos
3. Visão computacional
4. Sonar de varredura lateral
5. Sabedoria das Massas

I. Universidade Federal do Rio Grande, FURG. Centro de Ciências Computacionais, C3.

UNIVERSIDADE FEDERAL DO RIO GRANDE
Centro de Ciências Computacionais
Programa da Pós-Graduação em Computação
Curso de Mestrado em Engenharia de Computação

DISSERTAÇÃO DE MESTRADO

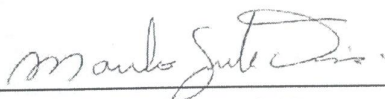
**Classificação de sedimentos superficiais de leitos oceânicos e de águas continentais:
Uma abordagem baseada em super segmentação e sabedoria das massas.**

Mauro Medeiros Barbat

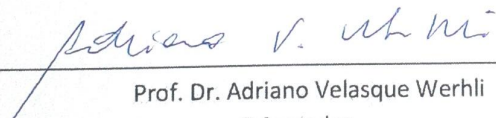
Banca examinadora:



Prof. Dr. Alessandro de Lima Bicho



Prof. Dr. Marcelo Sperle Dias



Prof. Dr. Adriano Velasque Werhli
Orientador



Prof. Dr. Lauro Antônio Saint Pastous Madureira
Coorientador

*Dedico este trabalho a minha mãe Ana Paula e ao meu avô Clodomiro os
quais com seu apoio e dedicação me ensinaram que os únicos obstáculos
que podem nós parar são aqueles impostos por nós mesmos.*

"A persistência é o menor caminho
do êxito"

Charles Chaplin

Agradecimentos

Agradeço ao meu orientador, ao meu co-orientador, aos meus colaboradores, à fundação que liberou verba para minhas pesquisas, aos meus amigos e à minha família.

Resumo

Observando o problema de classificar o substrato marinho, existem diversificadas técnicas que em grande parte consistem em metodologias tradicionais que demandam a necessidade de especialistas e/ou aquisição física de amostras para efetuar a classificação em laboratório, consumindo grande quantidade de recursos humanos e financeiros. No entanto, técnicas para classificação remota são capazes de proporcionar não apenas ganhos em relação a tempo e recursos mas também, são capazes de operar sobre grandes volumes de dados de maneira eficiente e eficaz, processo o qual poderia requerer um tempo considerável de trabalho. Neste intuito, este trabalho propõe o desenvolvimento de uma metodologia/ferramenta computacional focada no contexto de classificação de sedimentos superficiais de leitos oceânicos e de águas continentais, utilizando sonografias de sonares de varredura lateral (*Side Scan Sonar*) que consistem em imagens digitais de alta resolução.

Palavras-chave: Aprendizado de máquina, classificação de sedimentos, visão computacional, sonar de varredura lateral, sabedoria das massas.

Abstract

Observing the problem of marine bottom classification, there are various techniques that largely consist of traditional methods that require the knowledge of experts and / or ground proof of samples to their classification in the laboratory, consuming large amounts of human and financial resources. However, techniques for remote classification are not only able to provide gains of time and resources, but also are able to operate on large volumes of data efficiently and effectively process which would require a considerable amount of work. Therefore, so far this dissertation proposes the development of a computational methodology/tool focused on the context of surface sediment classification of ocean floors and continental waters, using as a basis the use of side scan sonar that consist of high-resolution digital images.

Keyword: Machine Learning, Sediments classification, computer vision, side scan sonar, wisdom of crowds.

Lista de Figuras

2.1	Exemplificação campos relacionados a análise de imagens digitais	11
2.2	Etapas gerais reconhecimento de padrões	14
2.3	Sistema de coordenadas para representação de imagens digitais.	16
2.4	Representação matricial.	17
2.5	Exemplificação convolução de filtro média	19
2.6	Exemplificação resultado da aplicação de filtro média	19
2.7	Exemplificação sequência completa filtragem no domínio da frequência	21
2.8	Exemplificação imagem representada no domínio da frequência	21
2.9	Exemplificação histograma de imagens em escala de cinza	22
2.10	Exemplificação alteração de brilho imagens	23
2.11	Exemplificação alteração de contraste em imagens	23
2.12	Segmentação de Imagens	26
2.13	Exemplificação imagem representada em três níveis de cinza na forma matricial e sua respectiva GLCM na direção 0	29
2.14	Ilustração texturas naturais e artificiais	30
2.15	Ilustração PCA	31
2.16	Exemplificação esquema para se obter a correlação de $f(x,y)$ e $w(x,y)$ no ponto (s,t)	35
2.17	Exemplificação blocos funcionais para reconhecimento de padrões na abordagem estatística, baseado em (Jain et al., 2000)	36

2.18	Ilustração efeito da rugosidade do fundo no retroespalhamento	39
2.19	Exemplo rugosidade relacionada as variadas ondulações resultantes por fenômenos hidrodinâmicos	40
2.20	Exemplo decomposição da rugosidade de fundo em suas diversas componentes espectrais	40
2.21	Exemplo Sonar de Varredura Lateral	41
2.22	Ilustração do funcionamento do sonar de varredura lateral - SVL	42
2.23	Exemplificação do imageamento acústico por sidescan sonar	44
2.24	Imagem de sonar de varredura lateral	45
2.25	Imagem de sonar de varredura lateral 2	46
2.26	Ilustração de razões para usar sistemas baseados em sabedoria das massas	48
2.27	Demonstração aplicação de sabedoria das massas em sistemas de classificação.	51
3.1	Sistema de aquisição de dados do sonar de varredura lateral.	60
3.2	Etapas processo de classificação supervisionada de objetos em imagens.	62
3.3	Sequência lógica de execução da metodologia de classificação proposta.	63
3.4	Amostras sedimentológicas coletadas no parcel do carpinteiro, visando relacionar com as sonografias adquiridas	64
3.5	Relação amostras <i>in situ</i> com o imageamento acústico	65
3.6	Ilustração processo de seleção conjunto de amostras para treino	65
3.7	Sobreposição superficial dos sedimentos que resulta em ecos de retorno com intensidade oscilante na linha de aquisição que constituem cada linha da imagem	67
3.8	Demonstração sonografias de alta e baixa frequência, 500KHz e 100KHz respectivamente	67
3.9	Aplicação filtro média	68

3.10	Máscara 5×5 da função Gaussiana aproximada com $\sigma = 1.0$	69
3.11	Aplicação filtro mediana	70
3.12	Ilustração tipos de ruídos mais comuns encontrados nestas imagens	71
3.13	Sonografia em escala de cinza antes e após a etapa de pré-processamento	71
3.14	Presença de <i>ripples</i> ocasionados por fenômenos hidrodinâmicos que alteram a textura da região.	72
3.15	Sonografias evidenciando mesma classe sedimentológica com variabilidade textural	73
3.16	Sonografias (2) evidenciando mesma classe sedimentológica com variabilidade textural	75
3.17	Seleção e extração de descritores	76
3.18	Ilustração comparação entre classificadores. Adaptado: Scikitlearn	80
3.19	Metodologia de treinamento proposta	81
3.20	Ilustração de segmentação por convolução	84
3.21	Ilustração convolução sobre regiões limítrofes entre classes	84
3.22	Demonstração da segmentação utilizando superpixels	85
3.23	Ilustração do processo de classificação proposto.	87
3.24	Ilustração das etapas de paralelismo durante a linha de execução do processo de classificação.	88
4.1	Demonstração do modulo de pré-processamento	90
4.2	Demonstração do modulo de treinamento	91
4.3	Demonstração do modulo de classificação	92
4.4	Comparativo de precisão entre classificadores com a técnica de Superpixels.	94
4.5	Comparativo de precisão entre classificadores com a técnica de <i>sliding window</i>	94

4.6	Comparativo entre os resultados dos classificadores independentes com um ponto de sabedoria das massas e da composição destes com a aplicação do segundo ponto, utilizando segmentação por <i>sliding window</i>	95
4.7	Comparativo entre os resultados dos classificadores independentes com um ponto de sabedoria das massas e da composição destes com a aplicação do segundo ponto, com segmentação por <i>superpixel</i>	96
4.8	Ilustração processo de composição resultado com dois pontos de sabedoria das massas.	96
4.9	Comparação entre o resultado da classificação humana e da metodologia apresentada.	98
4.10	Comparação entre o resultado da classificação humana e da metodologia apresentada.	99
4.11	Comparativo resultados qualitativos entre <i>superpixel</i> e <i>sliding window</i> . . .	101
4.12	Comparativo resultados qualitativos entre <i>superpixel</i> e <i>sliding window</i> . . .	102

Lista de Tabelas

2.1	Comparação entre os autores e as técnicas utilizadas para classificação de imagens de SSS.	55
-----	--	----

Sumário

1	Introdução	1
1.1	Justificativa	2
1.2	Motivação	3
1.3	Objetivo geral	5
1.4	Objetivos específicos	5
1.5	Desafios	6
2	Metodologia e trabalhos relacionados	9
2.1	Introdução	9
2.2	Análise de Imagens Digitais	10
2.2.1	Introdução	10
2.2.2	Metodologia	10
2.2.3	Aplicações	36
2.3	Acústica Submarina	37
2.3.1	Introdução	37
2.3.2	Metodologia	38
2.3.3	Aplicações	45
2.4	Sabedoria das Massas	46
2.4.1	Introdução	46
2.4.2	Metodologia	47

2.4.3	Aplicações	50
2.5	Trabalhos Relacionados	50
2.5.1	Introdução	50
2.5.2	Trabalhos publicados	52
3	Framework Proposto	57
3.1	Introdução	57
3.2	Engenharia de Software	57
3.2.1	Engenharia de Requisitos	58
3.2.2	Engenharia de Sistemas	59
3.2.3	Engenharia de Projeto	61
3.3	Seleção de Amostras Treinamento	61
3.4	Pré-Processamento	66
3.5	Extração de Características	72
3.6	Modificação de Características	76
3.7	Treinamento	77
3.8	Segmentação x Convolução	81
3.9	Classificação	83
4	Resultados	89
4.1	Introdução	89
4.2	O <i>Framework</i>	89
4.3	Resultados Quantitativos	91
4.4	Resultados Qualitativos	97
5	Discussão	103
5.1	Os resultados	103
6	Conclusões	107

Capítulo 1

Introdução

A necessidade de explorar recursos presentes sobre e sob o leito oceânico é uma realidade cada vez mais urgente com o passar dos anos, pois grande parte dos recursos como os combustíveis fósseis estão tendo suas reservas terrestres levadas a exaustão. Porém, a exploração em ambientes marinhos não é uma atividade simples, considerando que a maior parte do planeta é coberto por oceanos e mares e que grande parte destes recursos encontram-se em regiões de difícil acesso.

Neste contexto, nas últimas décadas um grande volume de trabalhos científicos foi produzido ao redor do globo focando suas atenções para os ambientes subaquáticos com diferentes propósitos. Neste trabalho é apresentada uma metodologia computacional capaz de classificar e mapear remotamente sedimentos presentes nas camadas superficiais do substrato marinho e de águas continentais valendo-se por base do uso de sonografias (imagens de alta resolução geradas através do eco-acústico) de sonar de varredura lateral, visando conciliar a eficiência dos métodos hidroacústicos para aquisição de dados e mapeamento em ambientes subaquáticos com a eficiência das técnicas de visão computacional, reconhecimento de padrões e aprendizado de máquina para diferenciação, caracterização e classificação de estruturas artificiais e naturais em imagens digitais.

Através do desenvolvimento de um *framework* computacional e uma metodologia para classificação supervisionada centrada na aplicação de conceitos de sabedoria das massas e super segmentação, de forma a compor uma solução global aprimorada a partir da união de soluções de classificadores especialistas locais.

1.1 Justificativa

A utilização de sonares de varredura lateral (*Side Scan Sonar*) para imageamento de alta resolução de leitos oceânicos e de águas continentais é amplamente explorada em diversas atividades, tais como: detecção de impacto de atividades humanas, identificação e inspeção de estruturas artificiais, estudos morfológicos, diferenciação e classificação dos sedimentos superficiais que compõem o substrato de determinada região, entre outras, ligadas a setores de defesa ambiental, engenharia oceânica e a indústria do petróleo, pois possibilitam o mapeamento de grandes áreas de maneira eficiente.

Metodologias tradicionais de classificação de sedimentos necessitam coletas *in situ* de amostras do substrato e demandam tempo e recursos complexos em equipamentos, embarcação e equipe especializada. Por outro lado, a elaboração de mecanismos para classificação remota, valendo-se de metodologias acústicas, torna este processo mais eficiente, em termos de tempo e custo de operação pois, possibilita a aquisição de dados com a embarcação em movimento (Oliveira Junior, 2012). Observando a relevância destas metodologias diretamente relacionadas a aspectos não apenas científicos e educacionais, mas econômicos, estratégicos e de gerenciamento, considerando o aumento das interações humanas sobre o substrato marinho, decorrente principalmente de procedimentos de exploração de recursos minerais e fósseis, nota-se a importância de ferramentas para monitoramento e inspeção destes ambientes.

Segundo Tęgowski (2005) e Anderson et al. (2008), os métodos acústicos são ferramentas eficientes, em termos de tempo e custo de operação, para mapear e monitorar

grandes áreas oceânicas quando comparados com ferramentas tradicionais, minimizando a necessidade de um grande número de amostragens físicas a partir de amostradores específicos. O retroespalhamento acústico de fundo que consiste, basicamente, em uma relação entre o sinal acústico incidente sobre o fundo e que retorna para o equipamento, pode ser utilizado para a classificação de sedimentos superficiais e, em alguns casos, o acompanhamento de comunidades biológicas.

Com avanço de tecnologias hidroacústicas como ecobatímetros e sonares de varredura lateral, diferentes métodos estatísticos, matemáticos e computacionais têm sido empregados para a classificação dos sedimentos superficiais que compõem o leito oceânico a partir dos dados fornecidos por estes equipamentos.

1.2 Motivação

Conforme definido no dicionário ([Michaelis, 2005](#)), engenharia é a arte de aplicar métodos científicos ou empíricos à utilização dos recursos da natureza em benefício do ser humano. No contexto de engenharia de computação, o pensamento computacional, assim como inúmeras técnicas computacionais, podem ser aplicadas visando contribuir em benefício do ser humano, proporcionando mecanismos mais eficazes em aplicações distintas como em problemas de tomada de decisão e classificação de objetos presentes em imagens.

Atualmente, existe uma vasta lista de aplicações como, por exemplo, a área médica, onde estas técnicas são utilizadas para a detecção de anomalias a partir de exames médicos específicos como radiografias, eletroencefalogramas, eletrocardiogramas, entre outros, pois proporcionam mecanismos eficientes no auxílio da detecção de estruturas anômalas que poderiam passar despercebidas em inspeções visuais, mesmo por profissionais treinados.

Visando a exploração e o estudo de recursos naturais marinhos, ferramentas que

apoiem o processo de classificação de sedimentos superficiais são importantes, pois possibilitam observar não somente classes de sedimentos, mas aspectos morfológicos que indiquem impactos provenientes de atividades humanas. Por exemplo, alterações no substrato geradas pela atividade pesqueira com redes de arrasto de fundo ou mesmo a detecção de resíduos fixados ao substrato, tais como vazamentos de petróleo.

De forma mais específica pode-se citar o processo de exploração do petróleo, no qual diversos estudos complementares são feitos de forma a determinar as características geológicas e geofísicas da área de exploração. Neste sentido, a união de métodos que possibilitem reunir informações sobre a geologia da superfície (Seabed Surveys), identificando fatores topográficos, físicos e sedimentares que podem proferir impactos nas atividades de exploração e de monitoramento tornam-se fundamentais.

Com o objetivo de complementar estes estudos, os métodos para diferenciação sedimentológica com base em dados acústicos são amplamente difundidos e utilizados, pela sua capacidade de fornecer informações pertinentes sobre a camada superficial do leito marinho, como granulometria e rugosidade. Além disso, este métodos possibilitam aos especialistas classificar e reconhecer em tempo real estruturas naturais e não naturais que possam ser de interesse.

No entanto, grande parte dos dados gerados por estes equipamentos, como no caso das sonografias resultantes por sonares de varredura lateral, são normalmente analisadas e interpretadas por operadores, demandando recursos humanos e de tempo para esta tarefa e, por consequência, sujeitando os resultados a possíveis falhas oriundas do desgaste humano intrínseco nesta abordagem. Neste contexto, a implementação e utilização de mecanismos automáticos que possibilitem reunir informações sobre a facilogia do leito oceânico são fundamentais, sendo capazes de identificar fatores físicos que possam influenciar em diferentes atividades impactando, assim, nos custos e na operacionalidade do processo.

1.3 Objetivo geral

O objetivo geral deste trabalho é o desenvolvimento de uma metodologia computacional inteligente capaz de efetuar a classificação supervisionada dos sedimentos superficiais presentes no substrato marinho e de águas continentais, sem a necessidade frequente de amostragens físicas, demandando o mínimo possível do tempo de especialistas humanos.

Para atingir este objetivo, diversas etapas intermediárias serão empregadas com base em técnicas de visão computacional, reconhecimento de padrões e aprendizado de máquina aplicadas em sonografias de sonares de varredura lateral.

1.4 Objetivos específicos

Os objetivos específicos estão assim identificados:

- Levantamento e resumo das principais tecnologias e trabalhos científicos na área de classificação subaquática;
- Apresentar um comparativo entre as principais classes de descritores e algoritmos encontrados na literatura para classificação de sedimentos a partir de sonografias;
- Desenvolver um *framework* configurável capaz de classificar sonografias de sonar de varredura lateral;
- Disponibilização de uma metodologia visando ampliar a precisão do processo do processo de classificação;
- Disponibilização de um conjunto de técnicas para extração de descritores de imagens;
- Disponibilização de técnicas para transformação dos espaços de características provenientes dos descritores utilizados;

- Disponibilização de um conjunto de classificadores supervisionados calibrados para o problema em questão;
- Possibilidade de fusão (*ensemble*) de resultados a partir diferentes combinações de descritores e classificadores distintos baseado em sabedoria das massas;
- Visualização dos mapas temáticos de classificação;
- Visualização da qualidade dos resultados referentes às diferentes configurações, através de testes de desempenho, eficiência e comparação com resultados de classificação através de resultados gerados por especialistas.

1.5 Desafios

O processo de classificação de imagens subaquática é uma área relativamente nova quando comparado com outras áreas de pesquisa de visão computacional, como por exemplo, a detecção e a classificação de estruturas artificiais e naturais terrestres em imagens ópticas que muitas vezes não sofrem interferência do meio externo ([ASM, Shihavuddin, 2014](#)).

Pensando neste contexto, a utilização de imagens óticas subaquáticas, e mais especificamente neste caso, sonografias geradas a partir do eco acústico, proporcionam um conjunto de desafios que necessitam estudo, análise e tratamento adequados *a priori*.

Alguns dos desafios apresentados pelo processo de classificação automática baseado em sonografias de sonares de varredura lateral incluem:

- Baixa diferenciabilidade entre sedimentos de granulometria e rugosidade próximos ou aglomerados;
- Ruídos ocasionados pela assimetria da intensidade de ensonificação proporcionada por desestabilidade do equipamento durante aquisição dos dados;

- Presença de sombras acústicas (Inexistência de eco, proporcionado por obstáculo na direção do feixe acústico);
- Presenças de *Ripples* (Ondulações proporcionadas por fenômenos hidrodinâmicos, alterando a textura de sedimentos leves);
- Presença de estruturas artificiais sobre o substrato;
- Presença de eventos biológicos (Corais, Algas, etc...) sobre o substrato;
- Complexidade na geometria espacial entre as classes;
- Alterações graduais entre sedimentos de mesma classe, proporcionados por fenômenos ambientais ou biológicos.

Estes desafios, entre outros, são objetos de pesquisa detalhada em diferentes áreas como visão computacional e hidroacústica. Neste trabalho, considerando o elevado número de abordagens e considerações intrínsecas a estes desafios, os mesmos serão tratados de forma a obter níveis de resultados satisfatórios, consistindo em adaptações de métodos encontrados na bibliografia e inserção de abordagens alternativas visando múltiplos aspectos do mesmo problema.

Capítulo 2

Metodologia e trabalhos relacionados

2.1 Introdução

O desenvolvimento de sistemas autônomos que sejam capazes de reagir a estímulos visuais reproduzindo a capacidade humana de receber, interpretar e reagir a estes, ainda é um desafio. A atribuição a um sistema a capacidade humana de operar sobre um grande volume de informações visuais, distinguindo-as e interpretando-as, é o foco da análise de imagens digitais.

O crescimento desta área pode ser observado pelo interesse de cientistas e profissionais de diversos domínios do conhecimento e também pelo avanço contínuo da tecnologia que tem permitido o desenvolvimento de recursos cada vez mais poderosos para a interpretação de imagens ([Pedrini, 2008](#)). No contexto oceanográfico, o emprego de tecnologias baseadas em imagens torna-se uma alternativa atraente quando observa-se a necessidade de explorar ambientes marinhos em profundidade e área, oferecendo a capacidade de cobrir grandes áreas e atingir grandes profundidades de maneira eficiente. Fato este, que a comunidade científica contribui ativamente no intuito de reduzir

a dificuldade de explorar estes ambientes.

Neste intuito, este capítulo apresentará um levantamento bibliográfico sobre o estado da arte das técnicas de análise de imagens digitais, acústica submarina e sabedoria das massas, direcionados para classificação de sedimento superficiais em ambientes subaquáticos.

Cabe salientar, que muitas das técnicas descritas são amplamente discutidas e estabelecidas nas suas respectivas áreas do conhecimento, sendo utilizadas para diferentes fins. Este capítulo apenas retrata os assuntos de maneira a oferecer uma referência básica para que o leitor possa acompanhar a metodologia empregada neste trabalho.

2.2 Análise de Imagens Digitais

2.2.1 Introdução

Uma das tarefas mais importantes em sistemas baseados imagens é a análise e a interpretação destas, com o principal objetivo de obter uma boa descrição contendo informações suficientes que sejam capazes de distinguir os objetos que compõem a cena de forma confiável e requerendo o mínimo de intervenção humana.

Neste intuito, a análise de imagens digitais emprega um conjunto de conhecimentos e modelos que intersectam distintas áreas do conhecimento. A figura 2.1, exemplifica os principais tópicos envolvidos no campo de análise de imagens digitais, ressaltando que estes campos usualmente interseccionam-se entre as áreas de visão computacional, processamento de imagens e reconhecimento de padrões.

2.2.2 Metodologia

Visão Computacional

De acordo com (Pedrini, 2008), Visão Computacional procura auxiliar na resolução de problemas altamente complexos, buscando imitar a cognição humana e a habilidade

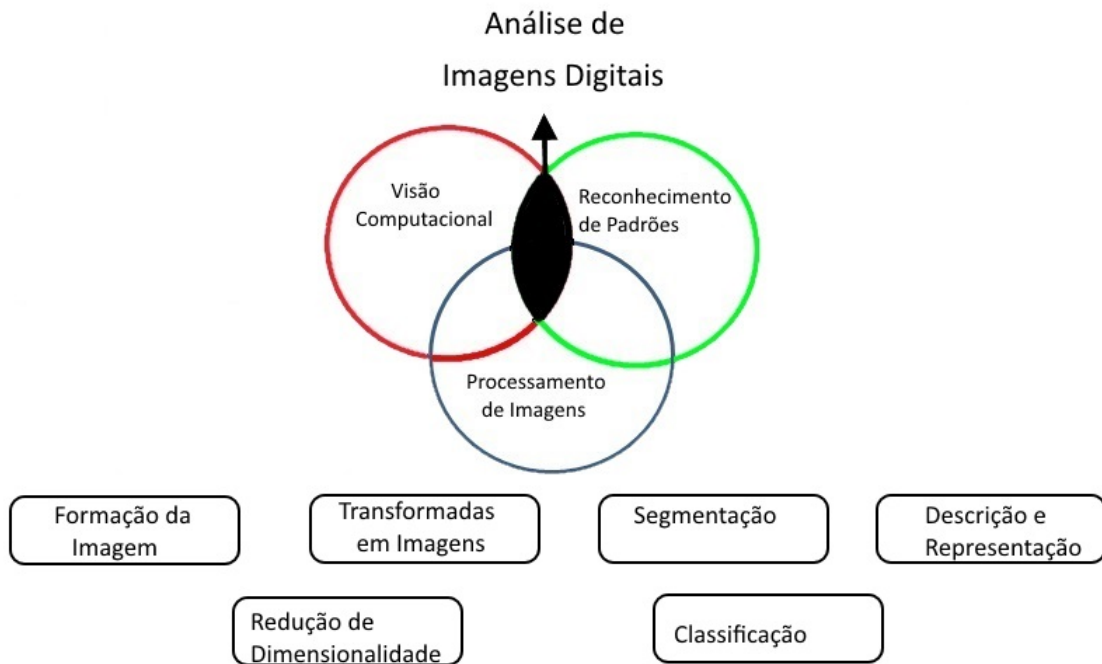


Figura 2.1: Campos relacionados a análise de imagens digitais.

do ser humano de tomar decisões de acordo com as informações contidas em imagens. Ou conforme (Forsyth and Ponce, 2011), visão computacional é agrupamento sólido de conhecimentos sobre câmeras e o processo físico de construção de imagens buscando obter simples inferências sobre conjuntos individuais de *pixels*, combinando-as de forma a extrair informações coerentes.

De forma mais ampla pode-se dizer que é a ciência responsável pela percepção visual de uma máquina ou seja, pela forma como esta enxerga o meio à sua volta. Através de informações significativas extraídas a partir de imagens capturadas por dispositivos como câmeras de vídeo, sensores, *scanners*. Estas informações permitem reconhecer, manipular e pensar sobre os objetos que compõem uma imagem (Ballard and Brown, 1982).

Como ciência, diferentemente do processamento digital de imagens o qual concentra-se em aplicar transformações em imagens visando construir outras, visão computacional

trata explicitamente da obtenção e manipulação das informações que podem ser extraídas de uma imagem. Conforme (Szaliski, 2010) o objetivo é descrever o mundo o qual visualizamos, a partir de propriedades extraídas de uma ou mais imagens.

Comparando a área de visão computacional com outras áreas de estudo, pode-se considerá-la com uma área relativamente recente. Seus trabalhos iniciaram-se por volta da década de 70, aliados a inteligência artificial, acreditando-se que seria possível representar completamente o sentido da visão em uma máquina (Honorato and Milano, 2010). Trabalhos como “*A framework for representing knowledge*” (Minsky, 1974) e “*The Psychology of Computer Vision*” (Winston, 1976) , destacam-se como precursores na área.

Processamento de Imagens

Considerado o primeiro estágio da grande maioria dos problemas que envolvem análise de imagens digitais, encontra-se o estágio de processamento de imagens. Centrado em representar e/ou transformar as imagens capturadas em novas imagens de forma que sejam mais adequadas para análise posterior (Szaliski, 2010). Podem ser citados como exemplos destas transformações operações como: correção de ruídos e balanceamento de cores. Algumas transformações direcionam-se ao ponto de vista estético da imagem enquanto outras, focam-se em salientar ou priorizar propriedades que possam ser utilizadas em etapas posteriores.

O emprego das técnicas de processamento de imagens permite ampliar a qualidade visual, assim como, identificar e extrair certos aspectos estruturais facilitando a percepção humana e a interpretação automática por meio de máquinas. Tendendo a ressaltar aspectos como formas, texturas e níveis de cor de objetos presentes na imagem.

Reconhecimento de Padrões

De acordo com ([Theodoridis and Koutroumbas, 2008](#)), reconhecimento de padrões pode ser definido como uma disciplina científica cujo o objetivo é a classificação de objetos entre um conjunto de categorias ou classes. Dependendo da aplicação estes objetos podem ser imagens, forma de ondas ou qualquer outro tipo de medidas que se deseje classificar.

A lista a seguir, visa esclarecer alguns conceitos fundamentais que são frequentemente empregados quando se aborda os temas de reconhecimento de padrões e classificação.

- **Padrão:** é uma entidade, objeto, processo ou evento, vagamente definido que pode assumir um nome.
- **Classe:** Conjunto de padrões que possuem características em comum.
- **Descritores:** Abrange um conjunto de algum tipo específico de características, por exemplo, descritores de textura.
- **Características** ou **Atributos:** dado extraído de uma amostra por meio de medida e/ou processamento. Em geral organizado na forma de um vetor de características.
- **Classificação:** atribuir classes para amostras, baseado em suas características.
- **Ruído:** distorção, falha ou imprecisão que ocorre na aquisição de dados.

O avanço de sistemas baseados em reconhecimento de padrões, vem acompanhado do avanço das tecnologias computacionais visto que muitos sistemas para classificação, principalmente os baseados em imagens de alta resolução requerem poder computacional elevado, observada a natureza matricial dos dados. Essencialmente sistemas baseados

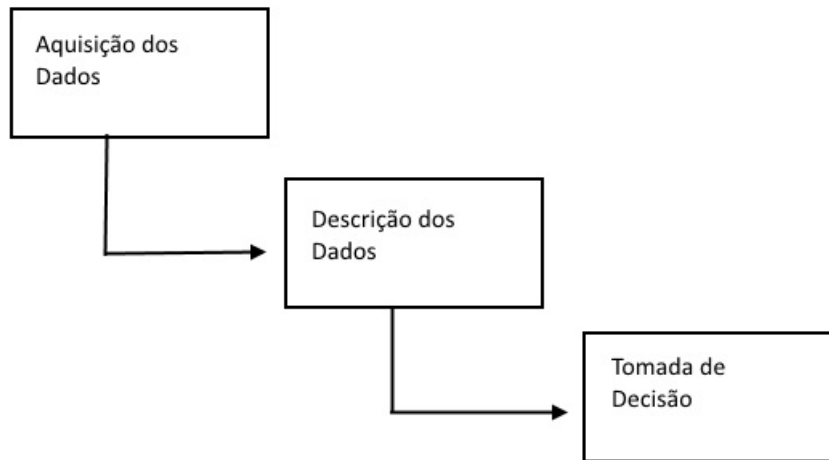


Figura 2.2: Etapas gerais reconhecimento de padrões.

em reconhecimentos de padrões, envolvem três etapas gerais: Aquisição dos dados, descrição dos dados e tomada de decisão. Como pode ser visto na figura 2.2.

A aquisição dos dados, centra-se na captura dos dados de acordo com o domínio do problema como baseados em imagens, textos, ondas sonoras entre outros, a etapa de descrição dos dados foca na extração de características capazes de definir de maneira adequada os objetos os quais se pretende diferenciar e classificar e a etapa de tomada de decisão, centra-se na especificação e construção de métodos classificadores capazes de diferenciar os objetos com base nas suas características intrínsecas (discutido nas próximas seções).

Logo, a definição de uma abordagem para o processo de reconhecimento de padrões não é uma tarefa fácil, não existindo uma abordagem “correta” e sim uma abordagem mais adequada para o problema em questão, o qual deve ser bem definido e suficientemente especificado.

Processo de classificação de padrões

Nas seções a seguir, serão descritos de maneira ampla as principais técnicas envolvidas no processo de classificação de padrões, visando servir como embasamento teórico básico para as metodologias e técnicas abordadas neste trabalho, as quais são novamente discutidas de maneira aplicada no capítulo 3.

Formação da Imagem

No centro de qualquer sistema que envolva análise de imagens digitais, as imagens podem ser definidas como uma função de intensidade luminosa denotada por $f(x,y)$, cujo valor ou amplitude nas coordenadas espaciais (x,y) fornece a intensidade ou o brilho na imagem naquele ponto (Pedrini, 2008).

Uma imagem digital pode ser obtida a partir do processo de digitalização o qual envolve dois passos, amostragem e quantização. O processo de **amostragem** consiste em discretizar o domínio de definição da imagem nas direções x e y , obtendo uma matriz bidimensional de $M \times N$ amostras, respectivamente. Enquanto o processo de **quantização** consiste em definir o número L de níveis de cinza (em uma imagem monocromática) para cada ponto da imagem.

Cada elemento $f(x,y)$ dessa matriz de amostras é chamado pixel (acrônimo do inglês *picture element*), logo a imagem contínua $f(x,y)$ é aproximada por uma matriz de pixels com dimensão M pixels na horizontal e N pixels na vertical.

Na figura 2.3, é ilustrada uma imagem e a orientação do sistema de coordenadas adotado e na equação 2.1 a transformação $f(x,y) \rightarrow M \times N$, respectivamente.

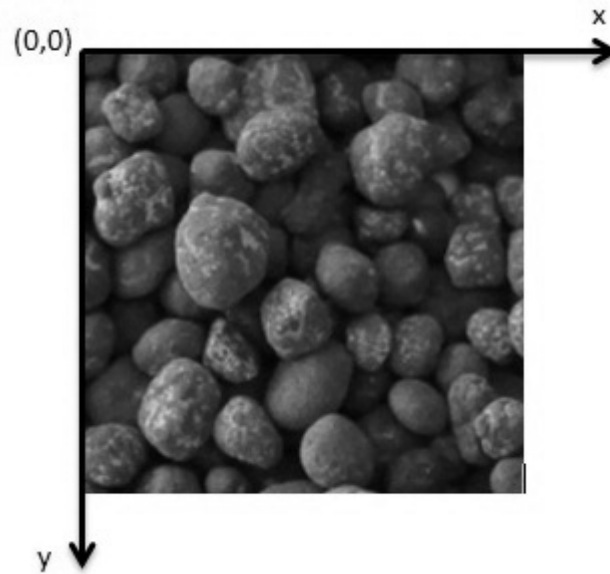


Figura 2.3: Convenção do sistema de coordenadas para representação de imagens digitais.

$$f(x,y) \approx \begin{bmatrix} f(0,0) & f(1,0) & \dots & f(M-1,0) \\ f(0,1) & f(1,1) & \dots & f(M-1,1) \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ f(0,N-1) & f(1,N-1) & \dots & f(M-1,N-1) \end{bmatrix} \quad (2.1)$$

A utilização do sistema matricial para representar imagens proporciona vantagens por se tratarem de estruturas de simples armazenamento, manipulação e visualização. Porém, sua invariabilidade espacial torna essa estrutura pouco adaptativa a eventuais irregularidades que possam existir na imagem.

Na figura 2.4 é exemplificada a representação matricial de uma imagem, uma pequena região é destacada, evidenciando números inteiros entre 0 e 255 referentes aos

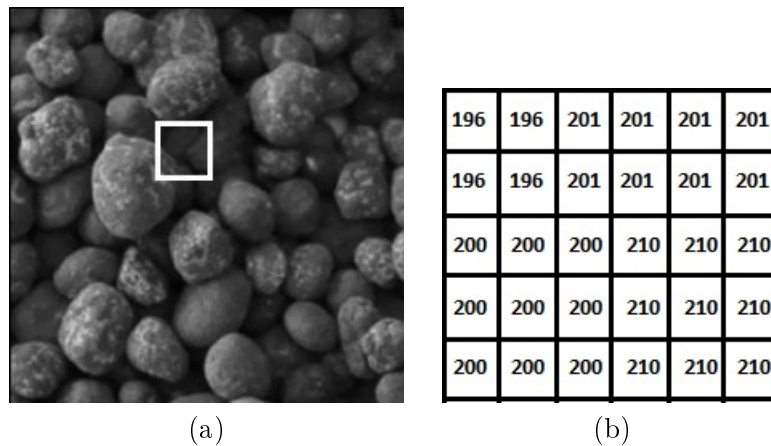


Figura 2.4: Representação matricial. (a) imagem. (b) níveis de cinza correspondentes à região da imagem em destaque.

níveis de tons de cinza dos pixels apresentados na região.

Em uma imagem monocromática, os valores dos pixels são escalares compreendidos entre L_{\min} e L_{\max} , porém imagens podem ser *multiespectrais*, *multibandas* ou coloridas onde nestas, cada pixel pode ser considerado com um valor vetorial $f(x,y) = (L_1, L_2, L_3, \dots, L_n)$, onde $L_{\min} \leq L_i \leq L_{\max}$ e $i = 1, 2, 3, \dots, n$. Usualmente em imagens coloridas cada ponto (x,y) é definido por três grandezas: luminância, matiz e saturação, associados ao brilho da luz, comprimento de onda dominante e o grau de pureza, respectivamente.

Imagens multiespectrais geralmente são muito utilizadas em sensoriamento remoto, no qual o sensor opera em diferentes faixas do espectro eletromagnético, denominadas bandas espectrais. Onde dependendo do alvo, tais como água, solo, vegetação, entre outras, a interação magnética produz diferentes respostas espectrais, com valores associados a posição espacial de um pixel de uma banda em particular.

Atualmente, existem diferentes técnicas para imageamento como o óptico e o acústico (seção 2.3), mais comumente encontrado, o imageamento óptico analógico e o digital envolvem fenômenos associados a luz e suas variações do espectro luminoso, que pode

ser descrito por seu comportamento ondulatório (óptica física) ou pela sua propagação no meio (óptica geométrica).

Transformadas em Imagem

A aplicação de transformadas em imagens resulta na alteração da representação inicial da imagem, de modo a prover informações que permitam uma análise mais acurada dos dados. Podendo ser aplicadas em uma vasta gama de aplicações no escopo de análise de imagens, por exemplo, utilizada na extração de características, compressão de dados, segmentação e filtragem de imagens (Pedrini, 2008).

Conceito fundamental do processamento de imagens são os filtros, técnicas de transformações aplicadas a cada pixel da imagem, considerando a relação entre o pixel e sua vizinhança. O processo de filtragem usualmente é feito utilizando matrizes denominadas máscaras de dimensões $m \times n$, as quais são convoluídas ou seja, deslocadas sobre a imagem original de dimensões $M \times N$, obtendo-se diferentes resultados de acordo com a máscara utilizada. As figuras 2.5 e 2.6, exemplificam a convolução e o resultado da aplicação de um filtro média sobre uma imagem, respectivamente.

Técnicas de filtragem são geralmente divididas em filtrações **lineares** e **não-lineares** e também agrupadas em dois domínios, o **domínio espacial** e o **domínio da frequência**.

A filtragem no domínio espacial refere-se ao próprio plano que compõem a imagem e um conjunto de valores que operam diretamente sobre os pixels da imagem, podendo ser expressa como demonstrado na equação 2.2 referente a uma convolução linear, onde h corresponde à máscara de acordo com filtro a ser utilizado, f é a imagem de entrada $N \times M$ com ruído, e g é a imagem resultante da filtragem no domínio espacial.

$$g(x,y) = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} f(i,j)h(x+i,y+i) \quad (2.2)$$

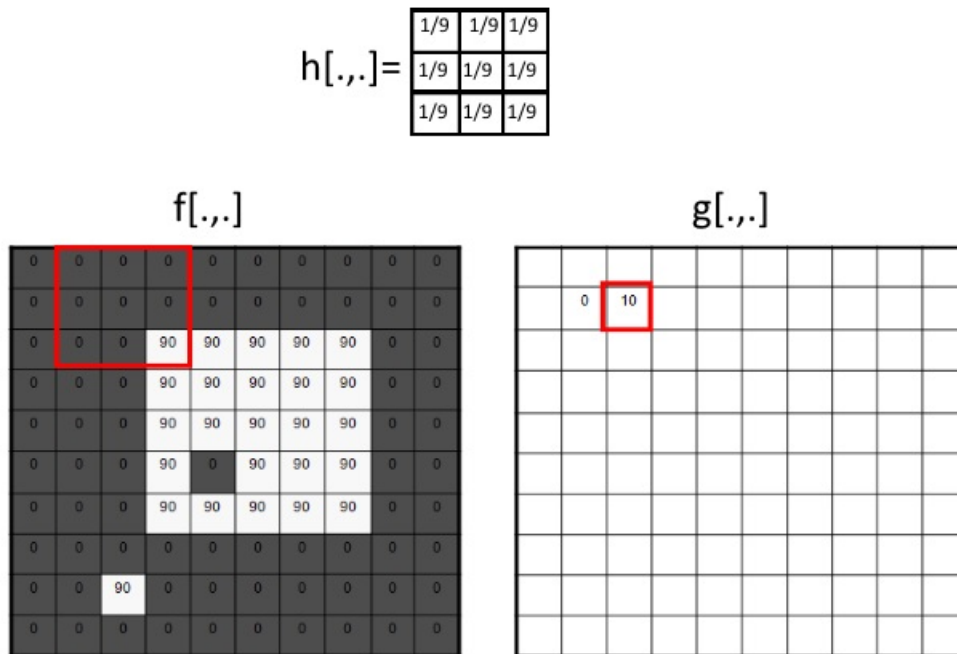


Figura 2.5: Exemplificação da convolução de um filtro média sobre uma imagem, matriz h máscara, f imagem original e g imagem filtrada.

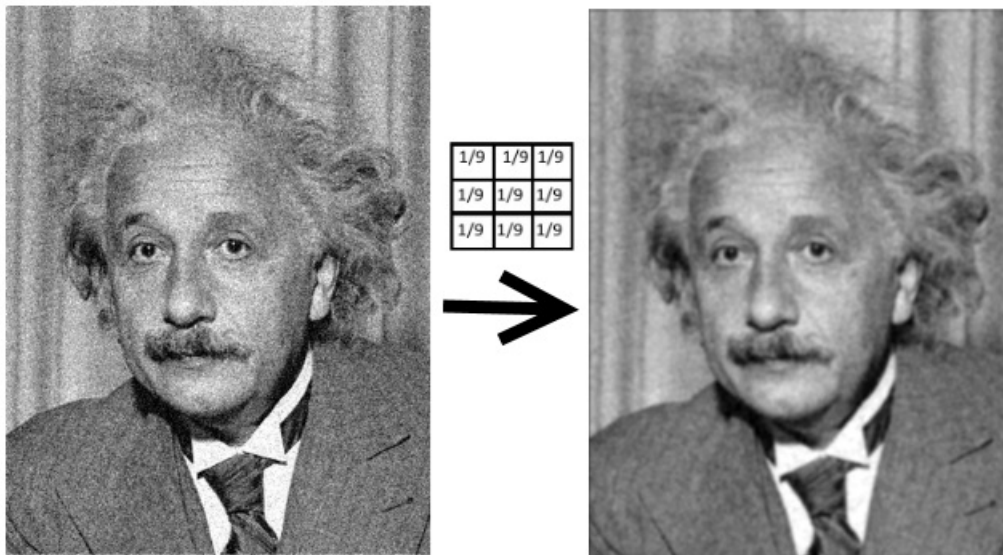


Figura 2.6: Exemplificação do resultado da aplicação de um filtro média sobre uma imagem.

Os filtros espaciais lineares suavizam e realçam detalhes da imagem, minimizando efeitos de ruídos sem alterar a média da imagem. Usualmente estes filtros são alocados categorias como:

- Filtros Passa-Baixa: Suavizam a imagem atenuando as altas frequências, que correspondem a mudanças abruptas.
- Filtros Passa-Alta: Realça de detalhes da imagem, tornando as transições entre regiões diferentes mais nítidas.
- Filtros Realce de bordas: Buscam realçar a cena, segundo direções preferenciais.

Os filtros espaciais não-lineares suavizam e realçam bordas na imagem, alterando a média da imagem. Usualmente estes filtros são alocados categorias como:

- Filtros Sobel: Realçam linhas horizontais e verticais mais escuras que o fundo, sem realçar pontos isolados.
- Filtros Morfológicos: Exploram as propriedades geométricas dos sinais (níveis de cinza da imagem), são exemplos comuns os filtros mediana, erosão e dilatação.

Para imagens digitais (uma função bidimensional discreta, $N \times M$), filtragens no domínio da frequência baseiam-se na modificação da imagem base, através da transformada de Fourier discreta (DFT - Discrete Fourier Transform, equação 2.3) para atingir um objetivo específico e na sequência é utilizada a sua transformada inversa (IDFT - Inverse Discrete Fourier Transform, equação 2.4) para retornar ao domínio espacial, conforme exemplificado na figura 2.7.

$$F(u,v) = \frac{1}{\sqrt{N} * \sqrt{M}} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} f(x,y) e^{\frac{-2\pi i}{N}(xu+yv)} \quad (2.3)$$

$$f(x,y) = \frac{1}{\sqrt{N} * \sqrt{M}} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} F(u,v) e^{\frac{+2\pi i}{N}(xu+yv)} \quad (2.4)$$

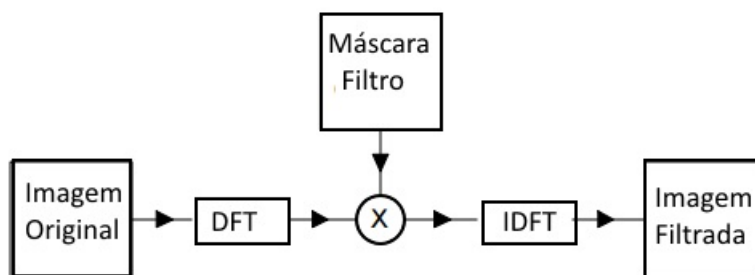


Figura 2.7: Exemplificação sequência completa filtragem no domínio da frequência.

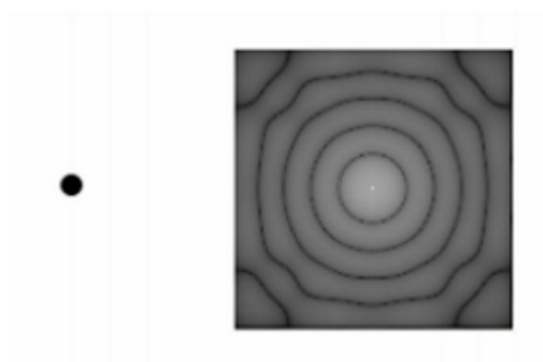


Figura 2.8: Exemplificação dimagem representada no domínio da frequência.

A aplicação destas transformadas, permite uma análise mais acurada dos dados. Por exemplo, a resultante da aplicação da transformada de Fourier pode ser utilizada como indicativo das variações presentes nos tons de pixels pois provê uma descrição quantitativa das frequências presentes na imagem (Pedrini, 2008).

É importante observar que durante as transformações não há perda de informação apenas, a informação visual que é representada de outra forma no domínio da frequência. Inicialmente parece complicado entender a visualização da imagem, pois, um ponto de uma imagem representada no domínio da frequência pode conter informações sobre toda uma imagem no domínio espacial (Figura 2.8), indicando quanto desta frequência há na imagem (Watt and Policarpo, 1998).

Entre as transformadas mais utilizadas encontram-se as baseadas em técnicas de

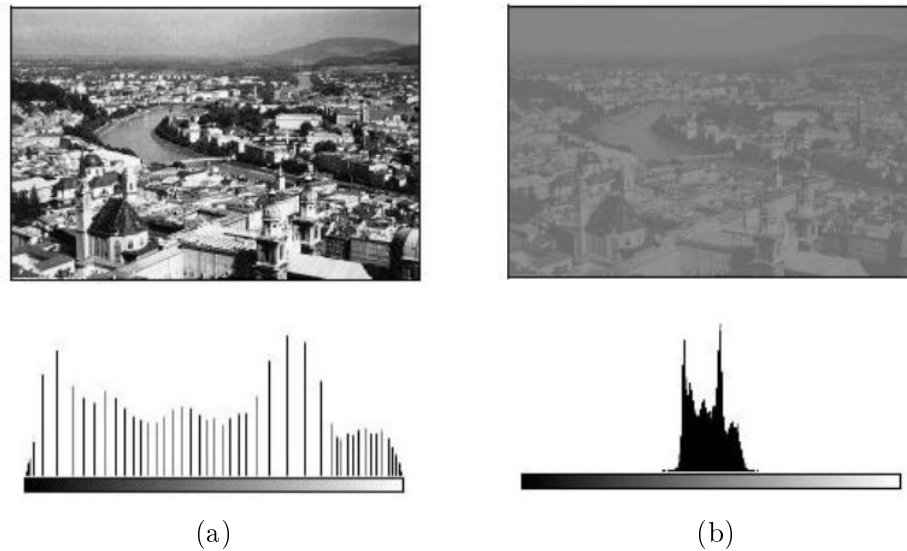


Figura 2.9: (a) imagem de entrada. (b) imagem saturada de níveis médios de cinza

realce de imagens que buscam acentuar ou melhorar a aparência de determinadas características na imagem, tornando-as mais adequadas a aplicação. Segundo (Pedrini, 2008), a aplicação de realce em imagens é necessária quando a imagem sofreu degradações em função de ruídos, perda de contraste, borramento, distorções causadas pelo equipamento de aquisição ou condição inadequada de iluminação.

Assim como as transformações, as técnicas de realce de imagens operam em duas categorias: métodos no domínio espacial e no domínio da frequência ambas discutidas no tópico anterior. Usualmente operações de realce agem diretamente através da manipulação do histograma da imagem, ou seja, através da distribuição dos níveis de cinza da imagem. O qual pode ser representado por um gráfico indicando a relação entre o número de pixels que compõem a imagem e quantidade de cada nível de cinza, como ilustrado na figura 2.9.

Dois conceitos fundamentais devem ser considerados quando o objetivo é realçar imagens digitais, **brilho** e **contraste**. O primeiro está associado à sensação visual da



Figura 2.10: (a) imagem de entrada. (b) imagem com brilho saturado

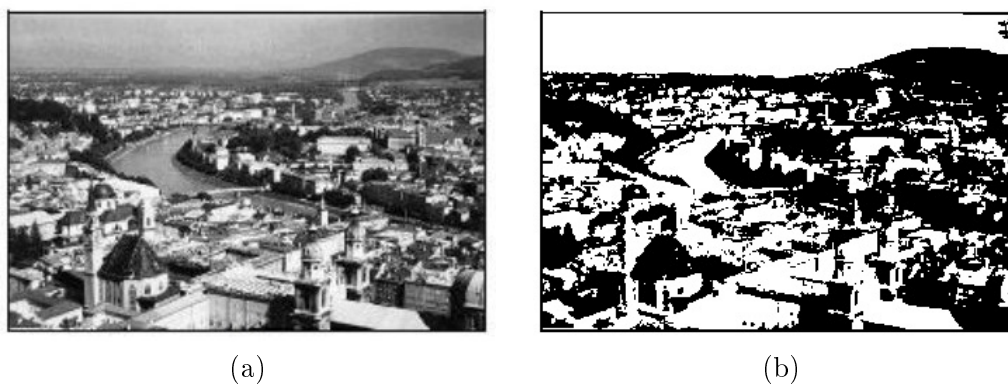


Figura 2.11: (a) imagem de entrada. (b) imagem com contraste saturado

intensidade luminosa de uma fonte (Figura 2.10) enquanto o segundo, está associado ao fato de que o brilho aparente de uma região depende fortemente da intensidade do fundo (Figura 2.11).

Operações baseadas em transformações de brilho e contraste aplicadas a histogramas de imagens, objetivam melhorar a qualidade das imagens sob critérios específicos, tornando mais fácil a percepção de informações contidas nas imagens. Algumas técnicas como equalização, hiperbolização e especificação de histogramas são de grande utilização em realce de imagens, com a finalidade de produzir imagens com histogramas de

formato específico.

De forma geral, na equalização de histogramas é buscado transformar a imagem original de modo que esta possua uma distribuição uniforme nos seus níveis de cinza, ou seja, os níveis de cinza devem aparecer aproximadamente na mesma frequência. A técnica de hiperbolização de histogramas, tem por objetivo melhorar a qualidade visual levando em conta aspectos relativos à percepção de brilho, ou seja o histograma é uniformizado baseado no brilho, cuja forma é hiperbólica e na especificação de histogramas é aplicado uma transformação de no intuito que o histograma apresente uma distribuição particular.

Segmentação

O processo de interpretação de imagens digitais, dado o volume de dados contidos nestas imagens tende a ser uma tarefa complexa. Uma etapa intermediária é necessária visando particionar o conjunto de dados em grupos que contemplem um conteúdo semântico semelhante, esta etapa é chamada segmentação.

De acordo com (Pedrini, 2008) um processo de segmentação que identifique corretamente a localização, a topologia e a forma dos objetos é um requisito de fundamental importância para que as informações resultantes de um sistema de análise de imagens sejam confiáveis. (Szaliski, 2010), argumenta que a segmentação de imagens é uma tarefa que busca agrupar pixels que “andam juntos”, na estatística este problema é conhecido como análise de *cluster* e é amplamente estudado para diversos fins.

Processos de segmentação geralmente abordam duas estratégias **descontinuidade** e **similariadade**, na estratégia por descontinuidade a imagem é particionada com base nas alterações bruscas de intensidade, caracterizados pela presença de pontos isolados. Na estratégia por similaridade o particionamento é construído com base na similaridade de determinado grupos de características entre os pixels, a seleção destas características está relacionada com a heterogeneidade das regiões.

Segmentação por regiões

Em contraste com técnicas de segmentação que buscam descontinuidades na imagem como contornos (Szaliski, 2010), métodos de segmentação por regiões detectam e delimitam áreas que compartilham características em comum diretamente na imagem, ou seja, os pixels são agrupados de maneira a estabelecer uma região comum. Diversas propriedades tem sido propostas para caracterizar regiões, tais como intensidade de níveis de cinza, cor, informação semântica ou textura (Haralick and Shapiro, 1992).

Definindo a própria imagem de entrada como uma região R , pode-se imaginar a segmentação como o processo de particionar R em n regiões R_1, R_2, \dots, R_n , tal que:

1. $\bigcup_{i=1}^n R_i = R$
2. R_i é uma função conexa, $i = 1, 2, \dots, n$
3. $R_i \cap R_j = \emptyset$
4. $P(R_i) = VERDADEIRO$ para $i = 1, 2, \dots, n$
5. $P(R_i \cup R_j) = FALSO$ para $i \neq j$ e R_i adjacente a R_j

Sendo $P(R_i)$ um predicado lógico sobre os pontos do conjunto R_i e \emptyset é o conjunto vazio. O predicado é utilizado como uma medida de similaridade para agrupar os pixels em uma região. A condição 1 indica que cada pixel deve pertencer a uma região da imagem, condição 2 requer que os pixels de uma região satisfaçam critérios de conectividade, 3 estabelece que as regiões devem ser disjuntas, 4 refere-se as propriedades que devem ser satisfeitas pelos pixels em uma região segmentada e 5 estabelece que regiões adjacentes R_i e R_j são diferentes em relação ao predicado P .

O processo de segmentação de imagens em regiões homogêneas disjuntas é ilustrado na figura 2.12

Diversas abordagens para segmentação têm sido extensivamente estudadas e revistas na literatura, para uma discussão mais detalhada, cabe consultar (Haralick and

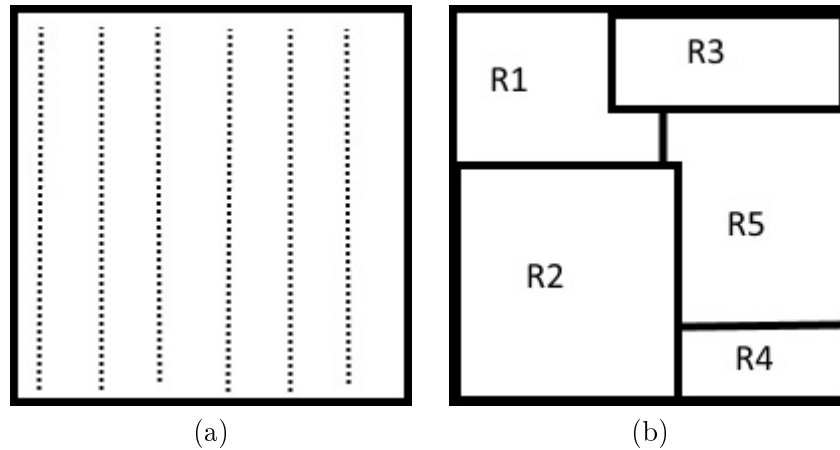


Figura 2.12: (a) imagem de entrada. (b) imagem segmentada em regiões

Shapiro, 1992), (Watt and Policarpo, 1998), (Theodoridis and Koutroumbas, 2008), (Pedrini, 2008) e (Szaliski, 2010).

Superpixels

Superpixel é uma técnica para segmentação de imagens baseada em regiões, usada para redução do número de *pixels* da imagem agrupando estes de acordo com propriedades semelhantes. Dentre as técnicas de *superpixels* mais utilizadas está a *Simple Linear Iterative Clustering - SLIC*, proposta por (Achanta et al., 2010).

O Algoritmo *SLIC* apresenta o método de agrupamento baseado no algoritmo *K-means* com um espaço de busca mais eficiente, visto que seu tamanho é limitado. A abordagem proposta para segmentação do *SLIC* pode ser descrita como:

- A partir de uma imagem de dimensões $N \times N$ são criadas k regiões retangulares (*Superpixels*) de dimensões $s \times s$, onde $s = \frac{N}{k}$. Cada *superpixel* é representado por um centróide definido como um vetor de cinco dimensões $C_j = L_j, a_j, b_j, x_j, y_j$, onde L é a luminância, a e b o intervalo de cores e x e y as coordenadas do *superpixel* j .

- A cada iteração, cada pixel i é associado com o centróide do *superpixel* mais próximo, considerando que cada superpixel possui uma área de busca de $2s \times 2s$. Na sequência, novos centróides são calculados como a média do vetor C_j , considerando todos os *pixels* do *superpixel*. Este processo é executado através da minimização da função D , definida pela equação 2.5.

$$D = \sqrt{d_c^2 + \left(\frac{d_s}{S}\right)^2 m^2} \quad (2.5)$$

Onde m controla a capacidade do *superpixel*, d_c é a distância no espaço de cores e d_s é a distância espacial, definidas pelas equações 2.6 e 2.7, respectivamente.

$$d_c = \sqrt{(l_j - l_i)^2 + (a_j - a_i)^2 + (b_j - b_i)^2} \quad (2.6)$$

$$d_s = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} \quad (2.7)$$

Descrição e Representação

Um dos problemas básicos, porém complexos, encontrados no desenvolvimento de sistemas para classificação, é a extração e seleção de um **conjunto de características** ou **espaço de características** que possam ser utilizadas com o propósito de descrever de maneira efetiva cada região contida na imagem.

Considerando que cada região na imagem pode ser descrita em termos de suas propriedades externas como bordas ou características internas como a relação entre o conjunto de pixels que a compõem, por exemplo texturas (Figura 2.14).

De maneira a exemplificar, considere apenas texturas que de acordo com (Haralick and Shanmugan, 1973), pode ser definida pela interação entre as primitivas tonais que a compõem, estas ocorrendo em diferentes formas e números. Características de textura podem ser extraídas utilizando o conceito de matrizes de co-ocorrência de níveis de cinza (*Grey Level Co-Occurrancy Matrix - GLCM*) descritas por (Haralick and Shanmugan,

1973) que consistem na contagem de quantas combinações de diferentes transições de níveis de cinza ocorrem na imagem.

A matriz de co-ocorrência é uma matriz quadrada com suas dimensões $N \times N$, sendo N referente a quantidade de níveis de cinza ($P_{i,j}$) da imagem analisada. As combinações são calculadas sobre ângulos 0, 45, 90 e 135 e os demais via simetria, conforme ilustrado na figura 2.13. Baseado na matrix de co-ocorrência é possível extrair vinte e quatro atributos distintos, porém, os mais representativos e utilizados são:

1. A energia também conhecida como segundo momento angular, permite identificar o grau de homogeneidade das medidas de distância.

$$\text{Energia} = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} P(i,j)^2 \quad (2.8)$$

2. Entropia, mede o nível de desordem das amostras.

$$\text{Entropia} = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} P(i,j) \log(P(i,j)) \quad (2.9)$$

3. variância (contraste), mede a variação entre as distâncias.

$$\text{Variância} = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} P(i,j)(i-j)^2 \quad (2.10)$$

4. Homogeneidade, indica se os valores entre as distâncias são constantes.

$$\text{Homogeneidade} = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \left(\frac{1}{1 + (i-j)^2} \right) P(i,j) \quad (2.11)$$

5. Dissimilaridade, semelhante a variância porém, com crescimento dos pesos linear.

$$\text{Dissimilaridade} = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} |i-j| P(i,j) \quad (2.12)$$

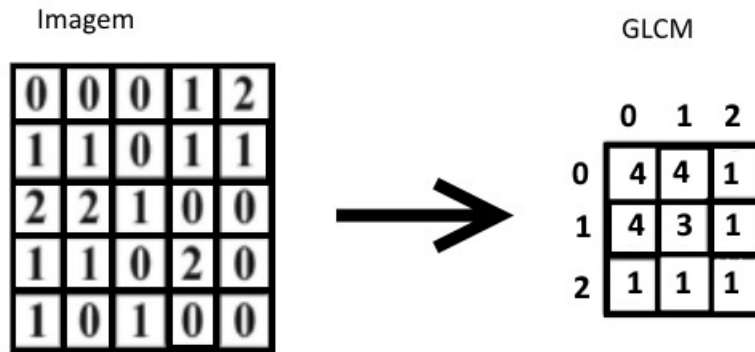


Figura 2.13: Exemplificação imagem representada em três níveis de cinza na forma matricial e sua respectiva GLCM na direção 0.

Como observado apenas baseado em textura e utilizando apenas uma técnica sobre texturas é possível extrair um número elevado de atributos. No entanto, a descrição depende da representação adotada e deve permitir a caracterização das amostras sem ambiguidades e com um número reduzido de características ou medidas extraídas da amostra evitando níveis de dimensionalidade elevados (Pedrini, 2008).

A questão é, quais características extrair de um objeto encontrado na imagem, de maneira que estas sejam suficientemente descritivas ao ponto de representar o objeto sem redundâncias? Descritores baseados em cores, níveis de cinza, padrões de forma, intensidade e textura são alguns dos mais utilizados, porém, geralmente requerem elevado poder computacional para operar sobre espaços de características de dimensões elevadas e que não garantem alto poder discriminatório, visto que nenhuma relação entre as amostras é considerada nesta etapa.

Para se obter apenas medidas que sejam representativas, se faz necessária a aplicação de uma etapa de seleção ou transformação sobre o espaço de características extraído, com o objetivo de selecionar o melhor conjunto de características de maneira reduzir sua dimensão e aumentar a acurácia.

Portanto é fortemente desejável a utilização de espaços de características de dimensionalidade reduzida e concisos que proporcionem nível satisfatório de separabilidade

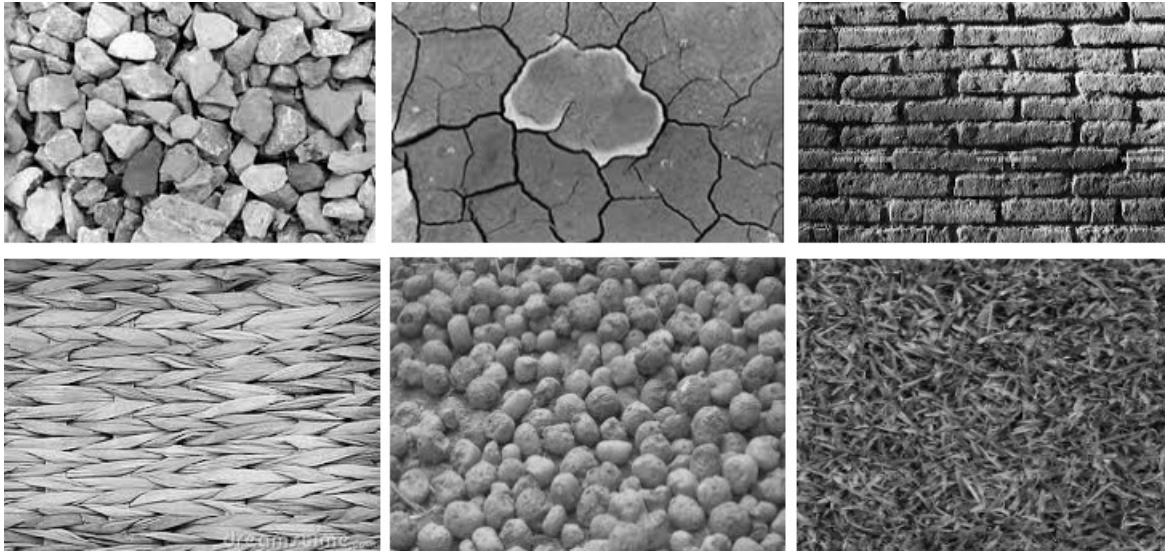


Figura 2.14: Ilustração texturas naturais e artificiais.

entre as amostras. A seção a seguir, foca neste conceito.

Redução de Dimensionalidade

De acordo com (Hair et al., 1995), a redução de dimensionalidade é uma das mais importantes formas de regressão, pois permite eliminar subconjuntos de atributos – dos processos de análise subsequentes – do conjunto original de atributos, muitas vezes de alta dimensão, que descrevem os objetos.

Quando o objetivo é descrever determinado objeto em termos de suas características, pode ser criado um conjunto de características de dimensões elevadas onde nem todas estas refletem propriedades realmente descritivas do objeto. Logo, o termo redução de dimensionalidade refere-se a redução do espaço de características de uma representação de padrões.

A redução de espaços de características de dimensões menos elevadas refletem diretamente sobre o custo de medição e precisão do classificador. Mantendo somente as características mais salientes, o classificador será mais eficiente, eficaz e ocupará menos memória.

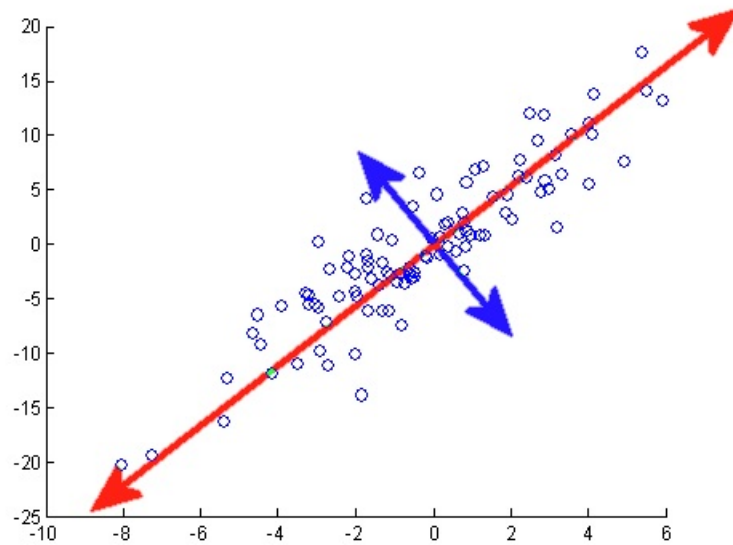


Figura 2.15: Ilustração PCA, linha vermelha mostra a distribuição principal dos dados e a linha azul a componente secundária.

Neste contexto, existem diferentes métodos para redução do espaço de características como *Rough Sets* baseado na seleção de características e a análise de componentes principais - ACP ou do inglês *Principal Component Analysis* - PCA baseado na transformação das características. Baseado na literatura, PCA é uma dos métodos estatísticos multivariados mais utilizados, o objetivo da análise de componentes principais é tomar n variáveis correlacionadas X_1, X_2, \dots, X_n e convertê-los em um conjunto menor de variáveis Z_1, Z_2, \dots, Z_m que descrevam a variação dos dados e não sejam correlacionados na sua ordem de importância (Araujo and Coelho, 2009), ilustrado na figura 2.15.

A análise de componentes principais é uma estratégia para identificar a relação entre características extraídas dos dados, a **componente principal** é o arranjo que melhor representa a distribuição dos dados e a **componente secundária** é perpendicular a componente principal.

O cálculo das componentes principais, tradicionalmente aborda os seguintes passos:

- Obter os dados ou as M amostras de vetores de dimensão n .

- Calcular a média ou o vetor médio destes dados.
- Subtrair a média de todos os itens de dados.
- Calcular a matriz de covariância usando todas as subtrações. Sendo esta, o resultado da média do produto de cada subtração por ela mesma e terá dimensões $n \times n$.
- Calcular os autovalores e auto vetores da matriz de covariância.
- Arranjar os elementos de modo que o elemento (0,0) seja o auto vetor correspondente ao maior autovalor, e assim, sucessivamente até o menor auto valor.

O auto vetor com o maior autovalor associado, corresponde à componente principal do conjunto de dados utilizado, significando que é o relacionamento mais significativo entre as dimensões dos dados. As componentes principais podem então ser usadas conforme a maneira desejada. Após transformados, o novo conjunto pode ser utilizado para visualização, para reconhecimento das principais características de medidas a serem usadas entre outros fins.

Para maiores discussões: ([Geladi et al., 1987](#)), ([Geladi et al., 1989](#)), ([Zortea and Haertel, 2004](#)), ([Borges, 2006](#)), ([Theodoridis and Koutroumbas, 2008](#)) e ([Araujo and Coelho, 2009](#)).

Classificação de Padrões

O processo de classificação de padrões visa determinar um mapeamento capaz de relacionar as propriedades extraídas de amostras com um conjunto de rótulos, apresentando a restrição de que amostras com características semelhantes devem ser mapeadas ao mesmo rótulo ([Pedrini, 2008](#)).

Quando agrupa-se diferentes amostras sob o mesmo rótulo diz-se que tais amostras pertencem a uma mesma classe, essa caracterizada por compreender elementos

que compartilham características em comum. Algoritmos e sistemas que se dispõem a estabelecer o mapeamento entre as propriedades do conjunto de amostras e a classe a qual apresenta maior afinidade são denotados algoritmos de classificação e sistemas classificadores, respectivamente.

Usualmente existem dois problemas intrínsecos a qualquer sistema de classificação: mapeamento dos objetos de acordo com as classes previamente estabelecidas e o agrupamento destes a partir de parâmetros significativamente discriminantes que os caracterizem. Embora os processos envolvidos nestas etapas sejam distintos ambos possuem por objetivo estimar o melhor rotulo para o objeto analisado.

Sistemas para classificação podem ser subdivididos em duas categorias, sistema de classificação **supervisionada** e **não supervisionada**. Em sistemas de classificação ditos supervisionados, para que seja possível obter os parâmetros que caracterizam cada classe, uma etapa denomina **treinamento** deve ser executada anteriormente a aplicação da etapa de classificação, etapa na qual as características que melhor descrevem os objetos a serem classificados são obtidos a partir de um conjunto de amostras previamente identificadas e devidamente rotuladas. Em sistemas não supervisionados, todos os parâmetros são obtidos a partir das próprias amostras a serem classificadas.

Na literatura referente a classificação de padrões, diferentes abordagens são inferidas para classificadores de modo a proporcionar maior nível de acurácia de acordo com a classe do problema abordado. Algumas das abordagens mais utilizadas são:

- Casamento de modelos (*Template Matching*).
- Classificadores de distância mínima.
- Casamento de modelos por correlação.
- Técnicas estatísticas.

O casamento de modelos é uma das técnicas mais simples utilizadas e baseia-se em

determinar a similaridade entre duas entidades do mesmo tipo baseando-se na utilização de funções de decisão ou discriminantes. Sendo o modelo tipicamente um protótipo aprendido a partir de um conjunto de treinamento, a similaridade é usualmente aferida em relação a uma função distâncias. Esse método é computacionalmente exigente, mas a atual disponibilidade de recursos computacionais permite com que essas abordagens viabilizem-se mais facilmente (Jain et al., 2000).

Classificadores de distância mínima, baseiam-se em atribuir uma classe ao objeto com base, na menor medida de distância entre os atributos que caracterizam o objeto e as possíveis classes. Através de distâncias como a Euclidiana ou a de Hamming, reduzindo o problema a computação de distâncias. Considere por exemplo, que cada classe seja representada por um vetor médio de características C com N características 2.13:

$$C_j = \frac{1}{N} \sum_{i=0}^{i=N} x \quad (2.13)$$

A distância pode ser aferida pela distância Euclidiana de acordo com a equação 2.14, sendo X a amostra a ser classificada. Na sequência atribui-se ao objeto X à classe C_j cuja $D_j(X)$ é a menor distância.

$$D_j = \|X - C_j\| \quad (2.14)$$

Em classificadores baseados em casamento de modelos por correlação, segundo (Gonzales and Woods, 2007), a correlação entre imagens é considerada a base para identificar o casamento de uma sub-imagem $w(x,y)$ de tamanho $m \times n$ dentro de uma imagem $f(x,y)$ de tamanho $M \times N$, sendo $m \leq M$ e $n \leq N$. De outra forma pode-se denotar a correlação entre $f(x,y)$ e $w(x,y)$ conforme apresentada na equação 2.15.

$$c(s,t) = \sum_x \sum_y f(x,y)w(x-s,y-t) \quad (2.15)$$

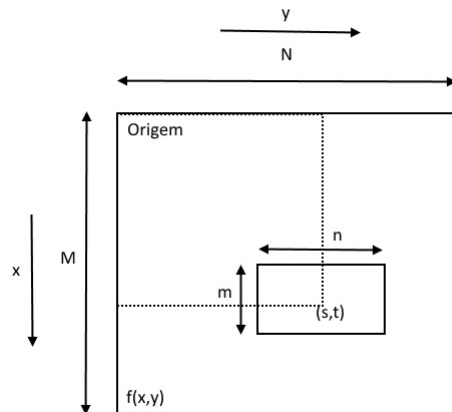


Figura 2.16: Ilustração esquema para se obter a correlação de $f(x,y)$ e $w(x,y)$ no ponto (s,t) , adaptado de (Gonzales and Woods, 2007)

onde $s = 0,1,2,\dots,M-1$ e $t = 0,1,2,\dots,N-1$ e a soma é realizada sobre a região em que as imagens f e w se sobreponham, conforme ilustrado na figura 2.16.

Utilizando classificadores baseados na abordagem estatística, cada amostra é representada em termos de suas características ou atributos. Tais características são expressas por meio de medidas compondo um espaço de características, que apresenta dimensionalidade igual ao número de medidas utilizadas para descrição das amostras (Pedrini, 2008). Idealmente, o espaço de características deve enfatizar a separabilidade entre as amostras com propriedades distintas e evitar a alta dimensionalidade, possibilitando a classificação em tempo aceitável.

Conforme (Jain et al., 2000), sistema de reconhecimento baseados na abordagem estatística operam em dois modos: Treinamento (aprendizagem) e Classificação (teste), conforme ilustrado na figura 2.17.

No intuito de otimizar um classificador supervisionado, muitas vezes o conjunto de treinamento é cuidadosamente selecionado visando diminuir possíveis anomalias presentes nas amostras, porém, como efeito colateral o sistema acaba perdendo a habilidade de generalização. Ocasionalmente problemas que devem ser considerados como o **super-treinamento** ou **super-adaptação**.

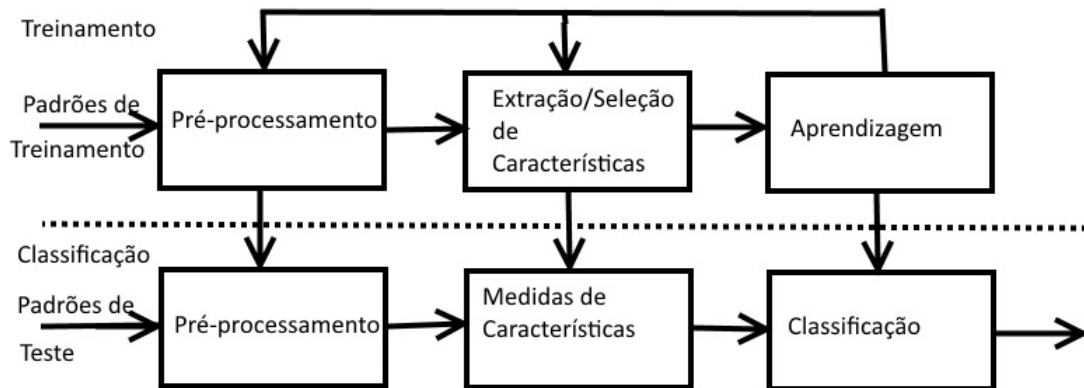


Figura 2.17: Exemplificação blocos funcionais para reconhecimento de padrões na abordagem estatística, baseado em (Jain et al., 2000)

A habilidade de generalização pode ser enfraquecida através de diversos fatores como utilizar conjuntos pequenos de amostras com muitas características para treinamento, grande número de parâmetros desconhecidos associados ao classificador ou um classificador ser intensivamente otimizado sobre o conjunto de treinamento.

2.2.3 Aplicações

Atualmente, nos deparamos imersos em sistema e aplicações que valem-se de análise de imagens digitais de inúmeras formas e com diferentes propósitos. Alguns destes sistemas, encontram-se de maneira tão amplamente inseridos no contexto diário de todos que muitas vezes passam de maneira despercebida ao observador.

São exemplos:

- Reconhecimento Facial
- Reconhecimento biométrico
- Diagnósticos médicos auxiliados por computador, a partir de exames baseados em imageamento digital.

- Mapeamento urbano, residencial e ambiental.
- Veículos de navegação autônoma, segurança automotiva.
- Interação humano-computador.
- Indústria, controle de qualidade.
- Interpretação de imagens de satélite.
- Segurança, detecção de intrusos.
- e muitos outros.

2.3 Acústica Submarina

2.3.1 Introdução

Conforme ([Kunzendorf, 1986](#)) na maioria dos casos, a exploração de ambientes marinhos começa com o mapeamento do leito valendo-se de métodos geofísicos e geológicos. Os quais, podem ser aplicados em paralelo para pesquisas topográficas resultando em dados na forma de imagens, vídeos, amostragem de sedimentos, entre outros.

De acordo com ([Anderson et al., 2008](#)), a classificação acústica do leito marinho (*Acoustic Seabed Classification - ASC*) é uma ciência em nascimento, visto que, o sensoriamento remoto do leito marinho baseado em diferentes métodos acústicos estão proporcionando soluções tecnológicas para iniciativas de mapeamento do habitat marinho.

A eficácia da propagação do som em ambientes subaquáticos ocasionou o avanço dos métodos hidroacústicos, utilizados durante a segunda guerra mundial com propósitos de reconhecimento de submarinos e de objetos sobre o leito oceânico. Neste contexto, foi proposto o conceito de força de espalhamento do inglês *scattering strenght*, desenvolvido para quantificar o espalhamento acústico quando incidido sobre o leito marinho ou

objetos encontrados na coluna d'água. Diversas medidas foram feitas durante o período de guerra e nas décadas seguintes visando categorizar a força do espalhamento acústico sobre diferentes tipos de sedimentos encontrados no substrato marinho como rochas, areia, cascalho e argila (Welsby, 1971).

Atualmente, a acústica submarina é aplicada em diversas atividades de pesquisa e exploração, não apenas para geofísica e geologia marinha mas também, como tecnologias de pesca para estimativas de abundância, mapeamento de comunidades biológicas, detecção de impacto humano, busca e recuperação de objetos naufragados, entre outros.

Objeto de interesse desta seção, a geoacústica, concentra-se em estudar os fenômenos que envolvem a propagação do som na água, assim como, a interação entre o substrato e as ondas mecânicas que constituem o som direcionada a investigação da superfície e subsuperfície de oceanos e águas continentais.

2.3.2 Metodologia

Propriedades Geoacústicas

Com base em diversos estudos, foram estimadas relações empíricas relacionando propriedades físicas dos sedimentos como porosidade, densidade de massa e tamanho médio do grão com propriedades geoacústicas os sedimentos (Welsby, 1971). De acordo com (Oliveira Junior, 2012), as propriedades geoacústicas podem ser subdivididas em três termos: impedância acústica, rugosidade e heterogeneidades volumétricas conforme detalhado a seguir.

Dentre as propriedades relacionadas à impedância acústica, destacam-se:

- Razão de velocidade do som, que a razão entre a velocidade do som no sedimento e na água sobre os mesmos níveis de temperatura, pressão e salinidade.
- Fator de atenuação ($k, dB/m$), que considera a influência da frequência do sistema de medição acústica no valor final da atenuação do sinal.

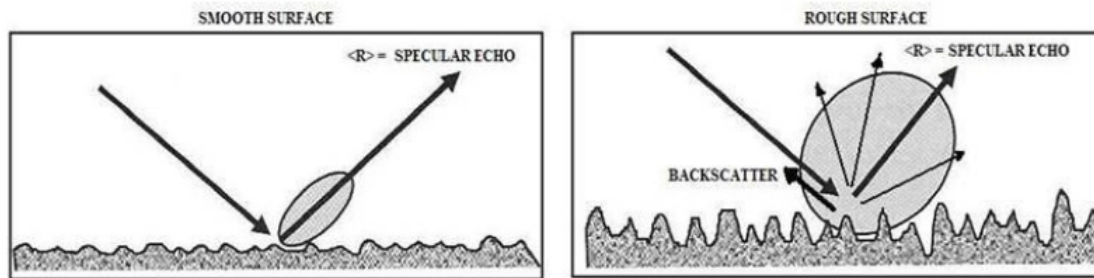


Figura 2.18: Ilustração efeito da rugosidade do fundo no retroespalhamento. Fonte: (Lurton, 2002)

A rugosidade, veja figura 2.19 presente no leito marinho é relacionada as variadas ondulações presentes na sua morfologia resultantes por fenômenos hidrodinâmicos, podendo ser representada estatisticamente por sua densidade de potência espectral (espectro de rugosidade, Figura 2.20). Para isso decompõem-se a rugosidade do fundo em suas diversas componentes espectrais, sendo as maiores amplitudes relacionadas aos maiores comprimentos de onda (Oliveira Junior, 2012).

A influência da rugosidade do fundo age diretamente sobre o retroespalhamento (*Backscatter*) do eco acústico que é retornado ao equipamento, havendo maior influência dependendo do tipo de sedimento e da frequência utilizada (Figura 2.18).

Em termos de heterogeneidade volumétrica dos sedimentos, as características dos sedimentos costumam apresentar variações de suas propriedades físicas a medida que são estudadas suas camadas mais profundas (Welsby, 1971).

Sonar de Varredura Lateral

O sonar de varredura lateral - SVL ou *sidescan sonar - SSS* (Figura 2.21), é um equipamento utilizado para busca, detecção e mapeamento de alta resolução de objetos e áreas submersas, realizando uma varredura lateral através da emissão de pulsos acústicos (Bulla et al., 2003).

Este tipo de sonar, acopla os transdutores na lateral do peixe do sonar (*Town Fish*)



Figura 2.19: Exemplo rugosidade relacionada as variadas ondulações resultantes por fenômenos hidrodinâmicos.

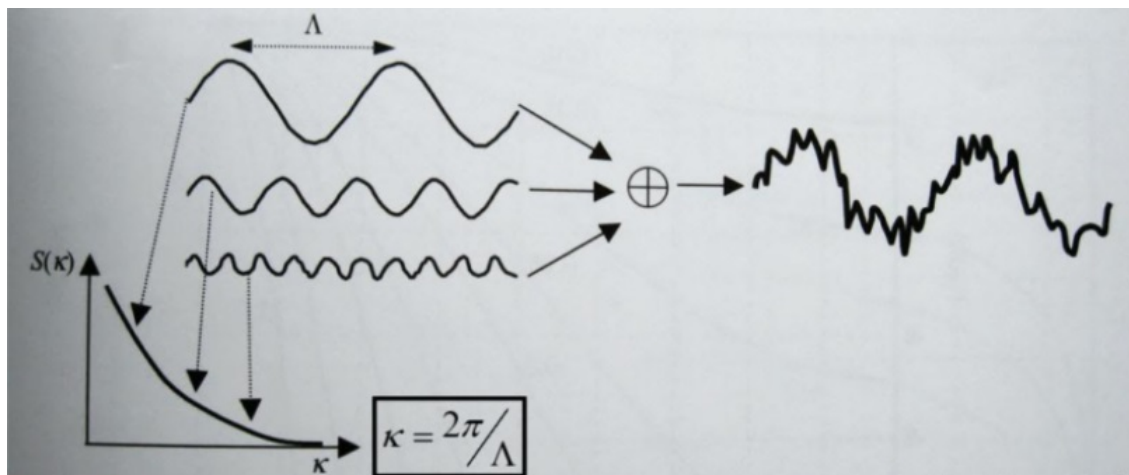


Figura 2.20: Exemplo decomposição da rugosidade de fundo em suas diversas componentes espectrais. Fonte: (Lurton, 2002)



Figura 2.21: Exemplo Sonar de Varredura Lateral. Fonte:<http://www.l-3mps.com/klein/sidescansonar.aspx>

que é rebocado por um navio, emitindo pulsos acústicos estreitos no plano horizontal e largos no plano vertical, ensonificando uma fina fatia do substrato a cada pulso. Quando a onda sonora encontra o solo ela é espalhada no caso de fundos rugosos ou refletida para longe, no caso de fundos lisos. O eco que retorna ao sonar é resultante das propriedades do fundo, como rugosidade e granulometria. O funcionamento do sonar de varredura lateral é ilustrado na figura 2.22

Imageamento acústico

Dado objetivos como mapear, monitorar ou inspecionar estruturas submersas ou o leito marinho, a obtenção de imagens em ambientes subaquáticos a partir de sistemas ópticos se torna uma tarefa complicada devido, principalmente, a baixa ou inexistente presença de luz, requisito básico para o trabalho com sistemas ópticos. A construção de imagens por equipamentos acústicos é um método baseado na utilização de feixes estreitos de energia acústica (som) transmitidos por um equipamento transdutor capaz de converter energia elétrica em acústica.

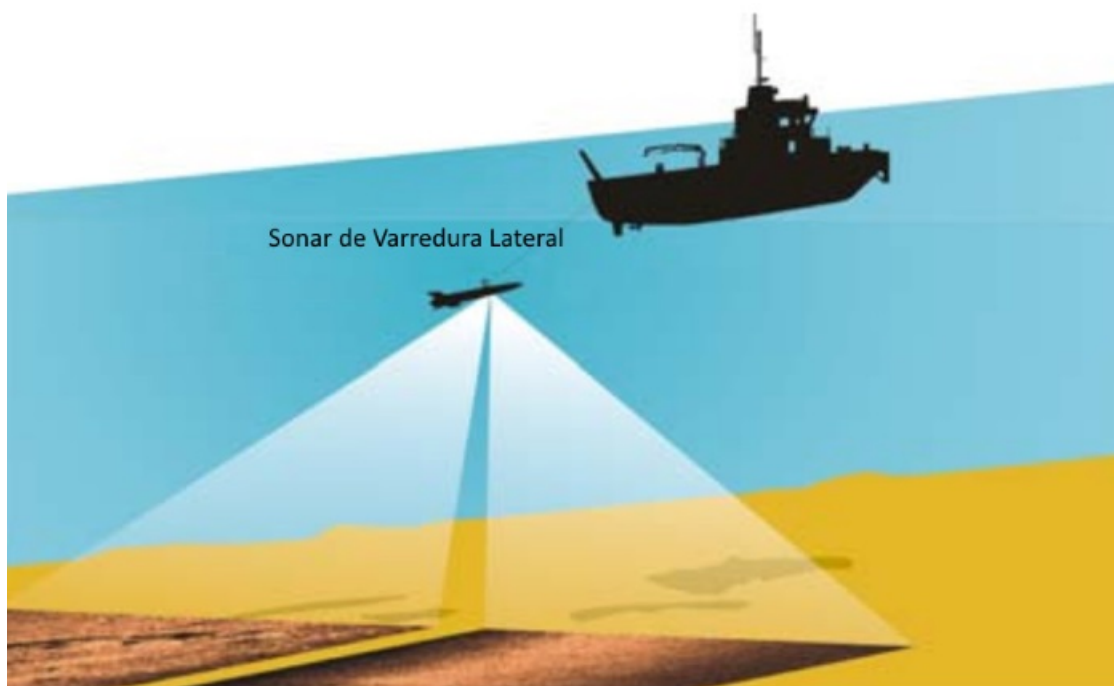


Figura 2.22: Ilustração do funcionamento do sonar de varredura lateral - SVL. Adaptado: sidescan sonar, Wessex University

O princípio é simples: um transdutor emite um pulso acústico através da coluna d'água, ao encontrar um alvo parte do pulso é refletido de volta. O tempo entre o envio e recebimento do pulso acústico é proporcional a profundidade do alvo ([Kunzendorf, 1986](#)).

De maneira menos abstrata, este transdutor converte os sinais elétricos gerados pelo equipamento acústico em uma onda mecânica (pulso acústico) que viaja em direção ao fundo identificando ecos em seu caminho. Estes ecos são o som refletido que retorna ao transdutor, onde é recebido e reconvertido para sinal elétrico que será então processado pelo equipamento acústico de maneira a formar imagens de acordo com a topografia do fundo.

Conforme a variação topográfica, a escala da imagem é determinada pelo tempo entre o envio e o recebimento do sinal acústico pelo transdutor, que transmite uma sequência de pulsos continuamente um após o outro. A imagem é gerada com base nessas variações linha após linha, onde objetos rígidos refletem mais pulsos, ocasionando imagens mais escuras. Enquanto objetos flexíveis, que absorvem mais a energia acústica não refletem grande quantidade de pulsos, ocasionando imagens mais claras ([Bulla et al., 2003](#)), conforme exemplificado na figura [2.23](#).

As imagens resultantes desta metodologia correspondem aos contrastes texturais da superfície do fundo marinho. A forte correlação entre os contrastes texturais observados nestas imagens esta diretamente relacionada com a rugosidade e dureza dos materiais que constituem o substrato ensonificado.

No entanto, o sistema acústico de construção de imagens é sujeito a uma série de desafios referentes ao processo de ensonificação. Como o equipamento opera sobre uma sequência de pulsos que são transmitidos em determinada frequência, estas devem ser previamente configuradas de acordo com a área que se deseja mapear.

Altas frequências geram imagens com maior resolução, mas também, são mais sensíveis a problemas como oscilação do equipamento, partículas suspensas na coluna d'água

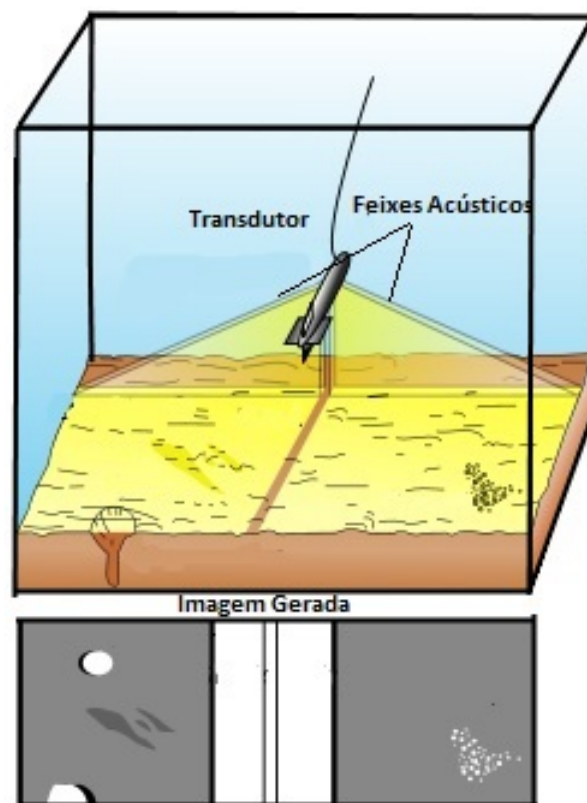


Figura 2.23: Exemplicação imageamento acústico por sidescan sonar. Adaptado: oceanexplorer.noaa.gov

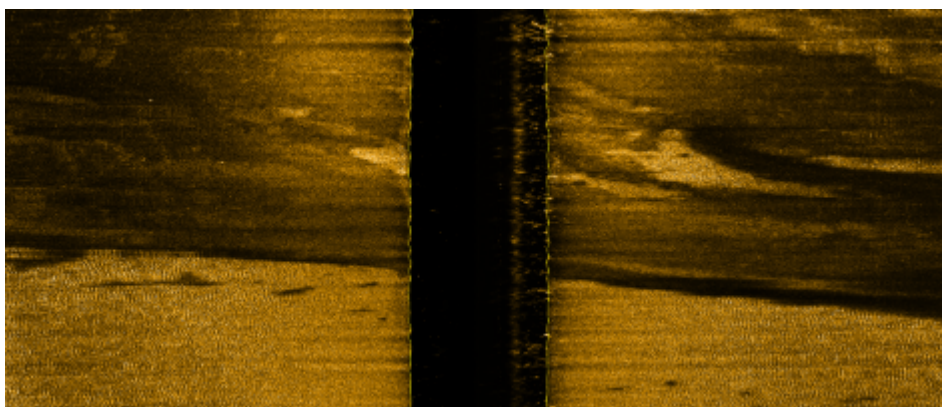


Figura 2.24: Imagem de alta resolução, obtida a partir de um sonar de varredura lateral. Fonte: edgetech.com

e a distância entre o transdutor e o fundo. Logo, por consequência durante a aquisição podem ser geradas texturas artificiais que prejudicam a interpretação destas imagens.

Nas figuras apresentadas é possível observar exemplos de imagens acústicas (sonografias) de alta resolução, obtidas a partir de sonares de varredura lateral. Nas quais, é possível observar não apenas a heterogeneidade dos sedimentos presentes no substrato (Figura 2.24) mas também o alto detalhamento de uma embarcação naufragada na qual é facilmente observável aspectos estruturais e geométricos da estrutura assim como, informações referentes a altura da embarcação com base em sua sombra acústica proveniente da ausência de ensonificação na área (Figura 2.25).

2.3.3 Aplicações

Existem inúmeras aplicações em que se atribui a utilização de sistemas hidroacústicos como: navegação, monitoramento de comunidades biológicas, atividades pesqueiras, localização de objetos submersos, mapeamento superficial, pesquisa por recursos fósseis presentes em grandes profundidades abaixo do leito marinho (sísmica), entre outros.

São exemplos:

1. Detecção e quantificação da abundância de recursos vivos marinhos.

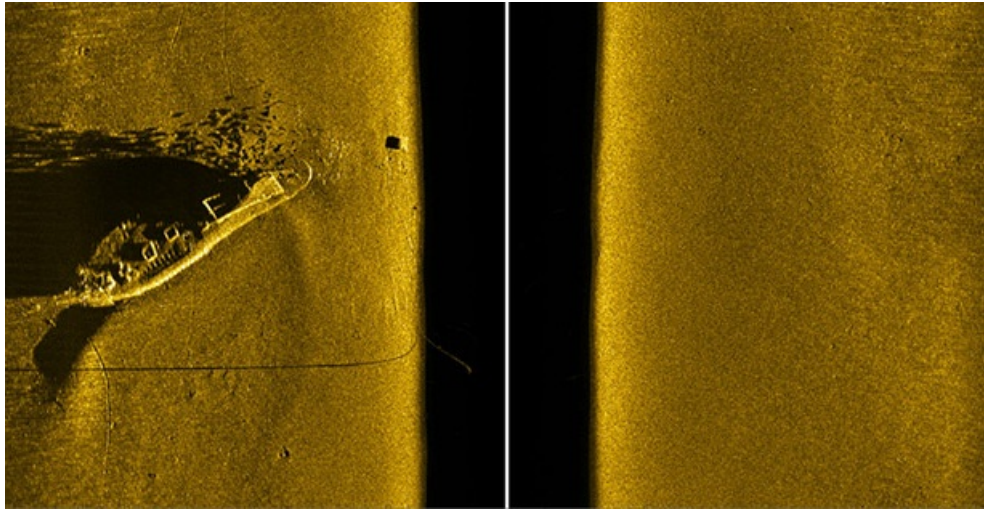


Figura 2.25: Imagem de alta resolução, obtida a partir de um sonar de varredura lateral.

2. Exploração de recursos minerais em camadas abaixo do leito oceânico.
3. Monitoramento de baleias, através do reconhecimento de padrões acústicos biológicos.
4. Navegação e detecção marítima e submarina.

2.4 Sabedoria das Massas

2.4.1 Introdução

De acordo com ([Surowiecki, 2005](#)), a teoria da sabedoria das massas, do inglês, *Wisdom of Crowds*, observa que o conhecimento que emerge a partir de uma decisão coletiva é geralmente mais precisa do que a decisão alcançada de maneira individual por qualquer especialista dentre seus membros. Em seu trabalho, ([Mendoza, 2014](#)), discute sobre melhorias de desempenho inferidas com base neste fenômeno, observando um conjunto de tarefas cujo cenário, é caracterizado por qualidades que tendem a fazer uma multidão inteligente, descentralizando a dependência de partes individuais. Resumindo as opiniões coletivas em um veredito coletivo.

Quando a sabedoria das massas é aplicada no contexto de combinar um conjunto de soluções de máquinas distintas visando melhorar o resultado final do sistema, evitando a perda de informação que pode ocorrer na escolha de um determinado estimador em detrimento de outros. Obtermos os sistemas chamados *Ensemble* ou no caso mais específicos em aprendizado de máquina *Ensemble Learning* ou Comitê de máquinas (Dieterich, 2000).

Conforme (Sharkey and Sharkey, 1997), a ideia de aproveitar, ao invés de perder, a informação contida em estimadores imperfeitos, é central para a utilização de máquinas em comitê. Em seu trabalho (Dieterich, 2000) demonstra que a combinação de estimadores exibe melhor generalização e robustez, sendo mais tolerantes a falhas por não utilizar informação de apenas um estimador.

Na literatura, são apresentados diversos trabalhos relacionando reconhecimento de padrões com sabedorias das massas, em seu trabalho, (Welinder et al., 2010) propõe uma metodologia para classificar conjuntos espaços de dados, baseado no reconhecimento de múltiplos autores, (Ozkan and Morency, 2011), propõem um modelo baseado em sabedoria das massas capaz de aprender automaticamente os padrões de protótipos dinâmicos escondidos entre diferentes modelos especialistas, no trabalho proposto por (Tomaso, 2013), é projetado um sistema para classificar padrões de comportamentos em ratos baseado na captura de múltiplos vídeos e na inferência probabilísticas que predizem os verdadeiros rótulos baseados na perícia de especialistas.

Logo, considera-se que tarefas baseadas em reconhecimento de padrões e classificação, valendo-se de sabedoria das massas para combinar diversas inferências de classificadores distintos pode resultar em ganhos em termos de acurácia e robustez.

2.4.2 Metodologia

Basicamente três razões são consideradas para que sistemas baseados em sabedoria das massas tendam a inferir melhores resultados do que sistemas baseados em modelos indi-

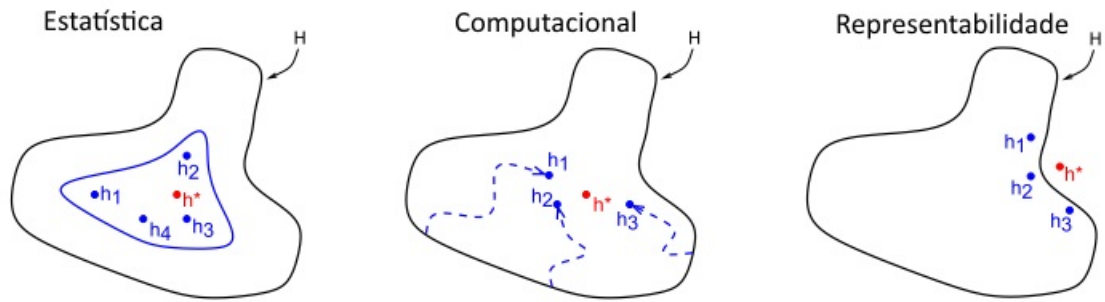


Figura 2.26: Ilustração razões para usar sistemas baseados em sabedoria das massas, sendo H espaço de possibilidades, h^* possível veredito e h_1, h_2, \dots, h_n possibilidades. Adaptado: (Mendoza, 2014)

viduais: estatística, computabilidade e representabilidade (Dietterich, 2000), ilustrado na figura(2.26).

1. **Estatística:** O veredito pode ser visto como pertencente a um conjunto bem definido de hipóteses. A aplicação de diversos vereditos simultâneos e independentes tendem a reduzir a relação da distância entre o conjunto de possibilidades e o veredito correto.
2. **Computacional:** Diferentes algoritmos podem tender a um veredito ótimo local, a união de diversas soluções ótimas locais tende a aproximar uma solução ótima global.
3. **Representabilidade:** Em alguns problemas de classificação pode acontecer de o espaço de classificadores não incluir uma solução ótima, como por exemplo classificadores lineares em problemas não lineares, no entanto, um conjunto de classificadores não-lineares ainda pode convergir para uma solução ótima.

Baseado no conhecimento de cada indivíduo independente, combinar estes conhecimentos de forma a inferir um veredito coerente é a tarefa conhecida como aprendizado em conjunto (*Ensemble Learning*). Diferentes métodos são propostos para combinar um grupo de hipóteses em uma única decisão, conforme (Mendoza, 2014) cita em seu

trabalho:

- **Voto da maioria:** Talvez o mais intuitivo, a decisão é tomada com base nos votos independentes mais frequentemente apresentados, ou seja, se uma determinada classe recebeu $50\% + 1$ votos esta é tomada como a decisão do grupo. Seja N o número de votantes e d um conjunto de possibilidades.

$$\text{Decisão} = \max \sum_{j=1}^N d(i,j) \quad (2.16)$$

- **Voto ponderado da maioria:** Semelhante ao anterior, porém, são atribuídos diferentes níveis de precisão. Neste caso, é possível existir um votante com o poder de decisão. Seja N o número de votantes, d um conjunto de possibilidades e w_i conjunto de pesos.

$$\text{Decisão} = \max \sum_{j=1}^N w_i d(i,j) \quad (2.17)$$

- **Combinação algébrica:** Neste caso a decisão é tomada com base em qualquer função algébrica definida como a média, média ponderada, regra do produto, mediana, entre outras.
- **Combinação Bayesiana:** Assumindo a condição de independência entre os votantes, uma decisão coletiva pode ser considerada baseada em um método de combinatório determinado. Considere d_i a classe a ser predita, w_i conjunto de pesos, L_i o conjunto de classificadores, considerando $P(d_i)$ a probabilidade de L_i rotule uma amostra x como d_i .

$$\text{Decisão}(x) = P(w_k) \prod_{i=1}^N P(d_i|w_k) \quad (2.18)$$

- **Funções de escolha social:** Derivada da teoria da escolha social que tem como foco agregar opiniões e preferências individuais em uma preferência coletiva. Estas preferências podem retratar um conjunto finito e linear de alternativas em um determinado cenário.

2.4.3 Aplicações

As aplicações de sabedoria das massas são encontradas em diversas áreas de estudo nas ciências exatas e nas ciências humanas e sociais. Quando aplicada ao contexto de classificação (também chamada *Ensemble Learning*) tem por objetivo construir uma curva de separabilidade mais refinada através de meta-heurísticas ou modelos matemáticos conforme demonstrado na figura 2.27 aumentando a precisão do processo. Outras aplicações cabe referenciar (Marbach et al., 2012), (Chen et al., 2013), (Velic et al., 2013) e (Mendoza, 2014).

2.5 Trabalhos Relacionados

2.5.1 Introdução

Nas últimas décadas, diversos trabalhos têm sido publicados focando suas atenções na crescente necessidade de classificar e mapear ecossistemas marinhos com escalas espaciais consideráveis (Anderson et al., 2008). Com base no volume de trabalhos publicados na área, é possível observar a evolução e a variedade de modelos teóricos focados em diferentes etapas do processo de classificação. Como principal desafio é observado a necessidade de obter bons descritores capazes de extrair informação útil, não redundante, contida nos dados sonográficos. Assim como, construir sistemas capazes de efetuar a tarefa em tempo computacional satisfatório.

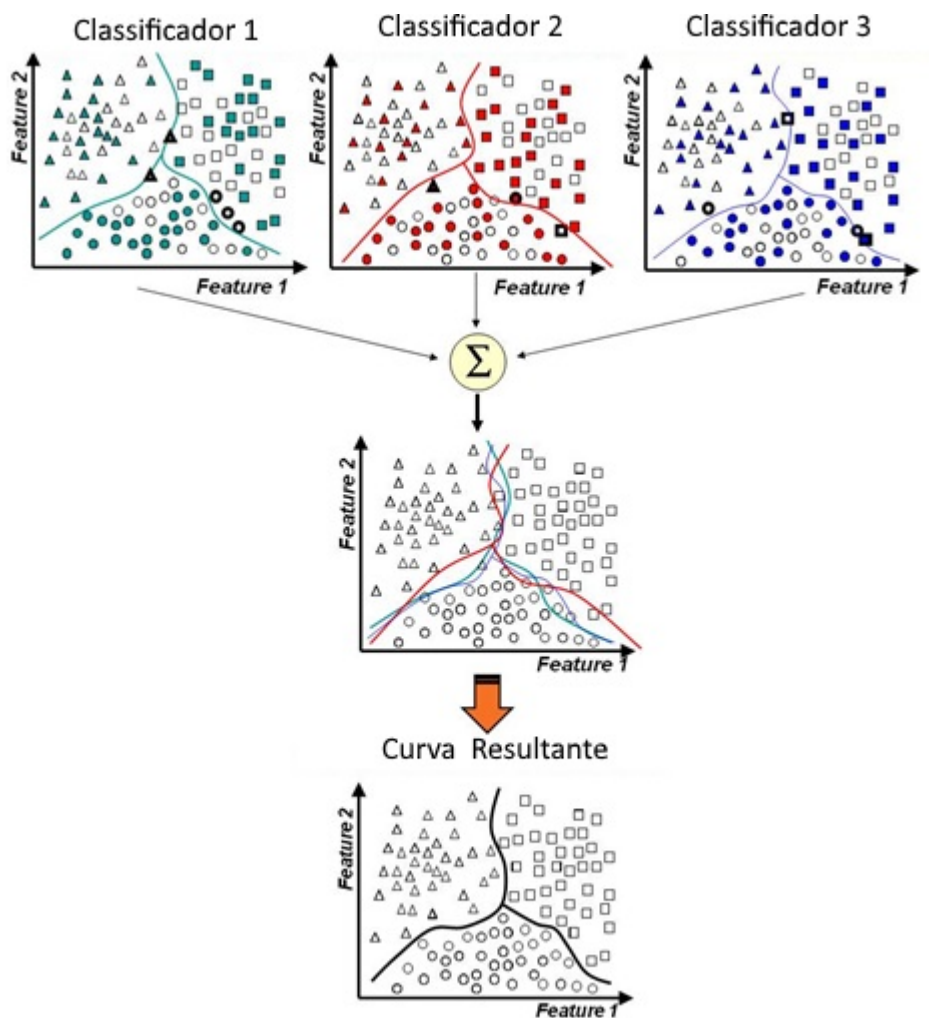


Figura 2.27: Demonstração aplicação de sabedoria das massas em sistemas de classificação. Adaptado:([Polikar, 2008](#))

2.5.2 Trabalhos publicados

Pace and Dyer (1979) descreveram em seu trabalho um método para estudo do leito do mar a partir de imagens de *side scan sonar*, que diminuísse a necessidade de especialistas para interpretar e identificar características geológicas, artificiais, topografia do fundo e tipos de sedimentos. Em seu trabalho, exploraram o potencial do uso de características de texturas para classificar pequenas áreas, utilizando procedimentos de reconhecimento de padrões e tomada de decisões. Como resultados, obtiveram considerável sucesso para pequenos números de características derivadas das variações relativas do sinal acústico, juntamente com uma regra de decisão simples (distância Euclideana).

Linnett et al. (1993) descrevem em seu trabalho o uso de técnicas de análise de textura para extrair informação a partir de sonografias de *side scan sonar*, utilizando o modelo fractal, demonstraram que a obtenção da dimensão fractal não é tão necessária quanto a direcionalidade das texturas, utilizando técnicas baseadas na teoria dos processos espaciais, para extrair valores relativos as variações de níveis de cinza pertencentes as texturas.

Subramaniam et al. (1993) utilizaram a abordagem textural para extrair catorze descritores de textura e reduziram a dimensionalidade do conjunto utilizando a transformada discreta de *Karhunen-loève*, para utilizar o algoritmo de agrupamento K-médias para efetuar a classificação de maneira não supervisionada. Como resultados, os autores notaram que análise de textura era uma técnica poderosa, porém a técnica de classificação não supervisionada era sensível a homogeneidade do leito marinho.

Stewart and Marra (1994) abordaram redes neurais para classificar imagens de *side scan sonar* e compararam medidas de textura em termos de classificação e precisão utilizando redes neurais *backpropagation*. Seus experimentos compararam a eficiência do método em relação a dois métodos estatísticos: Classificador *Fisher-pairwise* e algoritmo *nearest-mean*. Obtendo para três classes de sedimentos, uma acurácia total de 85.9%, 91.2% e 80.1%.

[Carmichael \(1998\)](#) apresentou em seu trabalho técnicas para classificação e segmentação de imagens e *side scan sonar* e para detecção de objetos sobre o leito marinho. Inicialmente, propondo um modelo logarítmico multivariado para caracterizar as propriedades estatísticas de retroespalhamento acústico a partir de diferentes sedimentos. Na sequência para detecção de objetos foi empregado a segmentação em regiões homogêneas, através da busca em cada segmento visando localizar qualquer objeto que correspondesse em forma e tamanho com o *template* definido previamente.

[Blondel et al. \(1998\)](#) apresentaram um método para análise textural de imagens de *side scan sonar*, testando sua eficácia em relação a dados obtidos com auxílio de veículos subaquáticos e amostragens *in-situ*. Para análise textural foi utilizada a abordagem baseada em Matrizes de co-ocorrência de níveis de cinza (GLCMs) conforme descrito por [Haralick and Shanmugan \(1973\)](#), a partir desta é quantificada a relação espacial entre os pixels medindo sua distância e ângulo em relação aos seus vizinhos. Obtendo um conjunto de mais 25 descritores estatísticos a partir das GLCCMs, para o TexAn, foram escolhidos apenas dois descritores (Homogeneidade e Entropia).

[Atallah and Probert Smith \(2004\)](#) descreveram em seu trabalho uma metodologia utilizando dados de amplitude e batimetria fornecidos por sonares de varredura lateral para classificação subaquática. Dessa maneira, adicionando mais uma componente 1D sobre imagens 2D provenientes destes equipamentos, proporcionando maior robustez na ausência de dados porém, tende a ser mais oneroso em termos computacionais. A metodologia, baseia-se em algoritmos supervisionados para a classificação e técnicas de análise de componentes principais para redução da dimensionalidade dos descritores texturais e batimétricos utilizados. Os resultados para os conjuntos de dados utilizados sugere que esta metodologia poderia ter importantes aplicações em mapeamento sedimentológico, inspeção subaquática e levantamentos ambientais.

[Imen et al. \(2005\)](#) apresentaram em seu trabalho um *framework* para avaliar similaridades a partir de estatísticas sobre matrizes de co-ocorrência de níveis de cinza entre

texturas de imagens de sonar de varredura lateral, buscando definir a similaridade entre texturas como sendo baseada na soma ponderada das distâncias entre as distribuições as co-ocorrências de níveis de cinza.

Capus et al. (2008) descreveram uma metodologia de pré-processamento de dados de *side scan sonar* com propósito de visualização, detecção e classificação. Focada nas características físicas e geométricas dos processos de sonar, foi proposto um método que estima fatores de correção separados em intervalos e dependências angulares diretamente das imagens de *side scan sonar*. Uma vez estimados, estes fatores podem ser aplicados sobre grandes conjuntos de dados. O método apresentado requer um bom algoritmo de detecção de fundo para a estimativa de altitude em cada tempo de transmissão e incorpora um esquema de reamostragem para o cálculo e aplicação dos fatores de correção angular-dependência. O método proposto fornece uma solução mais completa do que esquemas de reamostragem reportados e oferece melhorias significativas em termos de precisão, robustez, facilidade de utilização e execução.

Observando a diversidade de descritores e classificadores abordados na literatura, nota-se a importância da elaboração de técnicas que melhor descrevam e diferenciem as classes sedimentológicas, sendo capazes de extrair maior quantidade de informação a partir das sonografias de *sides can sonar* e também a obtenção de classificadores mais eficientes em termos de custo computacional.

Na tabela 2.1, é mostrada uma comparação entre os autores e as técnicas utilizadas para classificação de imagens de *sides can sonar*.

Autor	Descritores	Classificadores
Pace and Dyer (1979)	Matrizes de Co-ocorrência de níveis de cinza.	Agrupamento: distância Euclideana
Linnett et al. (1993)	Modelos Fractais, Características de Texturas, Matrizes de Co-ocorrência de níveis de cinza.	-
Subramaniam et al. (1993)	Matrizes de Co-ocorrência de níveis de cinza.	K-Means
Stewart and Marra (1994)	Estimativa espectral, comprimento de níveis de cinza, Matrizes de Co-ocorrência de níveis de cinza.	Redes Neurais
Carmichael (1998)	Matrizes de Co-ocorrência de níveis de cinza.	Agrupamento
Blondel et al. (1998)	Matrizes de Co-ocorrência de níveis de cinza.	-
Atallah and Probert Smith (2004)	Características de Textura	KNN/K-Means
Imen et al. (2005)	Matrizes de Co-ocorrência de níveis de cinza.	Agrupamento
Fakiris and Papatheodorou (2013)	Texturas locais e reverberação de características.	Agrupamento

Tabela 2.1: Comparação entre os autores e as técnicas utilizadas para classificação de imagens de SSS.

Capítulo 3

Framework Proposto

3.1 Introdução

Em seu trabalho ([Fayad and Schmidt, 1997](#)), definem *framework* (ou arcabouço) como um conjunto de modelos computacionais que colaboram para realizar uma responsabilidade em um domínio de aplicação. Em outras palavras, um *framework* reúne um conjunto de funcionalidades provenientes de diferentes projetos, provendo uma estrutura genérica para solução de problemas específicos.

Neste capítulo, será descrita a metodologia empregada no desenvolvimento do *framework* proposto, explorando e relacionando os modelos computacionais inseridos. As abordagens utilizadas neste trabalho para atingir o objetivo específico de classificação de sonografias de sonar de varredura lateral serão discutidas a medida que são apresentadas em seus respectivos módulos e etapas de processamento.

3.2 Engenharia de Software

Engenharia de software é uma disciplina de engenharia cujo foco está em todos os aspectos da produção de software, desde os estágios iniciais da especificação do sistema

até sua manutenção, quando o sistema já está em uso (Sommerville, 2011).

3.2.1 Engenharia de Requisitos

De acordo com (Pressman, 2006), a parte individual mais difícil da construção de um sistema é decidir o que construir. Nenhuma parte do trabalho danifica tanto o sistema resultante se for feita errado. Nenhuma outra parte é mais difícil de consertar depois. Nesta perspectiva, a engenharia de requisitos fornece mecanismos apropriados para auxiliar a compreender o que o cliente deseja, analisando as necessidades e avaliando a exequibilidade.

Como objeto base para o funcionamento do *framework* proposto neste trabalho, estão as sonografias provenientes dos sonares de varredura lateral estas por sua vez são produto de distintos equipamentos e são sujeitas a uma vasta lista de parâmetros de aquisição e visualização presentes nos softwares de controle destes equipamentos. Estes parâmetros, podem ocasionar alterações diretas sobre a imagem gerada.

Logo, tratando-se de uma ferramenta multidisciplinar a ser usada em um contexto prático, alguns requisitos funcionais e não funcionais foram definidos em conjunto com especialistas da área de emprego do software.

Requisitos Não-Funcionais:

- Sistema deve ser multi-plataforma ou seja, operar sobre diferentes plataformas computacionais como Windows e Linux.
- Possuir uma interface de fácil manipulação e configuração.

Requisitos Funcionais:

- Sistema deve operar sobre imagens extraídas diretamente dos sistemas de aquisição, sem nenhuma etapa de pós-processamento.

- As imagens a serem utilizadas para o treinamento do sistema e para classificação. Devem ser oriundas do mesmo processo de aquisição ou deve ser garantido que nenhum parâmetro de configuração ocasione em modificações nas sonografias.
- As imagens devem ser fornecidas com resolução adequada, de forma a preservar contrastes texturais e de intensidade. O próprio sistema se encarrega de efetuar as devidas transformações necessárias para otimizar o desempenho.
- O conjunto de imagens a serem utilizadas para o treinamento devem ser extraídas de múltiplas sonografias, priorizando a diversidade encontrada na mesma classe sedimentológica.

3.2.2 Engenharia de Sistemas

O processo de engenharia de sistemas tem por princípio que um sistema antes de ser submetido ao processo de desenvolvimento, deve ser entendido e relacionado com o meio no qual ele irá residir e operar, ou seja, suas necessidades técnicas, relação com usuários, conectividade com o meio, entre outros. Logo, todos os elementos relacionados ao sistema devem ser modelados, analisados e especificados a priori, de modo a garantir a estabilidade e confiabilidade do sistema.

Considerando que o emprego do *framework* na prática ocorrerá a bordo de navios de pesquisa, os quais frequentemente não possuem conexão com a internet devido a custos ou inviabilidade técnica, a construção de sistemas web não é uma opção, necessitando o emprego de softwares *stand-alone*, ou seja, autossuficientes em seu funcionamento. Baseado na arquitetura a qual o sistema de imageamento do sonar de varredura lateral opera (Figura 3.1), o *framework* pode ser instalado no computador do equipamento responsável pelo software de gerenciamento do sistema, facilitando sua integração ou aquisição direta das imagens possibilitando a classificação em paralelo com a aquisição.

Baseado nos requisitos discutidos entre outros, o *framework* proposto foi desenvol-



Figura 3.1: Sistema de aquisição de dados do sonar de varredura lateral.

vido inteiramente sobre a linguagem Python¹. Pois, além de características como ser uma linguagem de programação de alto nível, interpretada, imperativa, orientada a objetos, funcional, de tipagem dinâmica e forte, possui uma comunidade muito ativa, estando presente em inúmeras aplicações científicas e comerciais como Youtube, Google, Yahoo, Blender, Gimp, entre outros.

Logo, Python é uma linguagem de propósito geral podendo ser empregada em várias classes de problemas e por se tratar de uma linguagem de *scripting*, torna-se uma linguagem altamente produtiva e com desempenho satisfatório. Programas desenvolvidos em python pode ser distribuídos como aplicações *Stand-Alone* com a utilização de ferramentas como o *Py2exe*² sem que o usuário necessite instalar interpretadores ou efetuar quaisquer configurações adicionais, facilitando seu emprego.

¹<https://wiki.python.org/moin/>

²<http://www.py2exe.org/>

3.2.3 Engenharia de Projeto

O processo de engenharia de projeto é iniciado assim que as primeiras interações entre engenharia de requisitos e engenharia de sistemas são concluídas. O objetivo geral é modelar e estabelecer um conjunto de princípios que irá levar ao desenvolvimento de um software de qualidade em nível estrutural, abordando conceitos fundamentais que irão fornecer a estrutura básica necessária para a construção correta do produto.

Fundamentado na abordagem estatística supervisionada, a arquitetura base para classificação inserida neste *framework* contém as etapas gerais padronizadas de qualquer sistema de classificação supervisionado (ASM, Shihavuddin, 2014), conforme ilustrado na figura 3.2.

Como metodologia principal para classificação de sedimentos baseada em sonografias de sonar de varredura lateral, este trabalho apresenta uma nova abordagem que visa proporcionar maior índice de acurácia e desempenho computacional empregando técnicas baseadas em teoria das massas e super segmentação, discutidas em suas respectivas seções. Também, visando proporcionar ao usuário a possibilidade de inferir resultados personalizados de acordo com seu objetivo específico é oferecido um ferramental diversificado em todas as etapas do processo. A figura 3.3, ilustra a sequência lógica de execução da metodologia de classificação proposta, considere "?" como pontos de tomada de decisão.

3.3 Seleção de Amostras Treinamento

Fundamental para qualquer sistema de classificação supervisionado, é estabelecer um conjunto de amostras previamente definidas que melhor discriminem as classes que se pretende mapear. Considerando que o conjunto de amostras de treinamento é definido e selecionado pelo especialista anteriormente as etapas de treinamento e classificação, e a partir destes serão extraídas as características que serão consideradas como descri-

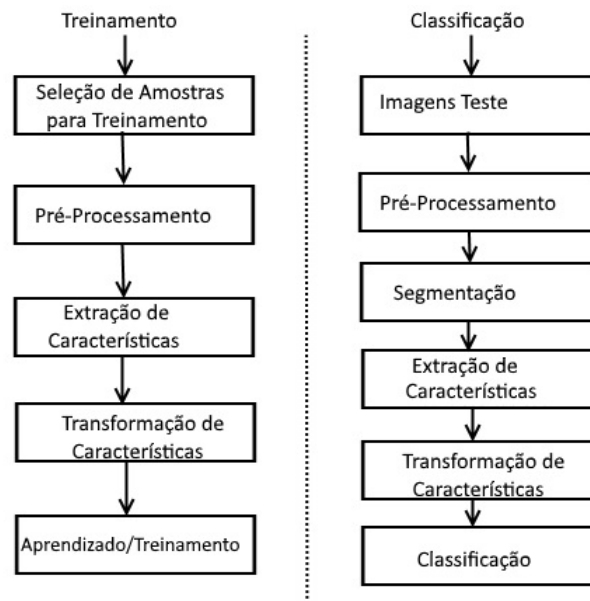


Figura 3.2: Etapas processo de classificação supervisionada de objetos em imagens.

tores pertinentes para cada classe, são necessárias algumas considerações especiais no momento de selecionar o melhor conjunto de amostras.

Para este trabalho foi definida a região do Parcel do Carpinteiro na plataforma continental interna a nordeste da Barra do Rio Grande-RS, como área de estudo por se tratar de uma região com maior heterogeneidade sedimentológica. Foram efetuados quatro perfis de mapeamento acústicos com o sonar de varredura lateral e três coletas *in situ* visando amostrar as três classes sedimentológicas visualizadas nas sonografias de acordo com a intensidade de retorno (Figura 3.5).

As amostras coletadas *in situ* com um amostrador do tipo *Box-Corer*, foram classificadas em laboratório como sendo a mancha clara referente a areia fina, mancha intermediária como areia condensada com lama e a macha mais escura como lama biodetrítica (Figura 3.4).

Em relação a escolha das amostras para classificação a partir das sonografias, estas devem ser extraídas da maior quantidade possível de sonografias (Figura 3.6), visando

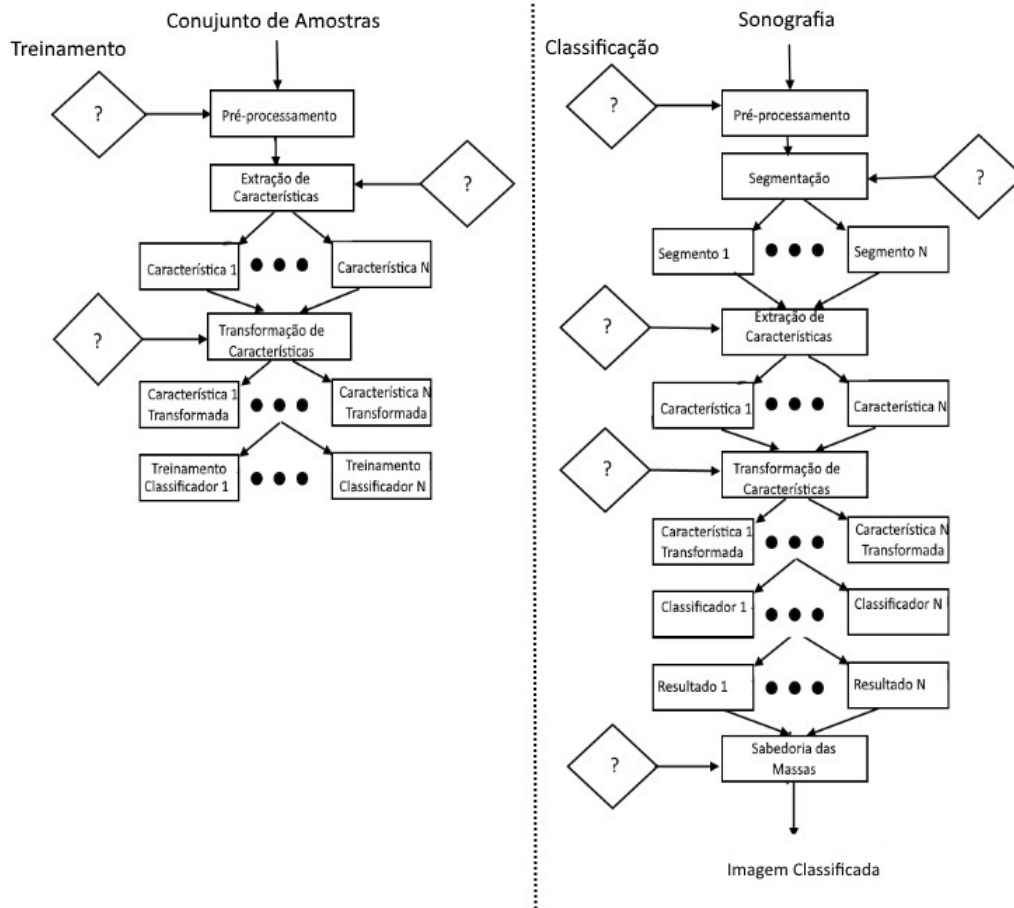


Figura 3.3: Sequência lógica de execução da metodologia de classificação proposta, considere "?" como pontos de tomada de decisão.



Figura 3.4: Amostras sedimentológicas coletadas no parcel do carpinteiro, visando relacionar com as sonografias adquiridas, classificadas em laboratório como sendo a mancha clara referente a areia fina, mancha intermediária como areia condensada com lama e a macha mais escura como lama biodetrítica.

proporcionar ao classificador a extração de características de um conjunto heterogêneo de amostras, evitando a super aprendizagem que por consequência aumenta o nível de especificidade do classificador. No entanto, o processo de escolha das amostras deve ser cuidadoso evitando rotular classes de maneira equivocada, de modo a não inserir ruído no conjunto de amostras resultando em um nível de aprendizado satisfatório.

Considerando que a maioria dos extratores de características inseridos operam no domínio espacial, uma vez definida as dimensões das amostras como “ 32×32 ” ou “ 64×64 ” pixels todas as amostras devem ser extraídas com as mesmas dimensões. Uma vez definidas as amostras e efetuado o treinamento, o classificador será independente e classificará as sonografias com bases no aprendizado, no entanto, uma vez que seja necessária a classificação em ambientes diversificados com classes novas o sistema deverá passar pela etapa de treinamento. Sendo aconselhável o treinamento sempre que forem utilizadas novos conjuntos de imagens, como um processo de calibração do sistema.

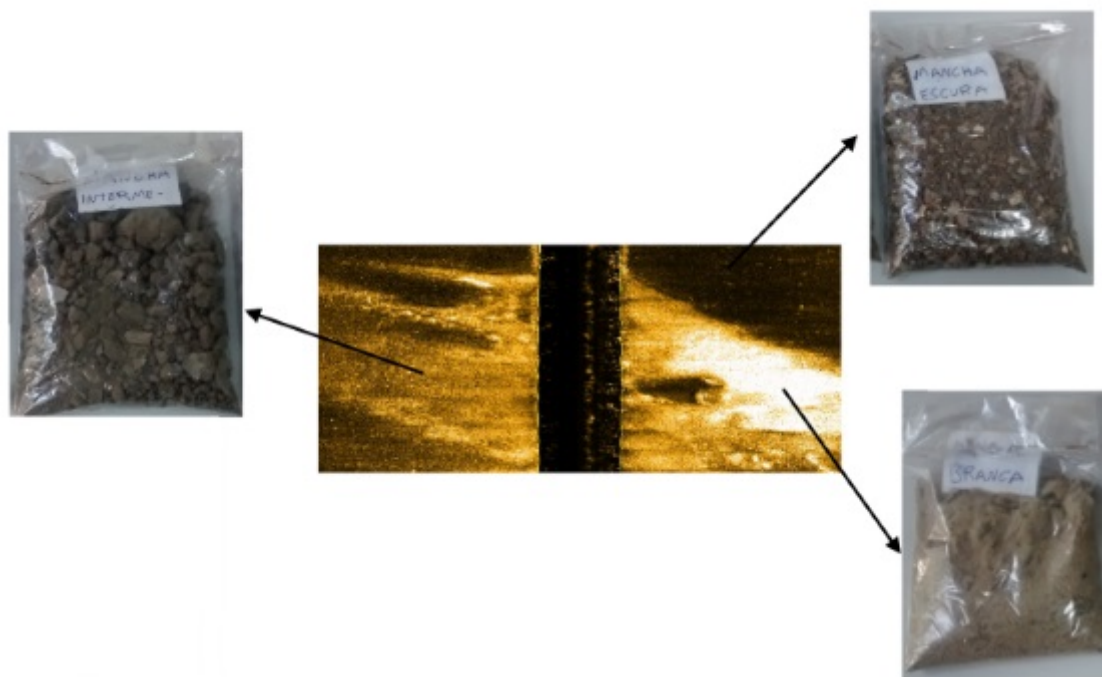


Figura 3.5: Relação amostras *in situ* com o imageamento acústico.

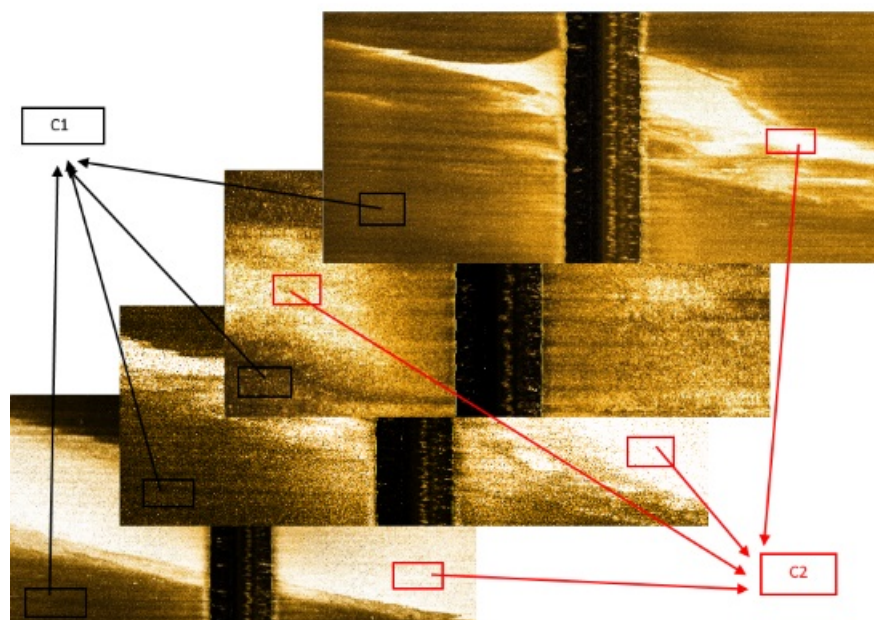


Figura 3.6: Ilustração processo de seleção conjunto de amostras para treino.

3.4 Pré-Processamento

Previamente a utilização dos conjuntos de imagens para treinamento e classificação são necessários ajustes que visem melhorar a qualidade das imagens realçando aspectos pertinentes a análise ou reduzindo ruídos que podem ser definidos como variáveis aleatórias caracterizadas por uma função de densidade de probabilidade. No contexto deste trabalho, visando ampliar o nível de separabilidade entre as distintas classes sedimentares contidas nas sonografias são considerados ruídos distintos aspectos presentes nas mesmas sendo alguns naturais como a formação de ondulações (*ripples*) sobre a superfície ocasionadas por fenômenos hidrodinâmicos e outros originados do meio acústico de aquisição, que necessitam de condições adequadas para aquisição dos dados como a estabilidade do *town-fish* do sonar de varredura lateral.

A figura 3.12, evidencia os dois tipos de estruturas prejudiciais mais comuns encontrados nestas imagens, no painel (A) ruído do tipo impulso (sal e pimenta) ocasionado pela sobreposição superficial dos sedimentos que resulta em ecos de retorno com intensidade oscilante na linha de aquisição que constituem cada linha da imagem (Figura 3.7), e o painel (B) gerado pela oscilação do *town-Fish* que resulta em linhas não ensonificadas. A figura 3.14, evidencia a presença de *ripples* ocasionados por fenômenos hidrodinâmicos que alteram a textura da região.

Os sonares de varredura lateral usualmente são multi-frequência operando em pelo menos duas frequências uma alta 500KHz ou mais e uma baixa de até 100KHz, imagens resultantes de altas frequências possuem resolução mais apurada no entanto, são mais sujeitas a problemas ocasionados durante a aquisição dos dados resultando em falhas na imagem, no contexto de classificação de sedimentos imagens, a utilização de imagens de baixa frequência se mostra mais eficaz sendo menos sujeitas a falhas e possuindo um alcance maior, conforme demonstrado na Figura 3.8.

Baseado em imagens de baixa frequência e no sentido de tratar diferentes tipos de

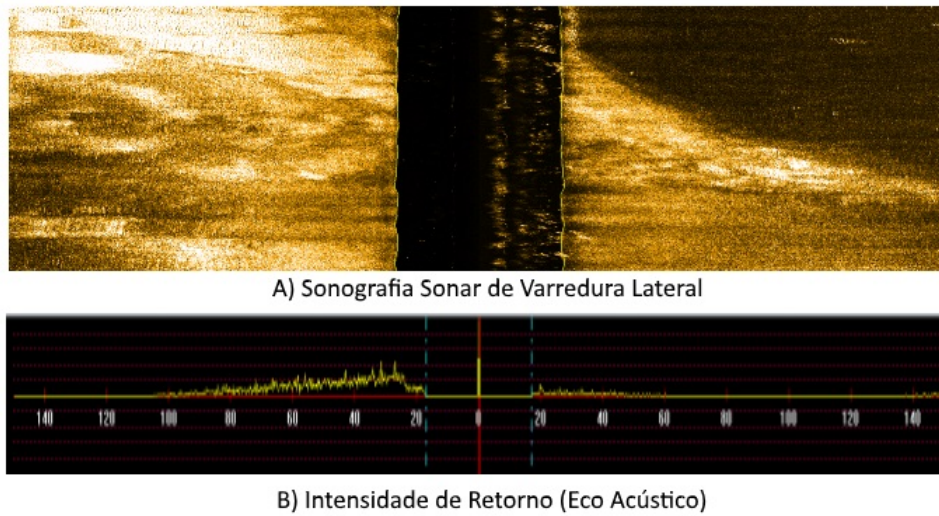


Figura 3.7: Sobreposição superficial dos sedimentos que resulta em ecos de retorno com intensidade oscilante na linha de aquisição que constituem cada linha da imagem.

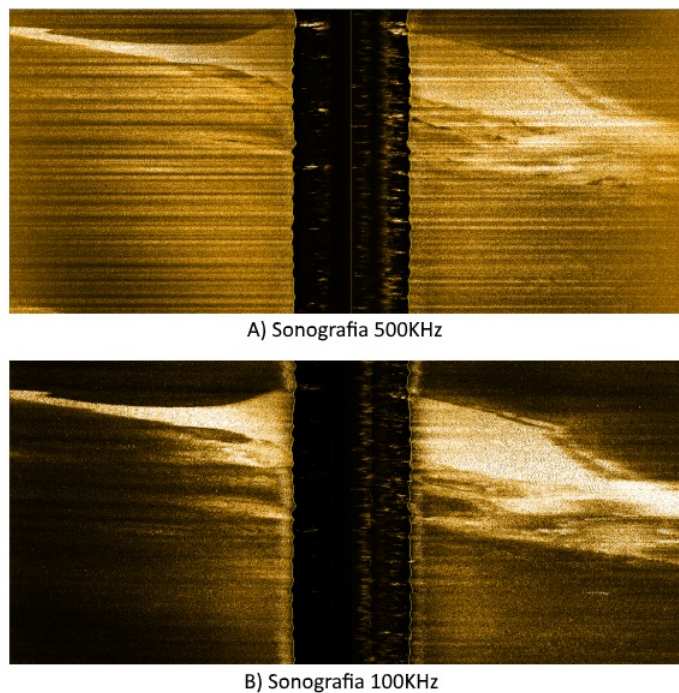


Figura 3.8: Demonstração de sonografias de alta e baixa frequência, 500KHz e 100KHz respectivamente

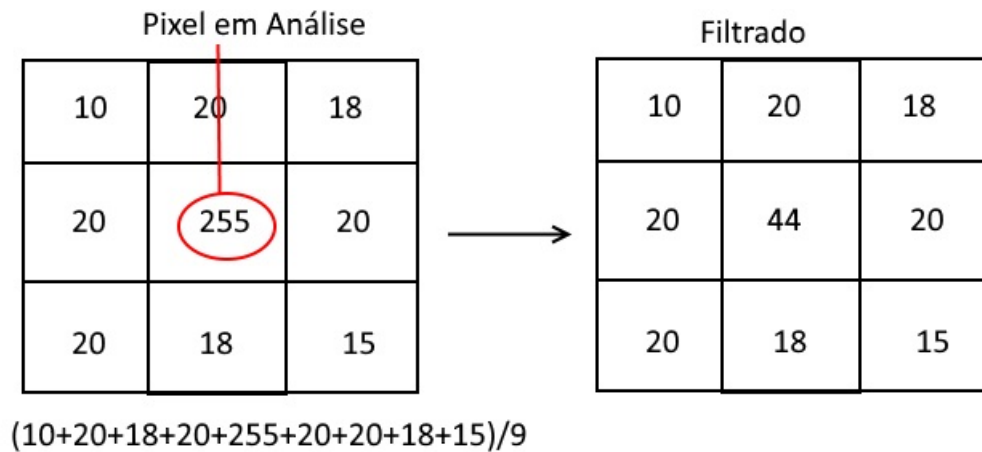


Figura 3.9: aplicação filtro média

ruídos, o *framework* proposto disponibiliza um conjunto filtros conforme listados a seguir, que podem ser utilizados individualmente ou em conjunto conforme a necessidade do usuário. No entanto, uma opção de filtragem pré-calibrada é disponibilizada como parte da metodologia inserida neste trabalho. A metodologia de filtragem estabelecida, visa subtrair o ruído do tipo impulso (sal e pimenta) comumente presente em sonografias de sonar de varredura lateral e ressaltar as variabilidade de contrastes entre as classes sedimentares presentes na imagem. Na figura 3.13, é demonstrada uma sonografia em escala de tons de cinza, antes e após a etapa de pré-processamento.

- **Filtro Média**, substitui cada pixel pela média aritmética da vizinhança, a quantidade de vizinhos é dado pelas dimensões da máscara de convolução. Ilustrado na figura 3.9.
- **Filtro Média Bilateral**, é um filtro não linear com o intuito de redução de ruídos e preservação de bordas. Onde cada pixel central é substituído pela média ponderada dos vizinhos. Os pesos dos vizinhos podem ser baseados em uma distribuição Gaussiana.
- **Filtro Gaussiano**, podendo ser utilizado como filtro passa-baixa, usa uma função

$\frac{1}{273}$	1	4	7	4	1
	4	16	26	16	4
	7	26	41	26	7
	4	16	26	16	4
	1	4	7	4	1

Figura 3.10: Máscara 5x5 da função Gaussiana aproximada com $\sigma = 1.0$

Gaussiana para obter valores de uma máscara a ser definida digitalmente. O filtro Gaussiano $2D$ pode ser definido como:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (3.1)$$

Onde x e y são as coordenadas da imagens e σ é o desvio padrão.

Considerando que a imagem é uma coleção de pixels discretos é possível produzir uma forma discreta da distribuição Gaussiana para obter uma máscara de convolução. A figura 3.10, mostra a máscara 5×5 da Gaussiana para $\sigma = 1.0$

A aplicação do filtro Gaussiano gera uma suavização (*smoothing, blur*) na imagem de forma semelhante ao filtro média, no entanto, por não possuir todos os pesos iguais como o filtro média, sua suavização é mais delicada preservando contornos. O resultado será mais suave a medida que σ cresce.

- **Filtro Mediana**, muito utilizado para suavizar ruídos do tipo impulsivo, sendo capaz de reduzir o ruído causando baixo nível de suavização. Nesta transformação, cada pixel central é substituído pelo nível de cinza central do conjunto ordenado

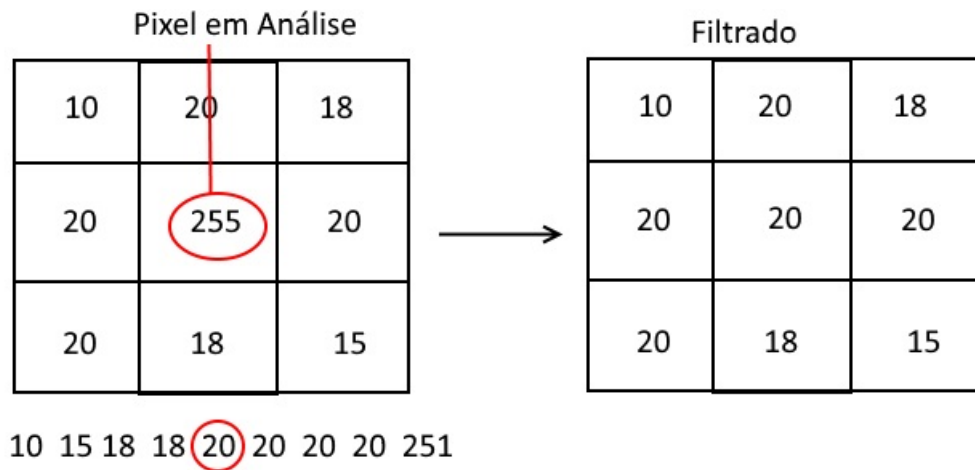


Figura 3.11: Aplicação filtro mediana

dos níveis de cinza da vizinhança. A figura 3.11, ilustra a operação do filtro mediana.

- **Filtro Contraste**, filtro linear que busca enfatizar a variações locais de níveis de cinza. Com o objetivo de aumentar a discriminação visual entre os os objetos presentes na imagem.

Neste contexto, é observado na bibliografia diversificadas alternativas que apresentam vantagens e desvantagens para tratar determinados tipos de ruídos. Nesta proposta, mediante experimentos foi verificado que a utilização do filtro mediana e do filtro contraste, respectivamente demonstraram os melhores resultados, considerando que a principal estrutura prejudicial encontrada e considerada neste trabalho como ruído em sonografias é o ruído do tipo impulso (*salt and pepper*). Problemas ocasionados por fatores hidrodinâmicos que alteram a textura de classes leves como areia fina como *ripples*, foram tratados mediante o aumento e a generalização do conjunto de treinamento.

Por fim, observa-se a relevância da etapa de pré-processamento, considerando esta fundamental para que seja possível a extração de bons descritores de características que

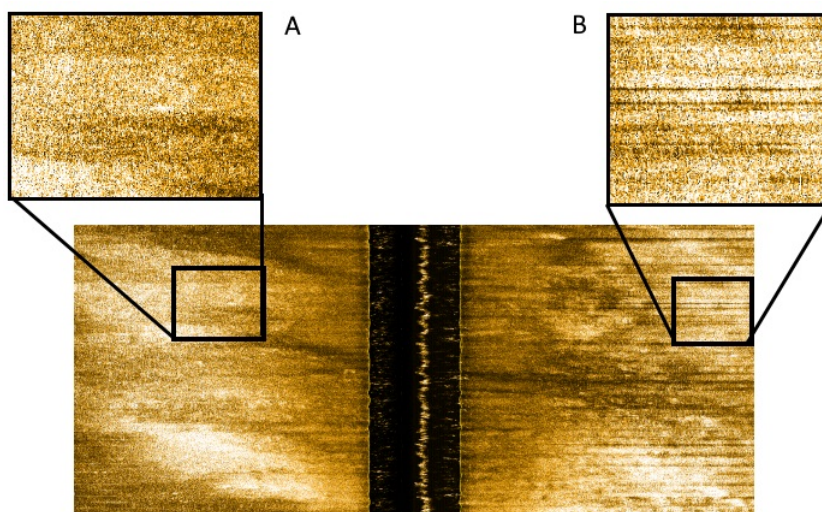


Figura 3.12: Tipos de ruídos mais comuns encontrados nestas imagens, no painel (A) ocasionado pela sobreposição superficial dos sedimentos que resulta em ecos de retorno com intensidade oscilante na linha de aquisição (que constituem cada linha da imagem), e no painel (B) gerado pela oscilação do *town-Fish* que resulta em linhas não ensonificadas.

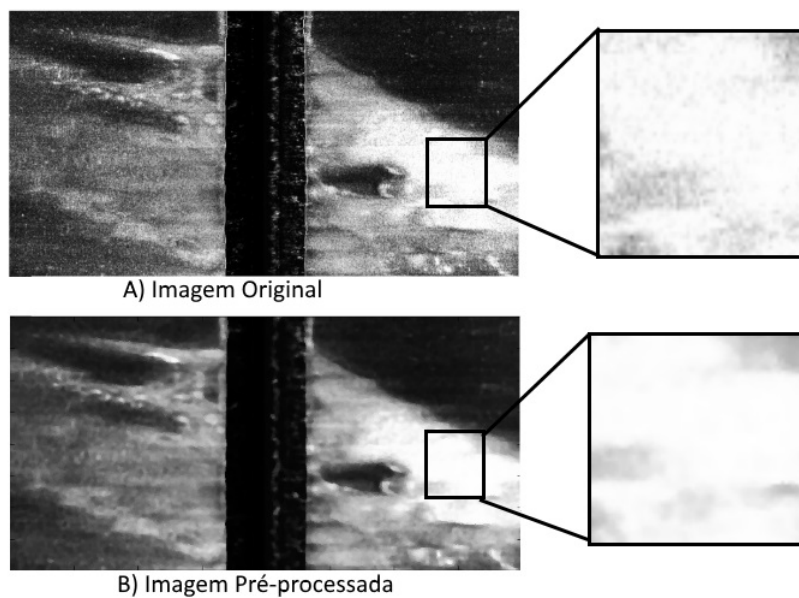


Figura 3.13: Sonografia em escala de cinza antes e após a etapa de pré-processamento.

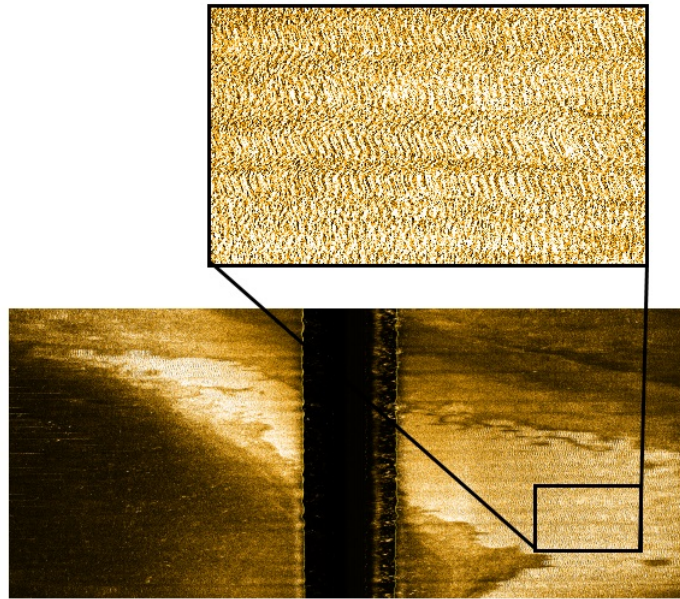


Figura 3.14: Demonstração da presença de *ripples* ocasionados por fenômenos hidrodinâmicos que alteram a textura da região.

sejam capazes de discriminar de maneira confiável diferentes classes sedimentológicas. Estes por sua vez, dado o volume de trabalhos sobre descritores de características, podem ser considerados o tópico mais importante do processo pois na ausência de características que diferenciem objetos estes tendem a pertencer a mesma classe.

3.5 Extração de Características

Quando o objetivo é a classificação automática de imagens de sonar varredura lateral, inúmeros trabalhos são encontrados na literatura visando estimar os melhores descritores capazes de diferenciar de maneira efetiva as classes sedimentológicas presentes nas mesmas, como discutido no capítulo 2.5.2. No entanto, o emprego destes descritores não é garantia de níveis adequados de sucesso, pois, os mesmos são sensíveis a condições anômalas que nem sempre são eliminadas nas etapas de pré-processamento.

Descritores baseados em textura são os mais utilizados em diversos sistemas para

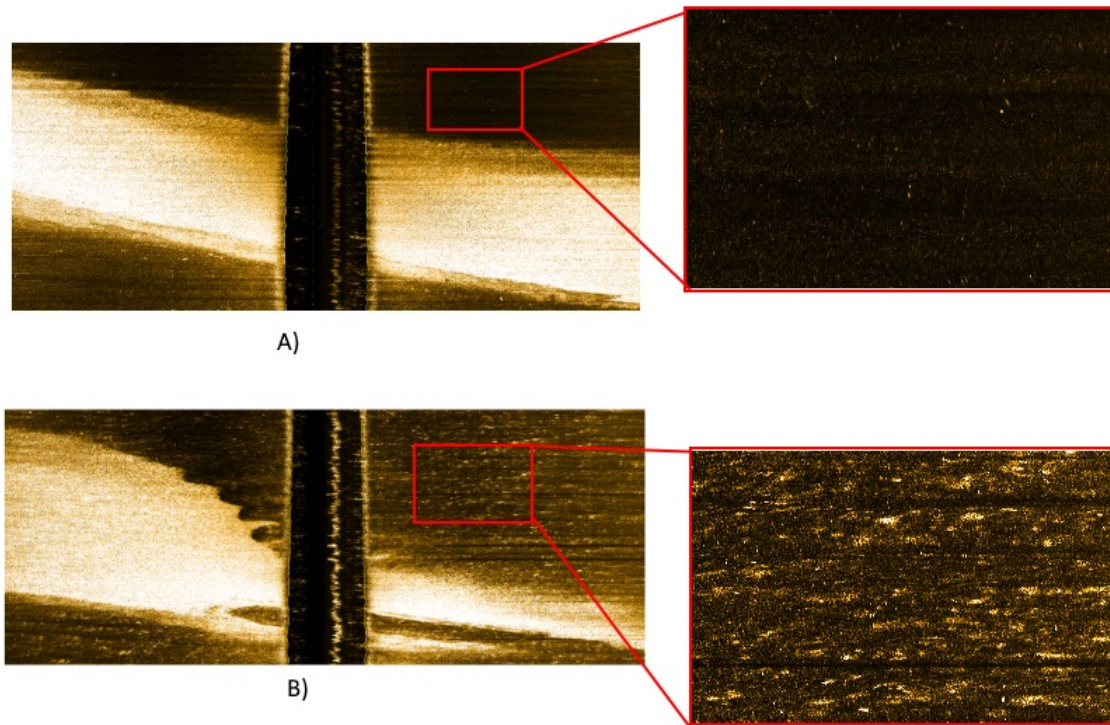


Figura 3.15: Sonografias evidenciando mesma classe sedimentológica com variabilidade textural.

classificação baseados em imagens. Tratando-se da classificação de sonografias, os descritores de textura apresentados por (Haralick and Shanmugan, 1973) são os mais utilizados, porém, isoladamente descritores de textura baseados em matrizes de co-ocorrência de cinza são sujeitos a situações nas quais determinada classe sedimentológica apresenta texturas completamente diferentes devido diferentes fatores naturais como compactação ou hidrodinâmica. A figura 3.15, demonstra a classe sedimentológica classificada como lama biodretrítica, é possível observar que a mesma classe sedimentológica apresenta respostas diferentes resultando em variabilidade textural.

No entanto, a união de descritores que visem características variadas tende a reduzir a especificidade do classificador, na figura 3.16 é possível observar a mesma ocorrência de variabilidade textural sobre outra classe. O emprego de descritores baseados na in-

tensidade do pixel ocasiona uma nova perspectiva para o classificador, neste sentido, a ideia básica da metodologia empregada neste *framework* é a construção de uma imagem resultante em função da utilização de múltiplos descritores resultando em visões especializadas e diferenciadas de classificação.

O *framework* apresentado contempla as seguintes classes de descritores e seus respectivos conjuntos de características:

- **Estatísticas Básicas:** Média, variância, obliquidade, curtose, entropia e moda. Como indicadores de variabilidade de rugosidade do sedimento inferidos sobre a intensidade do conjunto de *pixels*.
- **Estatísticas baseadas em histogramas:** Buscando mensurar a distribuição do retroespalhamento acústico, inferindo a Média, variância, obliquidade, curtose, entropia e moda diretamente sobre o vetor unidimensional referente a distribuição de níveis de cinza contidos na imagem.
- **Matrizes de co-ocorrência de níveis de cinza:** Estatísticas sobre a relação dos *pixels* e seus vizinhos em diferentes direções, são extraídos vinte e quatro atributos estatísticos mas utilizados seis considerados mais representativos para descrever texturas de acordo com ([Haralick and Shanmugan, 1973](#)) sendo estes momento angular, média, variância, entropia, correlação e obliquidade.
- **Potência espectral:** Relaciona a razão de energia contida dentro das diferentes bandas de frequências, observando-se, por exemplo, que fundos rochosos apresentam maior energia concentrada nas baixas frequências.
- **Dimensões Fractais:** Considera que as características da distribuição do retroespalhamento acústico são similares a medida que a imagem é vista em diferentes escalas.

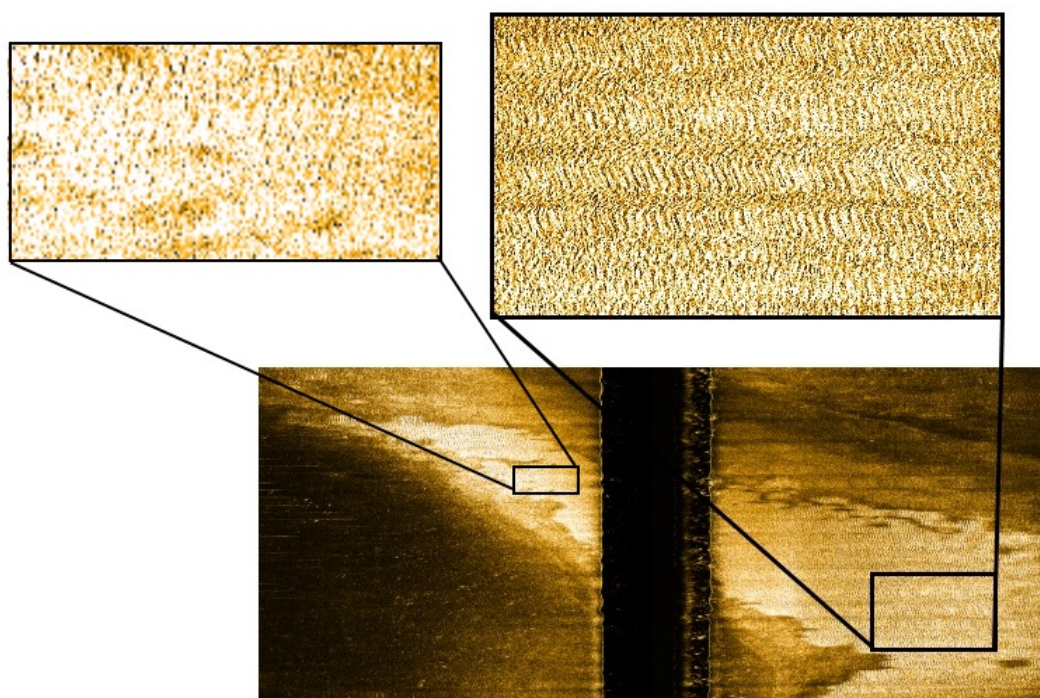


Figura 3.16: Sonografias (2) evidenciando mesma classe sedimentológica com variabilidade textural.

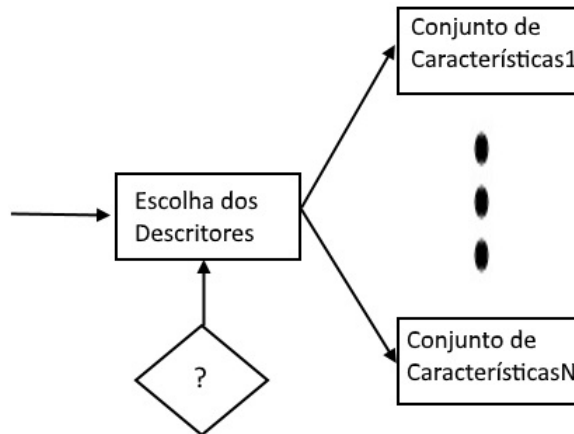


Figura 3.17: Seleção e extração de descritores, cada classe de descritores fornece um conjunto de características que serão utilizados pelos classificadores durante a diferenciação e predição.

Por consequência são extraídos N conjuntos de parâmetros de características de acordo com o número de descritores selecionado. Conforme exemplificado na figura 3.17.

3.6 Modificação de Características

Utilizando todos os descritores apresentados são construídos conjuntos de características com dimensões elevadas. No entanto, dimensões elevadas não garantem níveis elevados de discriminabilidade, pelo contrário, geram ruídos que interferem na qualidade do conjunto e reduzem o desempenho do sistema. Conforme demonstrado no capítulo 4.1, a união destes diferentes descritores tende a produzir resultados mais precisos aproveitando-se da especificidade de cada classe de descritores.

No entanto, como mencionado anteriormente, a utilização de muitas classes de descritores ocasiona um número elevado de características extraídas, como por exemplo os descritores de histogramas que proporcionam seis atributos estatísticos, matrizes de co-ocorrência vinte e quatro atributos. Logo, a utilização de espaços amplos de características além de não prover ganhos em termos de eficácia compromete muito o

desempenho do processo.

Neste sentido, existe a necessidade de reduzir a dimensão do espaço de características em um espaço reduzido que contemple apenas atributos significativos. Neste intuito é empregado o uso da análise de componentes principais transformando um conjunto de alta dimensionalidade de atributos possivelmente relacionados em um conjunto reduzido de atributos linearmente descorrelacionados, conforme apresentado na seção 2.2.2.

3.7 Treinamento

Etapa fundamental da classificação supervisionada, o treinamento visa o ajuste de um modelo de similaridade a partir dos descritores contidos no conjunto de treinamento. Este modelo de similaridade será utilizado durante a classificação como delimitador no momento de rotular corretamente o novo objeto a ser classificado, logo o conjunto de treinamento deve ser cuidadosamente selecionado de maneira a tornar o aprendizado coerente e genérico.

Neste trabalho, o conjunto de treinamento consiste em selecionar amostras de 32×32 ou 64×64 pixels das diferentes classes sedimentológicas a ser classificadas. Estas amostras são extraídas diretamente do conjunto de imagens a ser classificado, porém devem ser extraídas amostras genéricas de cada sedimento ou seja, amostras que apresentem variabilidade na mesma classe evitando que aconteça o sobre-ajuste (*Overfitting*) visando evitar que o modelo se ajuste em demasia ao conjunto amostral.

Na metodologia aqui proposta, tratando-se da classificação de sonografias de sonar de varredura lateral o treinamento pode utilizado para contornar problemas referentes a variações na textura de sedimentos leves causados por fatores hidrodinâmicos como os *ripples*. A inserção de amostras com a presença destes fatores diferenciados de textura confere ao sistema um nível de generalização sem prejudicar a classificação, já que a metodologia é baseada na composição de diferentes classes de descritores.

Em relação ao tamanho do conjunto amostral, neste trabalho foi experimentado diferentes configurações como 20, 50, 100 e 300 amostras de cada sedimento, no entanto, devido ao número reduzido de classes sedimentares (três) as quais havia disponibilidade para esse trabalho, conjuntos acima de 50 amostras de treinamento não apresentaram ganhos significativos e consomem tempo considerável no processo de aprendizado.

No entanto, acredita-se que com o aumento do número de classes a serem classificadas, haja a necessidade do aumento no conjunto amostral principalmente para classes sedimentares próximas ou mistas, como areia e areia lamosa ou lama e lama-biodetrítica. Neste sentido, futuramente com a inserção de novas classes novas avaliações e considerações deveram ser feitas em relação a seleção e a proporção de amostras visando aferir modelos mais precisos de diferenciação.

Após os conjuntos de características terem sido devidamente transformados, os mesmos são disponibilizados para a etapa de treinamento. Durante esta etapa são disponibilizados nove classificadores supervisionados que podem ser utilizados individualmente conjunto através da aplicação de sabedoria das massas, conforme demonstrado na figura 3.19.

Neste trabalho, foram abordados e disponibilizados os seguintes algoritmos de classificação, cada qual pré-calibrado para classificação de sonografias de sonar de varredura lateral:

- Support Vector Machine Classifier (SVM) : O algoritmo SVM é baseado na construção de hiperplanos de separabilidade é um classificador versátil pois permite utilização de diferentes funções de decisão (*Kernel Functions*).
- K-nearest Neighbors Classifier (KNN) : O algoritmo KNN tem por principio encontrar a qual classe de treinamento o novo ponto encontra-se mais próximo. Esta distância pode ser qualquer métrica como a distância Euclideana.
- Decision Tree Classifier (DT): Algoritmos baseados em árvores de decisão como o

J48 constroem modelos de predição baseados nos atributos de entrada tomados como hipóteses. Aproximando a curva de predição baseado em conjuntos de regras de decisão.

- Decision Tree Regressive Classifier (DTR): Semelhante a algoritmo anterior porém especializado em extrair modelos de predição, através de regressões lineares locais.
- Random Forest Classifier (RF): O algoritmo Random Forest é baseado em *ensemble learning* ou seja, seu resultado é composto da união dos resultados obtidos de múltiplas árvores de decisão.
- AdaBoost Classifier (AB): É um classificador meta-heurístico baseado e *ensemble learning*, utilizado para aumentar a performance de outros algoritmos de classificação compondo sua predição a partir do resultado de múltiplos classificadores.
- Linear Discriminant Analysis Classifier (LDA): É um método estatístico que busca encontrar uma combinação linear de características capaz de caracterizar a separabilidade entre duas ou mais classes, o resultado desta combinação geralmente é utilizado para redução de dimensionalidade antes da classificação.
- Quadratic Discriminant Analysis Classifier (QDA): É uma especialização do método LDA, concentrado em separar duas ou mais classes através de superfícies quadráticas.
- Logistic Regression Classifier (RL): É um classificador estatístico que possui por objetivo produzir, a partir de um conjunto de observações, um modelo que permita a predição de valores tomados por uma variável, frequentemente binária, a partir de uma série de variáveis explicativas contínuas e/ou binárias.

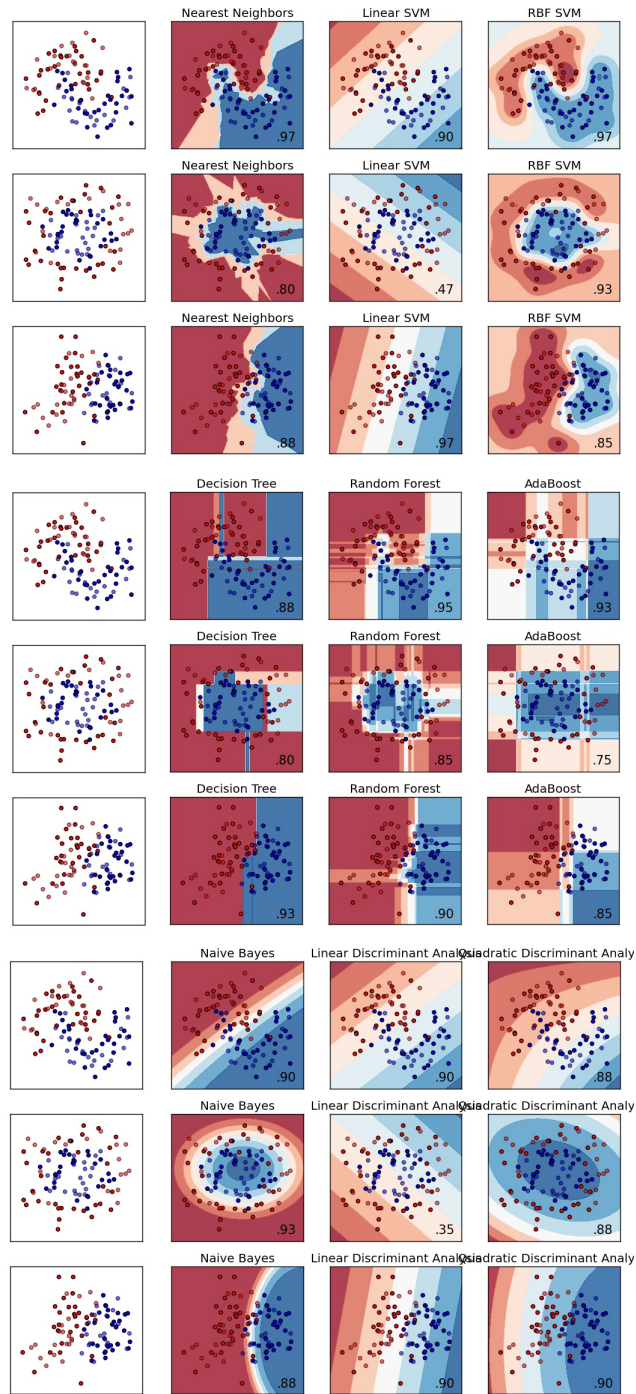


Figura 3.18: Ilustração comparação entre classificadores. Adaptado: Scikitlearn

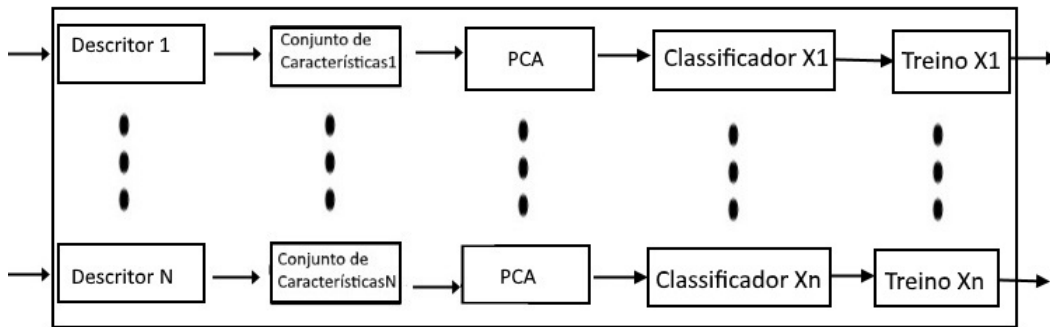


Figura 3.19: Metodologia de treinamento proposta, a partir da seleção de um número de classificadores superior a um a metodologia aplica a sabedoria das massas para unir os distintos preditores.

3.8 Segmentação x Convolução

Anteriormente a etapa de classificação, a imagem a ser classificada deve ser devidamente subdividida em regiões de interesse de maneira a extrair características destas regiões. Usualmente, é observada a utilização de matrizes de dimensões $N \times N$ que são deslocadas sobre a imagem original, assumindo que a região delimitada pela matriz possui características comuns (Figura 3.20). No entanto, a estrutura rígida desta representação está sujeita a situações como a captura de regiões limítrofes entre classes, como ilustrado na figura 3.21 resultando em características extraídas de maneira inconsistente.

Em relação ao desempenho a utilização da convolução, está relacionada com dois aspectos as dimensões da matriz de convolução e o passo de deslocamento, ou seja, quantos pixels a matriz ignora entre seus deslocamentos. Na situação ideal, a matriz seria unitária cobrindo apenas 1 pixel e se deslocaria com passo 1 de pixel em pixel. No entanto, essa abordagem é impraticável em relação ao tempo para imagens de dimensões elevadas e não seria capaz de extrair características discriminantes de um conjunto unitário de pixels sabendo que os descritores baseiam-se nos pixels e na sua vizinhança.

Usualmente é concebida a utilização de matrizes e deslocamentos maiores, resul-

tando em níveis de desempenho computacional mais consideráveis. Resultando não apenas em níveis baixos de precisão no processo de classificação, assim como, resultados com bordas serrilhadas devido a estrutura rígida da janela de convolução.

Em termos assintóticos a complexidade computacional inerente da técnica de convolução pode ser inferida considerando o deslocamento de uma matriz de dimensões $m \times m$ sobre uma imagem n de dimensões $N \times N$ e considerando o passo de deslocamento k , como:

$$F(Fn) = \frac{1}{k} \times (N \times N) \times (m \times m) \quad (3.2)$$

No pior caso, deslocando um matriz de convolução $m \times m$ sobre a imagem " i " de dimensões $N \times N$ com passo de deslocamento $k = 1$ temos:

$$O(N^2 \times m^2) \quad (3.3)$$

A utilização de métodos de segmentação baseados em superpixels, proporcionam ganhos consideráveis em relação ao desempenho. Sabendo que um superpixel é construído a partir do agrupamento de pixels em blocos (*clusters*) que compartilham propriedades semelhantes, conforme demonstrado na figura 3.22. O Resultado é um conjunto reduzido de regiões a serem manipuladas nos processos subsequentes e problemas como de regiões limítrofes discutidos na convolução, não ocorrem já que estes contem apenas pixels com propriedades semelhantes.

Em termos assintóticos a complexidade computacional inerente da técnica de superpixels baseada no algoritmo de agrupamento (*K-means*) pode ser inferida considerando como entrada uma imagem " n " de proporções $N \times M$ a qual são construídas " k " regiões retangulares (superpixels), cada superpixel " S " é constituído por

$$S = \frac{(N \times M)}{k} \quad (3.4)$$

pixels e seja " N_i " o número de iterações para realocar os grupos e $P_{i,j}$ o pixel da posição " i,j " temos:

$$F(Fn) = N_i \times k \times S_k P_{i,j} \quad (3.5)$$

operando sobre N_i e K constantes obtemos:

$$F(Fn) = C \times C \times S_k P_{i,j} \quad (3.6)$$

$$F(Fn) = C \times S_k P_{i,j} \quad (3.7)$$

$$F(Fn) = S_k P_{i,j} = n \quad (3.8)$$

$$F(Fn) = n \quad (3.9)$$

onde " Fn " é conjunto total de pixels presentes na imagem, Logo:

$$O(n) \quad (3.10)$$

O *framework* apresentado, disponibiliza a utilização das duas metodologias para segmentação das imagens a serem classificadas. No entanto, os resultados obtidos com o processo de segmentação são mais eficientes, resultando em tempo computacional consideravelmente reduzido e proporcionando a criação de mosaicos de classificação mais delicados, respeitando a complexidade geométrica presente na distribuição dos sedimentos sobre o fundo.

3.9 Classificação

Na etapa de classificação os segmentos da imagem são rotulados em uma das classes sedimentológicas previamente aprendidas através da etapa de treinamento. A metodo-

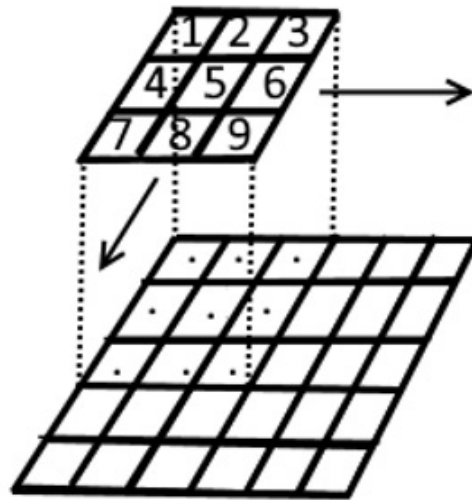


Figura 3.20: Ilustração de segmentação por convolução.

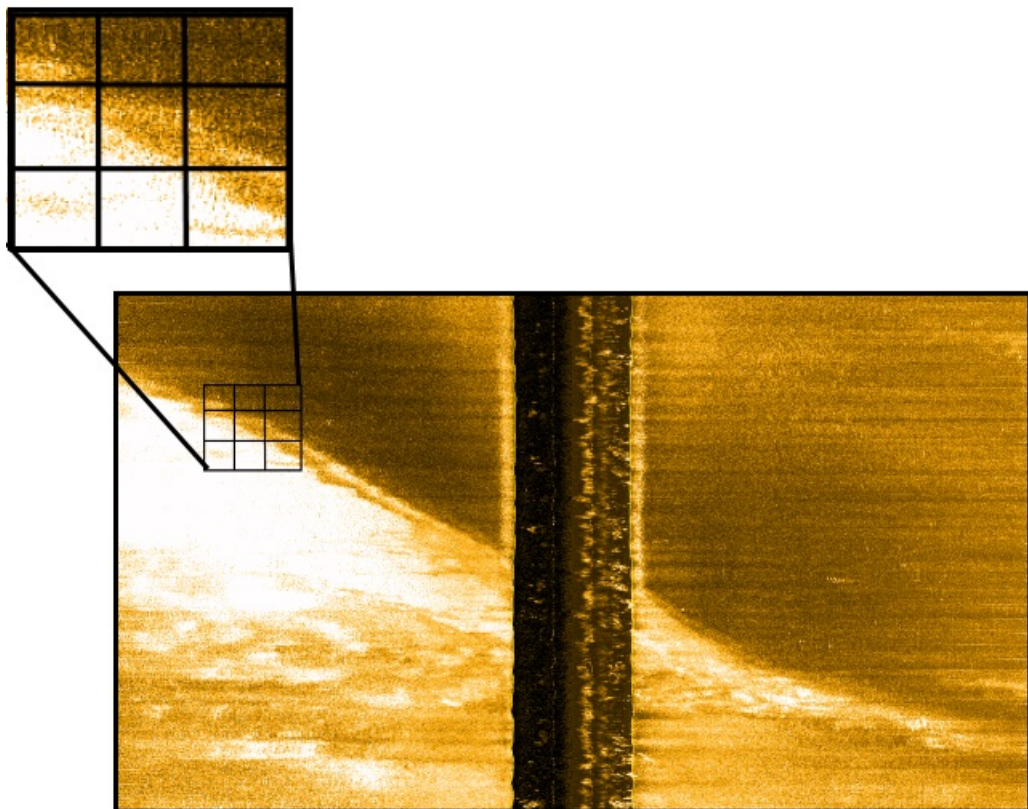


Figura 3.21: Ilustração convolução sobre regiões limítrofes entre classes.

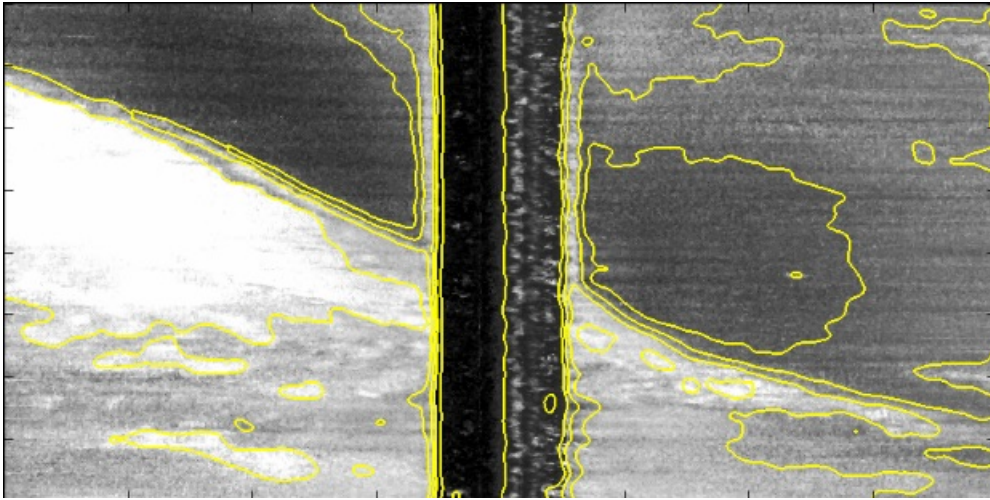


Figura 3.22: Demonstração da segmentação utilizando superpixels.

logia de classificação apresentada neste trabalho, consiste na utilização da sabedoria das massas (ou *ensemble learning*) para unir o conjunto de soluções oriundas da utilização de diferentes descritores e/ou classificadores em uma única solução final, resultando em ganhos de precisão. Em outras palavras, no centro da metodologia proposta a sonografia devidamente classificada é o resultado do processo de interpolação entre diferentes configurações de descritores e classificadores, visando compor uma solução global resultante da união de soluções especialistas locais.

A interpolação é feita a partir de diversas imagens resultantes de classificadores baseados em descritores especializados, ou seja, são produzidas múltiplas imagens em paralelo resultantes da classificação com uma única classe de descritores. Na sequência estes resultados são unidos pelo voto da maioria compondo uma solução global em termos de descritores, e no caso da utilização de múltiplos classificadores diferentes ao final é feito um segundo ponto de união entre os resultados dos múltiplos classificadores que por sua vez são resultantes da união de múltiplos descritores.

O processo geral de classificação proposta neste trabalho pode ser ilustrado conforme demonstrado na figura 3.23, no qual é possível observar a utilização de dois pontos

de sabedoria das massas, o primeiro para unir os resultados oriundos da utilização de classes de descritores distintas e o segundo para unir as predições dos distintos classificadores.

Esta metodologia foi planejada visando proporcionar ao sistema, um nível maior de abstração, pois, compõe uma única solução global a partir de especialistas locais. No entanto, a utilização de múltiplos descritores e classificadores ocasiona perdas não apenas em relação ao desempenho, mas também, em relação ao consumo de recursos computacionais, pois, trabalha-se com múltiplas imagens. Estas perdas são compensadas com a utilização de técnicas computacionais que visam o ganho de desempenho como a utilização transformadas espaciais nas imagens originais, técnicas de super segmentação e o paralelismo de diferentes etapas do processo.

O paralelismo conforme demonstrado na figura 3.24 é inserido em todos os módulos, uma vez que os múltiplos classificadores são independentes entre si, até o momento da composição da solução no segundo ponto de sabedoria das massas. A utilização da técnica de superpixel além de preservar a complexidade geométrica dos objetos na imagem, efetua a operação de separação dos *clusters* (superpixels) em tempo linear evitando a necessidade de paralelismo neste ponto. Ao utilizarmos a segmentação por janela de convolução (*Sliding Window*) o tempo computacional elevado dado a natureza assintótica cúbica da metodologia a qual pode ser reduzida a linearidade valendo-se de paralelismo em nível de GPU (*Graphic Processing Unit*).

Logo, a etapa de construção do resultado final (*Ensemble*) é o resultado da aplicação das técnicas de sabedoria das massas quando o sistema é configurado com múltiplos descritores visando aprimorar o classificador a partir de resultados distintos centrados em aspectos como textura, luminosidade, energia de espectro entre outros. A união dessas abordagens resulta em uma imagem de classificação mais precisa contemplando todos os aspectos descritos.

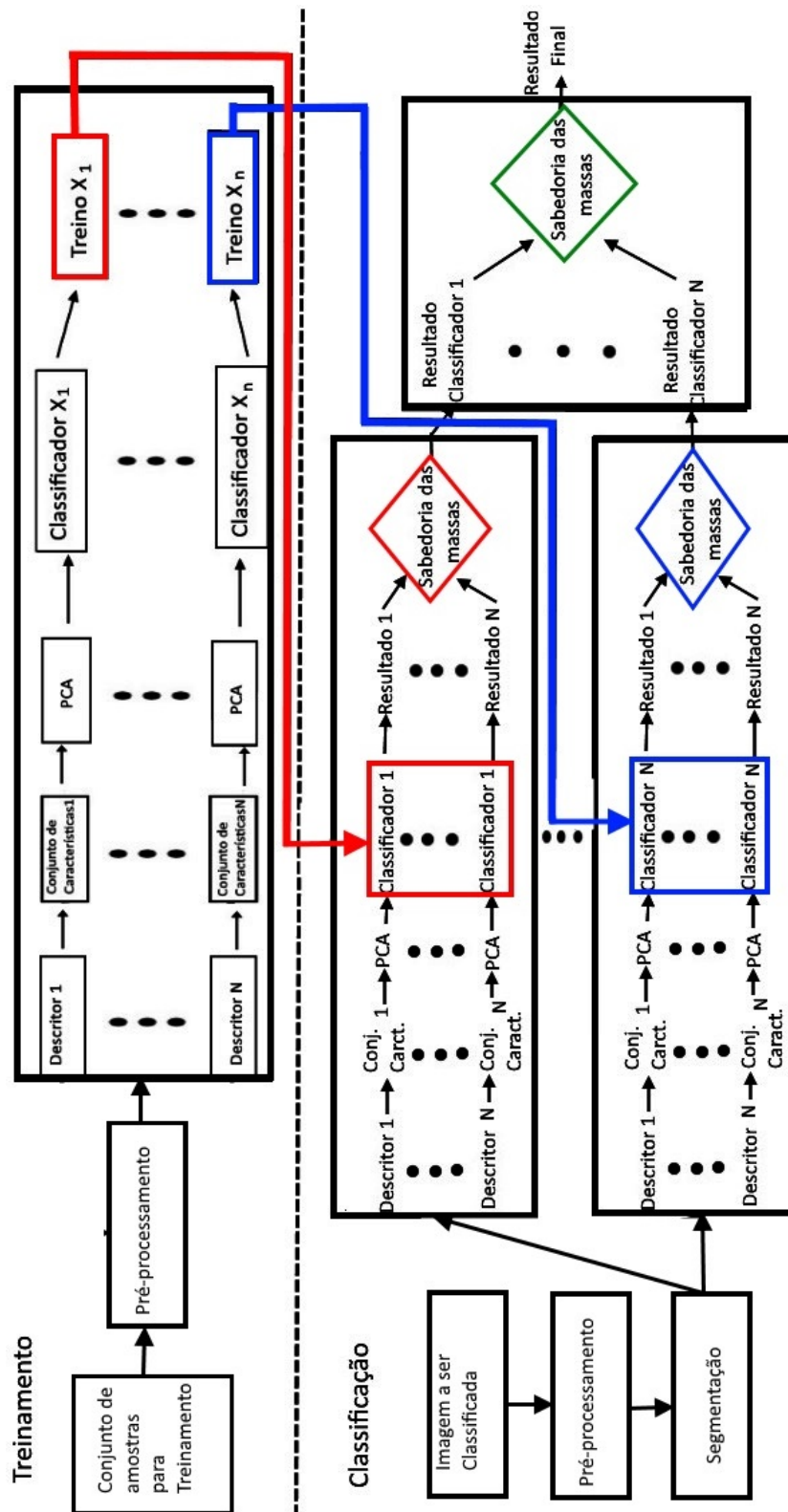


Figura 3.23: Ilustração do processo de classificação proposto.

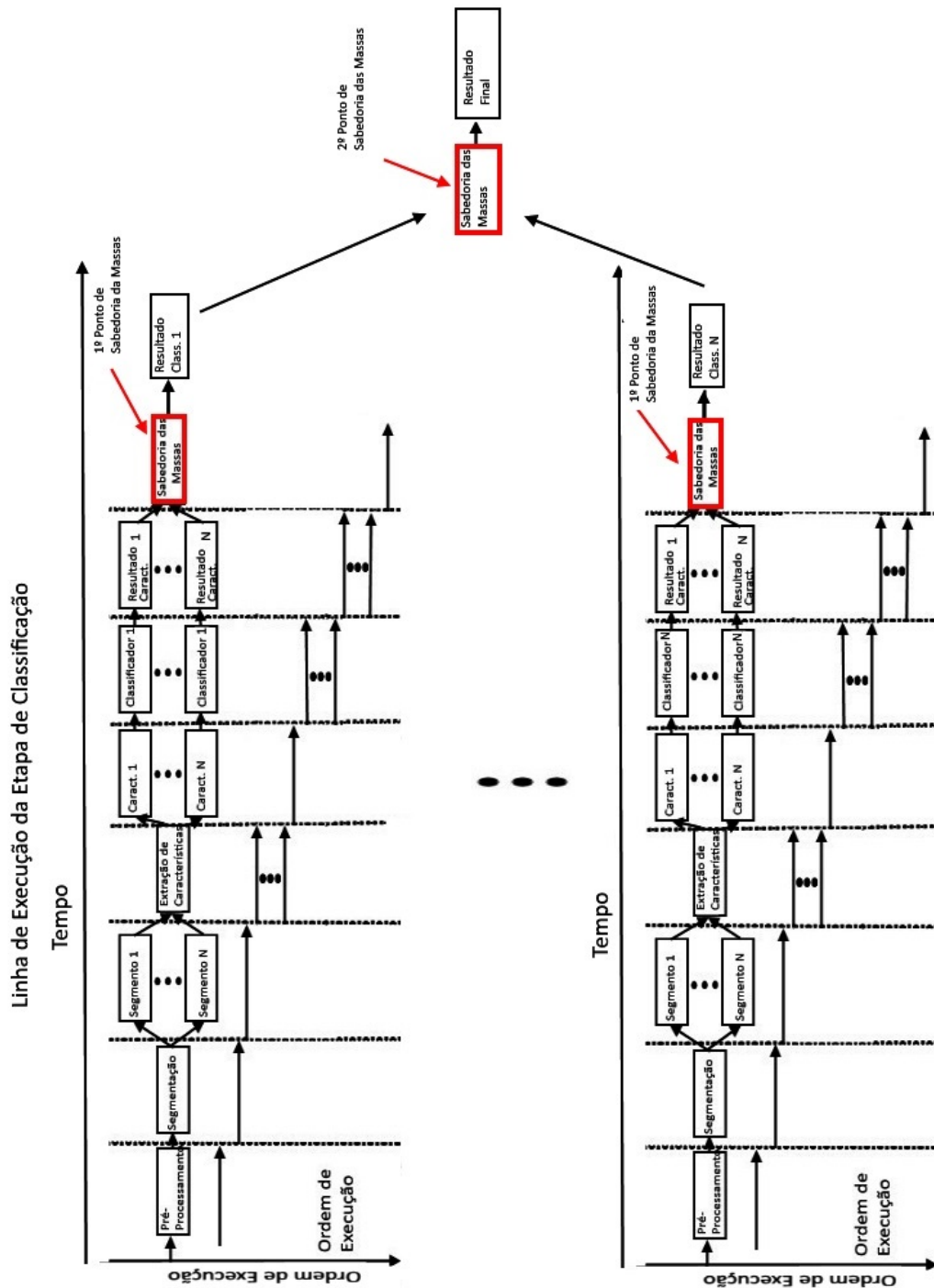


Figura 3.24: Ilustração das etapas de paralelismo durante a linha de execução do processo de classificação.

Capítulo 4

Resultados

4.1 Introdução

Neste capítulo será apresentado o *framework* desenvolvido neste trabalho, assim como, os resultados atingidos por meio da metodologia proposta, estes resultados serão apresentados de maneira qualitativa através da apresentação das sonografias classificadas e também de maneira quantitativa, explorando os resultados numéricos obtidos valendo-se de diferentes classes de descritores e diferentes tipos de algoritmos de classificação. A discussão dos resultados de maneira mais profunda será abordada no próximo capítulo [5](#).

4.2 O *Framework*

Nesta seção, será apresentado o *framework* computacional desenvolvido neste trabalho, assim como os seus módulos conforme discutido no [3](#).

Na figura [4.1](#) é apresentado o módulo de pré-processamento, este módulo é responsável pelos passos iniciais em direção a classificação. Nele é contido um conjunto de filtros (1) para imagens digitais visando a correção de ruídos e o realce de características

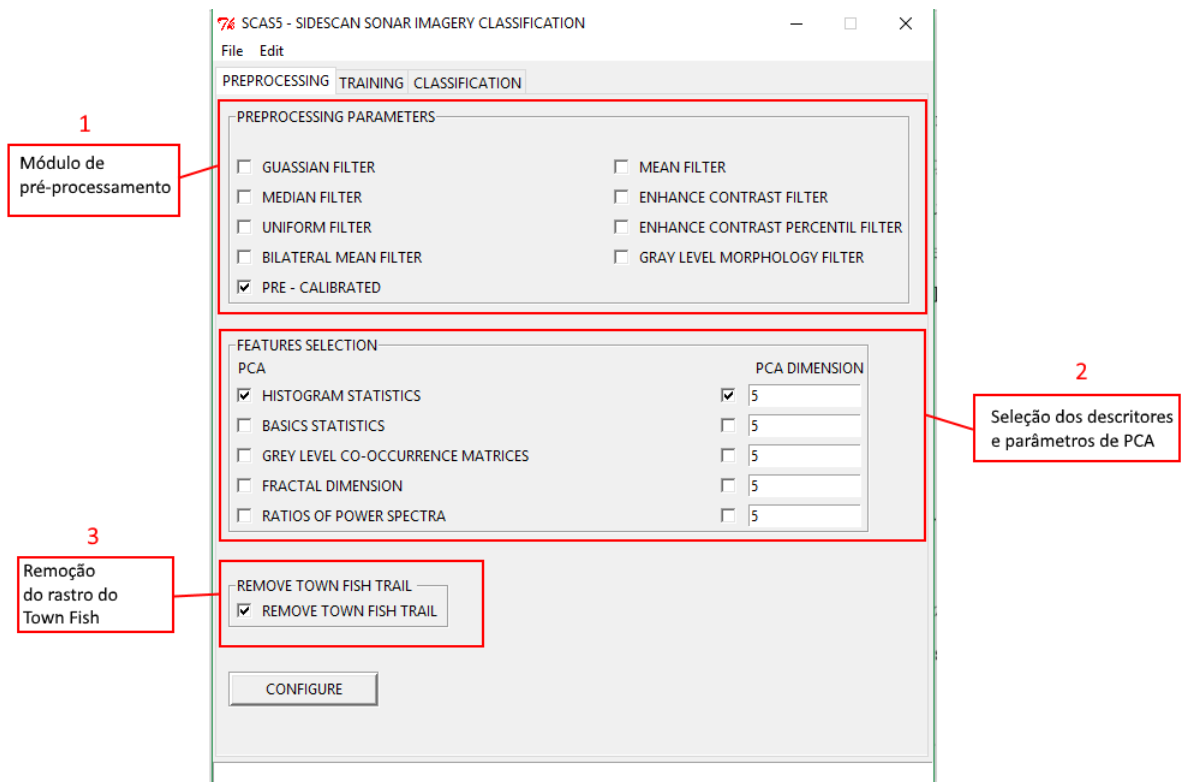


Figura 4.1: Demonstração do modulo de pré-processamento.

que serão exploradas em etapas subsequentes.

Neste módulo também é feita a escolha das classes de descritores (2) que serão empregadas na classificação, assim como, a escolha da utilização da análise de componentes principais para redução da dimensionalidade do espaço de características. Uma vez selecionado um número maior do que um de descritores, o sistema interpretará automaticamente que no final a classificação será empregada para cada classe de descritores e por fim será feita união destes resultados através da sabedoria das massas.

Na figura 4.2 é apresentado o módulo de treinamento, neste módulo são inseridos os parâmetros de treinamento (1) como a escolha do conjunto de amostras de treinamento, o número de classes amostrais e o número de amostras de treinamento. Por fim são selecionados os classificadores (2) que serão utilizados, de forma análoga a escolha dos descritores, se for escolhido um número maior do que um de classificadores, ao final,

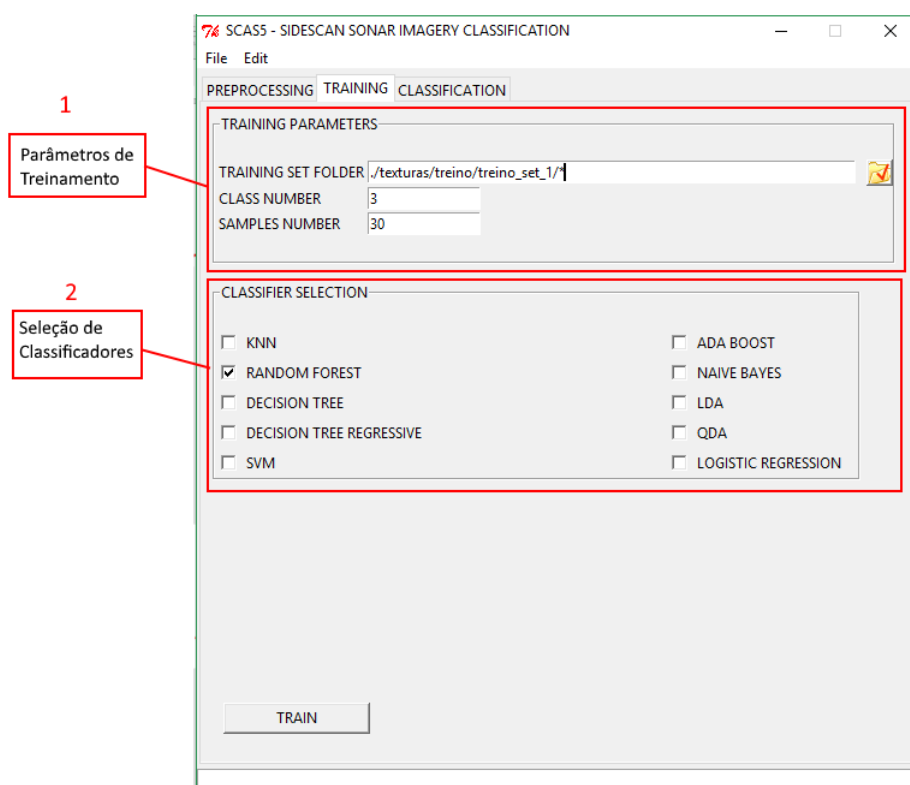


Figura 4.2: Demonstração do módulo de treinamento.

será empregado um segundo ponto de *ensemble* entre os classificadores.

A figura 4.3 exemplifica o módulo de classificação, neste módulo são selecionados a sonografia a ser classificada e opcionalmente uma sonografia já classificada de modo a aferir a precisão do sistema a fim de testar diferentes configurações para o processo (1). No módulo de classificação também é feita a seleção e configuração da técnica de segmentação escolhida (2) e apresenta um área de log que irá informar o usuário sobre o status do sistema e os resultados numéricos obtidos.

4.3 Resultados Quantitativos

Com o objetivo de mensurar o percentual de *pixels* classificados corretamente entre as classes sedimentares contempladas neste trabalho, foi utilizado como métrica principal

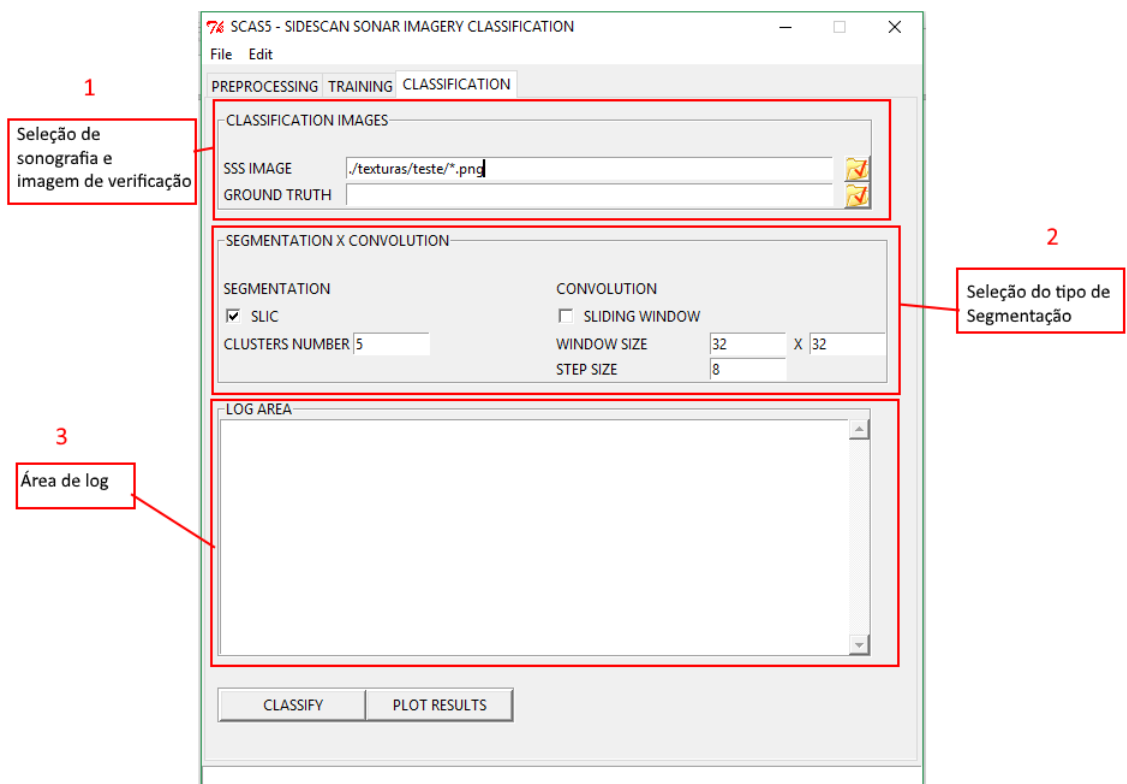


Figura 4.3: Demonstração do módulo de classificação.

a precisão, conforme demonstrado na equação 4.1, onde VP é a quantidade de *pixels* classificados corretamente pelo sistema comparando *pixel* a *pixel* com a imagem classificada pelo especialista humano e TP é a quantidade total de *pixels* que compõem a imagem.

$$Precisão = \frac{VP}{TP} \quad (4.1)$$

Nas figuras 4.4 e 4.5 são apresentados comparativos gráficos referentes a precisão entre os diferentes algoritmos de classificação utilizando-se das técnicas de *superpixel* e *sliding window*, respectivamente. Visando demonstrar a precisão atingida com diferentes algoritmos de classificação valendo-se de cada classe de descritores de maneira independente e também com a utilização da sabedoria das massas (*Ensemble*) através do voto da maioria.

Em ambos os gráficos mencionados, o eixo das abcissas representa os diferentes algoritmos de classificação disponibilizados, o eixo das ordenadas representa a precisão média e sobre as barras é apresentado o desvio padrão, ambos em porcentagem. Conforme observado, para cada classificador utilizado são apresentadas cinco barras de precisão, sendo as quatro primeiras referentes a precisão obtida a partir da utilização do classificador em questão com cada classe de descritores isoladamente, por exemplo, AB com estatísticas de histogramas, AB com fractais e assim sucessivamente e a quinta (Barra vermelha) referente a precisão obtida com a união das quatro primeiras através da aplicação de sabedoria das massas (*Ensemble*).

Nas imagens 4.6 e 4.7, são demonstrados os resultados de precisão a partir da utilização da sabedoria das massas em dois pontos do processo, (i) na composição de uma solução para cada classificador através da união de soluções isoladas com cada classe de descritores, conforme demonstrado nas figuras 4.4 e 4.5 e (ii) na união dos resultados de cada classificador, conforme ilustrado na figura 4.8.

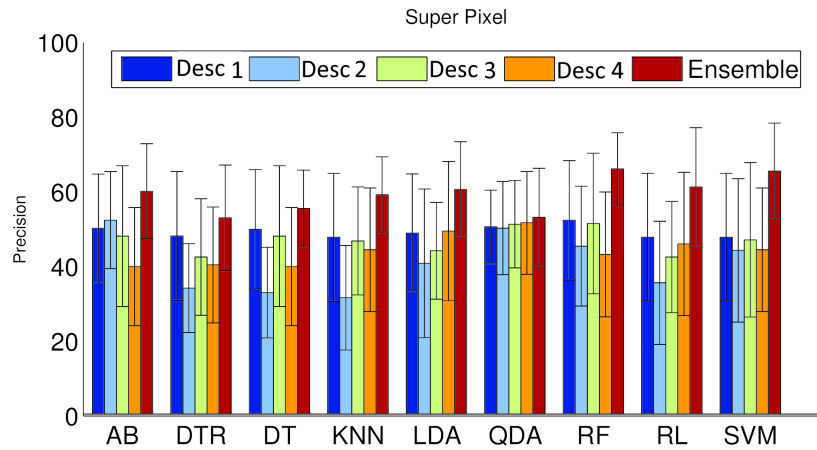


Figura 4.4: Comparativo de precisão entre classificadores com a técnica de Superpixels, onde eixo das abcissas representa os diferentes algoritmos de classificação disponibilizados, o eixo das ordenadas representa a precisão média atingida em porcentagem e sobre as barras é apresentado o desvio padrão. As quatro primeiras barras representam a precisão média atingidas utilizando os descritores independentemente sendo eles: Desc1 (Estatísticas de histogramas), Desc2 (Estatísticas de pixels), Desc3 (Matrizes de co-ocorrência de níveis de cinza) e Desc4 (Potência espectral) e a quinta barra representa o resultado de precisão obtido após o primeiro ponto de sabedoria das massas.

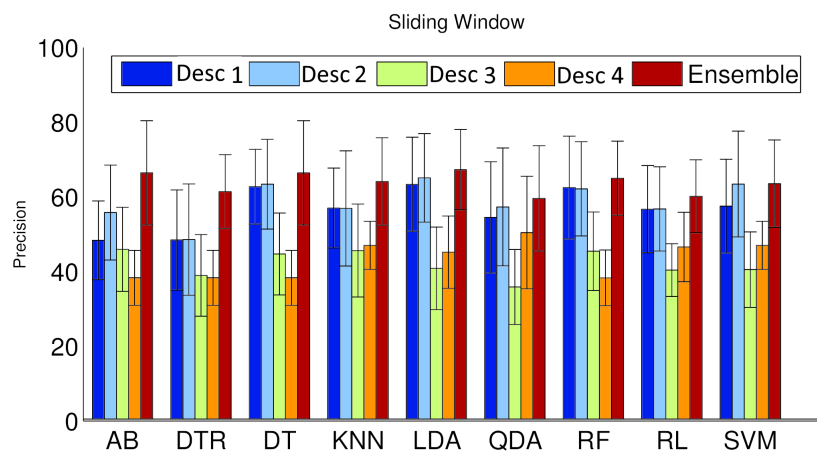


Figura 4.5: Comparativo de precisão entre classificadores com a técnica de *sliding window*, o eixo das abcissas representa os diferentes algoritmos de classificação disponibilizados, o eixo das ordenadas representa a precisão média atingida em porcentagem e sobre as barras é apresentado o desvio padrão. As quatro primeiras barras representam a precisão média atingidas utilizando os descritores independentemente sendo eles: Desc1 (Estatísticas de histogramas), Desc2 (Estatísticas de pixels), Desc3 (Matrizes de co-ocorrência de níveis de cinza) e Desc4 (Potência espectral) e a quinta barra representa o resultado de precisão obtido após o primeiro ponto de sabedoria das massas.

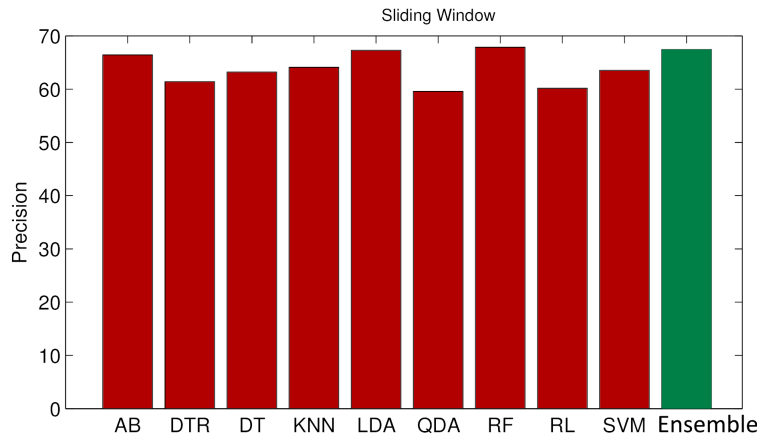


Figura 4.6: Comparativo da precisão entre classificadores independentes com a utilização do primeiro ponto de sabedoria das massas e da aplicação do segundo ponto de sabedoria das massas onde são unidos os resultados dos classificadores independentes (Barra verde), com segmentação por *sliding window*.

Nestes gráficos (4.6 e 4.7), o eixo das abcissas representa os diferentes algoritmos de classificação disponibilizados e o eixo das ordenadas a precisão média atingida em porcentagem após a utilização do primeiro ponto de sabedoria das massas (Barras vermelhas) e por último a precisão média obtida aplicando o segundo ponto de sabedoria das massas (Verde).

No entanto, mesmo tendo demonstrado ganhos na precisão a adição deste ponto de *ensemble* não demonstrou ganhos significativos em relação a utilização do classificador RF ou AB que também consistem em classificadores do tipo *ensemble*, observando que o ganho de precisão não compensa a perda de desempenho computacional na adição desta etapa.

Logo, na sequência será abordada apenas a utilização de um ponto de *ensemble* entre as classes de descritores e será utilizado como algoritmo de classificação padrão o RF por ter demonstrado os melhores resultados em termos de precisão entre os classificadores testados.

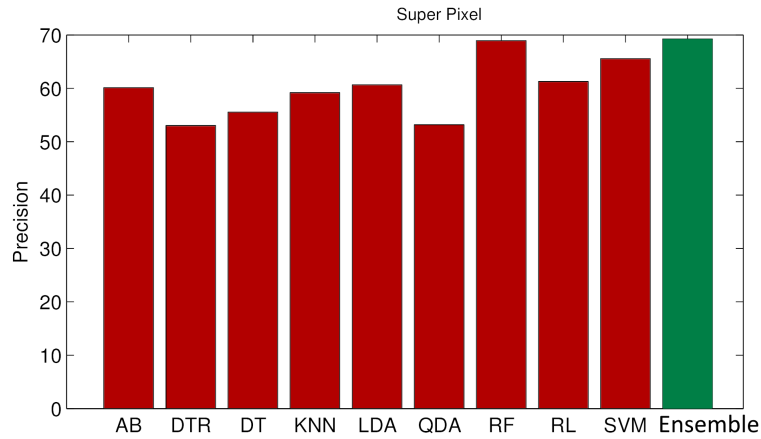


Figura 4.7: Comparativo da precisão entre classificadores independentes com a utilização do primeiro ponto de sabedoria das massas e da aplicação do segundo ponto de sabedoria das massas onde são unidos os resultados dos classificadores independentes (Barra verde), com segmentação por *superpixel*.

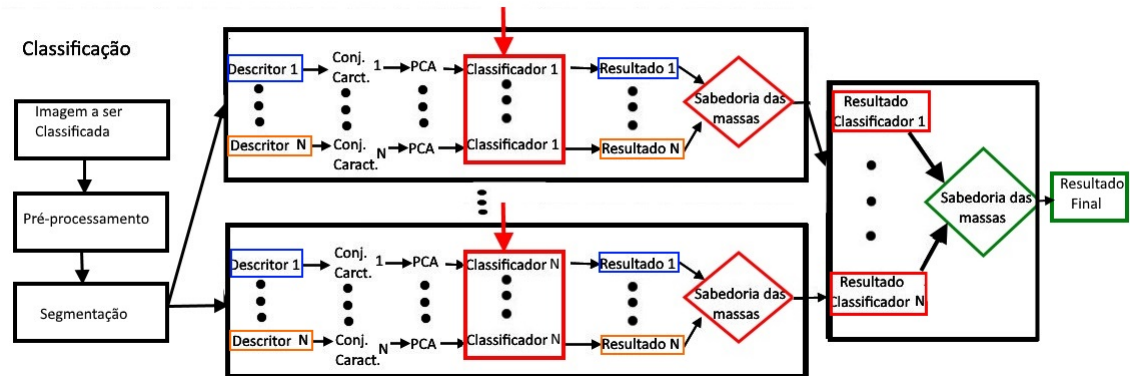


Figura 4.8: Ilustração processo de composição resultado com dois pontos de sabedoria das massas.

4.4 Resultados Qualitativos

Nesta seção serão demonstrados os resultados qualitativos através das sonografias classificadas, conforme destacado anteriormente dentre distintos algoritmos de classificação, o algoritmo *Random Forest* apresentou melhor precisão logo os resultados a seguir são baseados no mesmo utilizando-se um ponto de *ensemble*. Nas figuras 4.9 e 4.10 são demonstrados os resultados da classificação de uma área contemplando as três classes sedimentológicas foco deste trabalho, nestas imagens é possível observar uma comparação direta entre a classificação manual feita pelo especialista humano e também os resultados atingidos com a metodologia proposta utilizando-se a técnica *superpixel* e *sliding window* para segmentação.

Conforme observado, ambos os resultados foram bastante aproximados da classificação manual feita pelo especialista no entanto, em termos qualitativos a classificação com a segmentação baseada em super segmentação apresenta um mapa temático mais suave preservando a complexidade geométrica da transição entre as classes sedimentológicas. Fato este devido a característica de agrupamento dos *pixels* utilizado na super segmentação baseado no algoritmo *k-means* que agrupa os *pixels* pela distância Euclideana gerando segmentos complexos e com propriedades semelhantes, evitando o problema das zonas limítrofes entre classes.

Na sequencia a figura 4.11, demonstra os resultados de classificação obtidos com o algoritmo *Random Forest* a partir da utilização de classes de descritores distintas de maneira independente, ou seja, existe uma instância do classificador para cada classe de descritores. Por fim, são demonstradas imagens resultantes da união dos resultados prévios através da sabedorias das massas utilizando o voto ponderado, com segmentação por *superpixels* e segmentação com janela de convolução.

É possível observar que a utilização de classificadores especializados em uma única classe de descritores, como por exemplo, descritores especializados apenas em textura

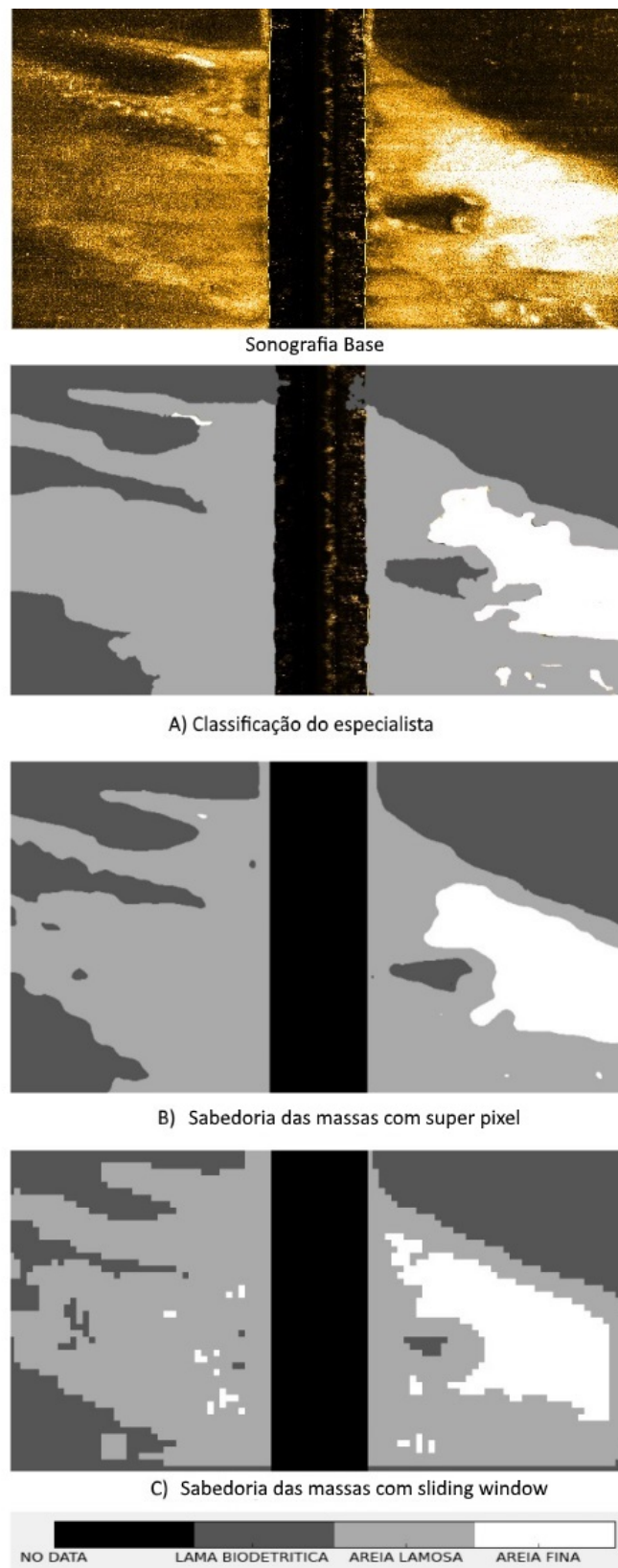


Figura 4.9: Comparação entre o resultado da classificação humana e da metodologia apresentada.

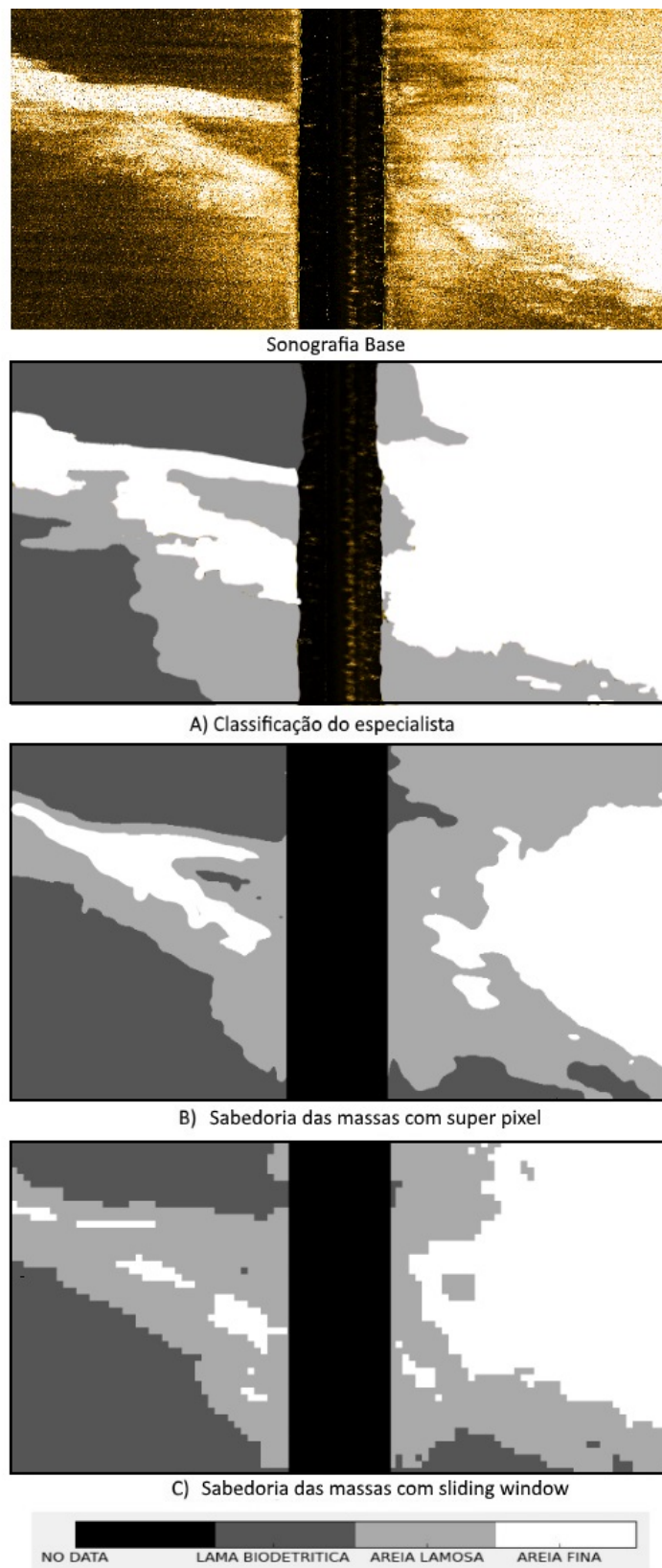


Figura 4.10: Comparação entre o resultado da classificação humana e da metodologia apresentada.

não são capazes de descrever de maneira adequada as distintas classes sedimentares presentes. Fato este, devido a natureza heterogenia dos elementos que compõem cada cena, sendo estes suscetíveis a pequenas alterações na textura ou intensidade ocasionados por ruídos ou estruturas naturais do meio.

Também é possível observar que devido a etapa treinamento ser uma etapa rígida, em que o sistema aprende um número pré-determinado de classes. Cenas que apresentam um número inferior ou superior de classes tende a apresentar resultados inferiores, pois, na presença de classes com baixa variabilidade o sistema tenderá a convergir para a classe mas próxima no momento. Logo, é essencial um novo treinamento na presença de novas classes sedimentares ou em áreas que apresentam um número reduzido ao treinado.

Conforme observado também na imagem [4.12](#), os resultados obtidos através da sabedoria das massas são mais interessantes em termos qualitativos e quantitativos, assim como o emprego da segmentação por *superpixels* preserva a complexidade geométrica entre os sedimentos proporcionando imagens classificadas mais acentuadas qualitativamente enquanto o uso da segmentação por *sliding window* tende a construir bordas serrilhadas não preservando detalhes ou a presença de objetos discretos na imagem.

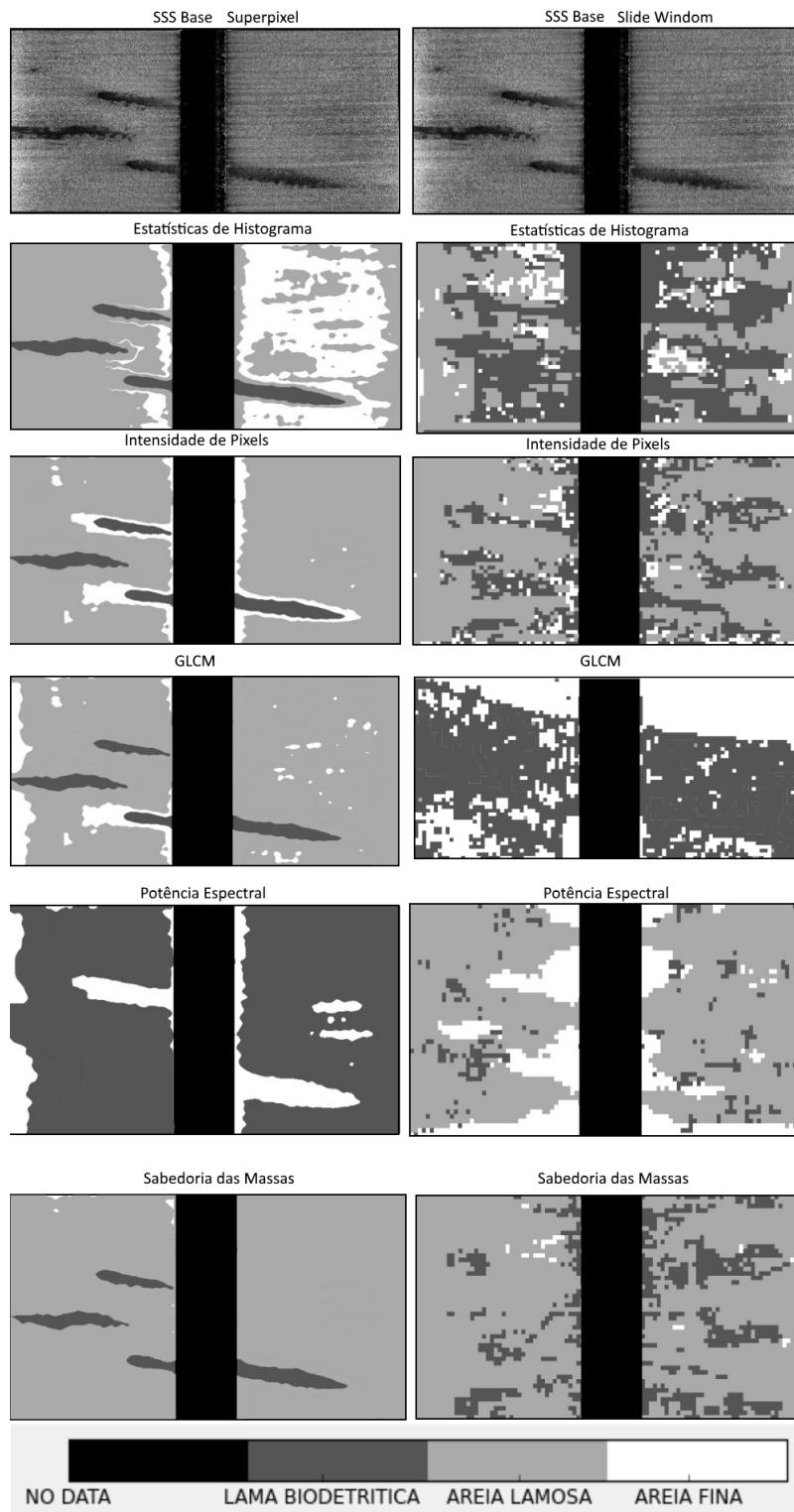


Figura 4.11: Comparativo resultados qualitativos entre *superpixel* e *sliding window*, cada imagem representa um resultado de classificação utilizando-se um classificador especialista para cada classe de descritor independentemente e no final a união com sabedoria das massas.

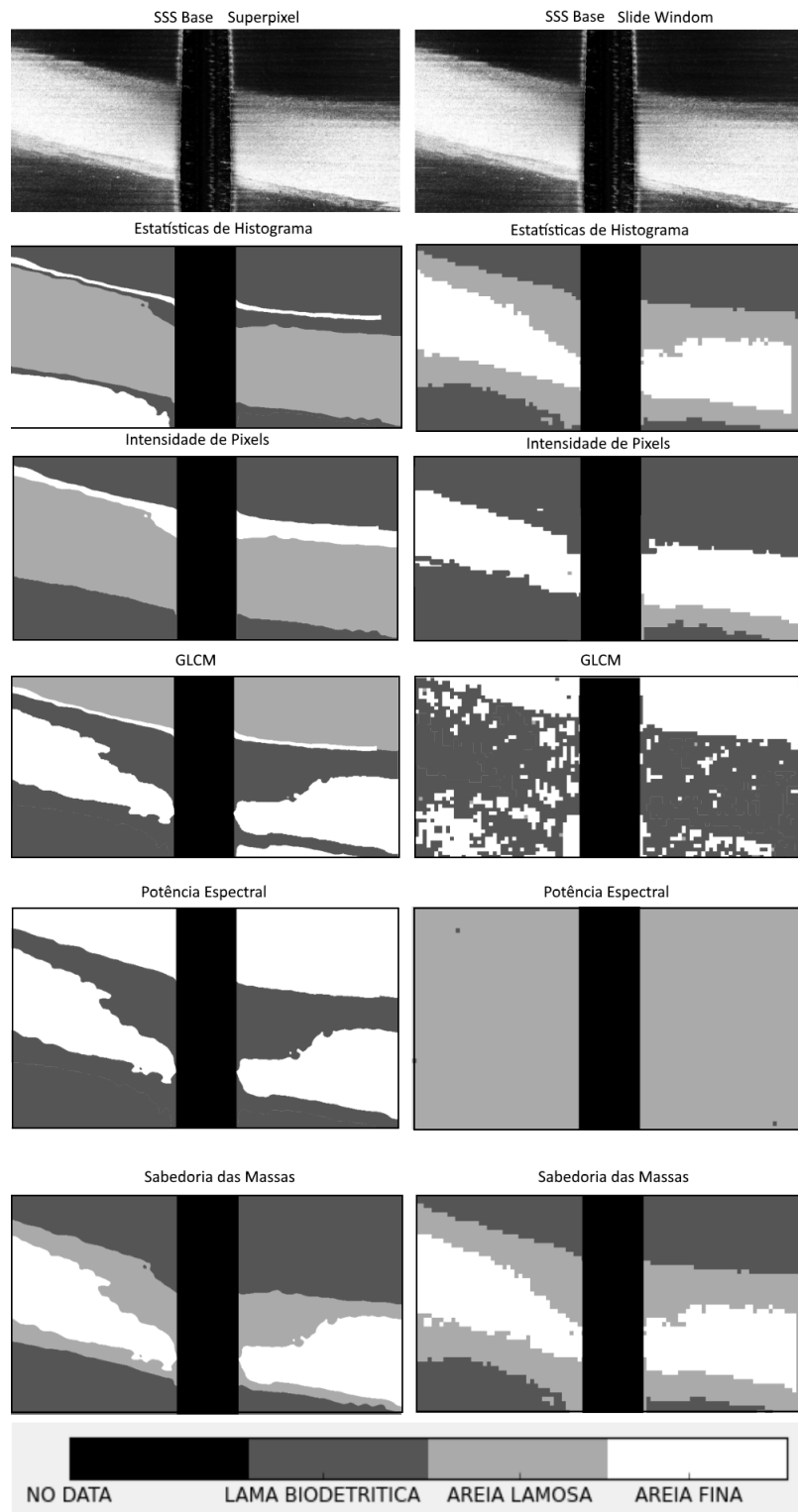


Figura 4.12: Comparativo resultados qualitativos entre *superpixel* e *sliding window*, cada imagem representa um resultado de classificação utilizando-se um classificador especialista para cada classe de descritor independentemente e no final a união com sabedoria das massas.

Capítulo 5

Discussão

5.1 Os resultados

Conforme proposto, foi desenvolvido um *framework* totalmente na linguagem Python contendo uma metodologia pré-configurada baseada na sabedoria das massas para classificação de sedimentos superficiais a partir de sonografias de sonar de varredura lateral. O *framework* é dividido em módulos de pré-processamento, treinamento e classificação cada módulo apresenta um conjunto de pontos de tomada de decisão para o usuário proporcionando o ajuste do processo de classificação de acordo com as suas necessidades.

No caso base, o *framework* é pré-configurado para efetuar a classificação através da metodologia proposta com dois pontos de sabedoria das massas, valendo-se de múltiplas classes de descritores e múltiplos classificadores supervisionados.

Conforme apresentado no capítulo 4, os resultados obtidos em termos quantitativos e qualitativos com a metodologia proposta foram satisfatórios quando considerado a baixa variabilidade entre as classes sedimentológicas analisadas, tratando-se apenas de areia, lama e lama-biodetrítica.

Conforme demonstrou-se nos gráficos 4.4 e 4.5 a precisão (Barra vermelha) inferida a partir da união dos resultados locais através da sabedoria das massas é superior

a precisão obtida ao se utilizar classes de descritores de maneira independente, isto deve-se a especialização de cada classe de descritores como por exemplo as matrizes de co-ocorrência de níveis de cinza especialista na extração de características de textura e nas estatísticas de pixels especializadas na extração de características de intensidade, que quando unidos resultam na composição destas especialidades.

Também é possível observar nos gráficos 4.4 e 4.5 que em relação a precisão a técnica de janela de convolução demonstrou-se superior em termos quantitativos, fato este observado devido a simetria constante da janela, ou seja, sempre operando sobre segmentos uniformes de 32×32 *pixels* sendo esta uma característica importante para a utilização de descritores baseados na dimensão fractal, enquanto, a super segmentação por *superpixels* não garante proporções simétricas entre seus segmentos quando operada com um número reduzido de segmentos no entanto, com a elevação do número de segmentos o ganho em relação a simetria e a precisão é elevado no entanto o desempenho é reduzido.

De forma geral, em termos quantitativos a metodologia proposta apresentou precisão média de 68% quando comparada pixel a pixel com as sonografias classificadas pelo especialista, porém, manteve-se superior a utilização de qualquer classificador independente validando a hipótese de que a união de soluções oriunda de classes distintas de descritores resultaria em ganhos de precisão. Neste sentido, diferentes técnicas de *ensemble* foram testadas como a média dos resultados, voto da maioria e o voto ponderado da maioria no entanto devido a especialização das diferentes classes de descritores o voto ponderado da maioria apresentou melhores resultados na composição da classificação final.

Em termos qualitativos, é possível observar na figura 4.9 que a técnica de super segmentação por *superpixels* preserva as transições suaves entre as classes sedimentológicas, possibilitando a análise de regiões limítrofes entre classes e também possibilitando a construção de mosaicos temáticos de áreas amplas. Enquanto a técnica de janela de

convolução (*Sliding Window*) resulta em mapas de classificação serrilhados fato este ocasionado pela dimensão da janela de convolução, quanto maior a janela melhor o desempenho computacional e pior o resultado qualitativo.

É interessante observar que a aplicação da técnica de janela de convolução sobre imagens que apresentam amostras de sedimentos de tamanho reduzido pode ocasionar a perda ou sobreposição destas amostras conforme demonstrado nas figuras 4.12 e 4.11. Isto ocasiona-se devido as dimensões rígidas da janela de convolução que em regiões limítrofes entre duas ou mais classes pode ocasionar falha na predição, devido a presença de características distintas dentro da mesma área cobertura da janela de convolução.

Capítulo 6

Conclusões

A crescente necessidade pela exploração em ambientes subaquáticos demanda a pesquisa e desenvolvimento de recursos tecnológicos que auxiliem a romper as barreiras impostas para exploração nestes ambientes. Neste contexto, a especificação da composição do substrato marinho é de suma importância em diversos setores como setores ambientais, engenharia oceânica e a indústria do petróleo.

Usualmente estudos do fundo marinho que valem-se de amostragem física coletadas *in situ* que demandam elevados recursos humanos e financeiros, visto a necessidade de embarcações e equipes especializadas. Técnicas baseadas em amostragem remota são alternativas eficientes e custo reduzido quando comparadas com as metodologias tradicionais, proporcionando a cobertura de grandes áreas sem a necessidade de paradas ou equipes especializadas.

Considerando a indústria do petróleo, a especificação dos sedimentos que constituem as camadas superiores do leito oceânico é uma das etapas iniciais na exploração de recursos fósseis, pois através dela é possível planejar não apenas sobre como a perfuração procederá mas também fatores ligados a segurança na instalação de plataformas e equipamentos.

Tratando de sistemas para classificação remota, dado base desta metodologia as

sonografias de sonar de varredura lateral são imagens de alta resolução formadas a partir do eco acústico retornado ao equipamento. Este eco retorna amortecido de acordo com o tipo de obstáculo ao qual foi incidido. Criando desta forma a variabilidade visualidade nas sonografias. Esta técnica de imageamento é suscetível a uma série de problemas do processo de aquisição resultando em um elevado índice de ruídos a serem tratados, dificultando a interpretação destas imagens.

Neste sentido, este trabalho apresentou uma ferramenta metodologia computacional capaz de efetuar a classificação supervisionada de sedimentos superficiais de leitos oceânicos e de águas continentais, empregando conceitos de visão computacional, aprendizado de máquina e sabedoria das massas desenvolvendo uma abordagem baseada na composição de um solução global a partir de soluções locais provenientes de especialistas.

A abordagem é centrada na utilização de classes de descritores especialistas em diferentes atributos presentes em imagens como intensidade de pixels e texturas e na utilização de diferentes algoritmos de classificação como algoritmos probabilísticos e de regressão. A união destes diferentes especialistas resulta na composição das características individuais ampliando a precisão do resultado final.

A metodologia proposta mediante experimentos, apresentou resultados considerados satisfatórios quando observada a baixa variabilidade de classes sedimentares disponíveis para este estudo. Os experimentos efetuados demonstram que a composição do resultado através de técnicas de sabedoria das massas resulta em ganhos de precisão, gerando resultados superiores a qualquer classe de descritores utilizados independentemente e a utilização de segmentação por superpixel confere ao sistemas ganhos em relação ao desempenho quando comparado à técnicas de segmentação clássica como as janelas de convolução e também proporciona resultados visuais mais suaves nas transições entre as classes observadas nas imagens.

Por fim, este trabalho possui um enfoque na multidisciplinaridade, conceito no qual

foi elaborado e desenvolvido, visando explorar a natureza do trabalho de engenharia. Com a finalidade de contribuir não apenas na área da computação, mas também em outras áreas como oceanografia e geologia marinha.

Com intuito de servir de recurso em empreendimentos os quais necessitam/demandam novas tecnologias para o auxílio em tarefas de pesquisa, monitoramento e exploração em ambientes marinhos e de águas continentais que requerem a interpretação de imagens de sonar de varredura lateral.

As principais contribuições deste trabalho podem ser listadas conforme apresentado a seguir.

- Desenvolvimento de um *framework* computacional totalmente operacional, o qual oferece uma nova metodologia para classificação de sedimentos superficiais de maneira eficiente e também a possibilidade de ser ajustado de acordo com as necessidades e preferências do usuário;
- O sistema preza pelo desempenho e acurácia, visando a possibilidade de uso a bordo de cruzeiros oceanográficos;
- O sistema possibilita a mensuração de similaridade em contraste com imagens de sonar de varredura lateral classificadas por especialistas humanos, possibilitando calibrar a ferramenta e também verificar a acurácia do sistema durante o processo de classificação;
- Como resultados, o *framework* apresentado disponibiliza a criação de mapas temáticos coloridos de acordo com as classes as quais se deseja classificar, facilitando desta forma a visualização e interpretação dos dados obtidos pelo sistema de sonar de varredura lateral.

Referências Bibliográficas

- Achanta, R., Shaji, A., Smith, K., Luchhi, A., Fua, P., and Susstrunk, S. (2010). SLIC Superpixels. *EPFL Technical Report 149300*.
- Anderson, J. T., Van Holliday, D., Kloser, R., Reid, D. G., and Simard, Y. (2008). Acoustic seabed classification: Current practice and future directions. *ICES Journal of Marine Science*, 65(6):1004–1011.
- Araujo, W. O. D. and Coelho, C. J. (2009). Análise de Componentes Principais (PCA).
- ASM, Shihavuddin (2014). *Automated Underwater Object Classification using Optical Imagery*. PhD thesis, University of Girona.
- Atallah, L. and Probert Smith, P. (2004). Automatic seabed classification by the analysis of sidescan sonar and bathymetric imagery.
- Ballard, D. H. and Brown, C. M. (1982). *Computer Vision*. Prentice Hall. ISBN 0-13-165316-4.
- Blondel, P., Parson, L., and Robigou, V. (1998). TexAn: textural analysis of sidescan sonar imagery and generic seafloor characterisation. *IEEE Oceanic Engineering Society. OCEANS'98. Conference Proceedings (Cat. No.98CH36259)*, 1.
- Borges, H. B. (2006). *Redução de Dimensionalidade em Bases de Dados de Expressão Gênica*. PhD thesis, Pontifícia Universidade Católica do Paraná.

- Bulla, L. a. S., Cunha, R. P., Griep, G. H., and Hartmann, C. (2003). Uso Do Sonar De Varredura Lateral Na Avaliação De Uma Área De Mineração De Areia No Rio Jacui - Rs .
- Capus, C. G., Banks, A. C., Coiras, E., Ruiz, I. T., Smith, C. J., and Petillot, Y. R. (2008). Data correction for visualisation and classification of sidescan SONAR imagery. *Iet Radar Sonar and Navigation*, 2:155–169.
- Carmichael, D. (1998). Image processing techniques for the analysis of sidescan sonar survey data. *IEEE Colloquium on Underwater Applications of Image Processing*, 1998:3–3.
- Chen, H., De, P., and Hwang, B.-h. (2013). Wisdom of crowds : the value of stock opinions transmitted through social media.
- Dieterich, T. G. (2000). Ensemble Learning in machine learning. pages 1—15.
- Dietterich, T. G. (2000). Hierarchical Reinforcement Learning with the MAXQ Value Function Decomposition. *Journal of Artificial Intelligence Research*, 13:227–303.
- Fakiris, E. and Papatheodorou, G. (2013). SonarClass : A MATLAB TOOLBOX FOR THE CLASSIFICATION OF SIDE SCAN SONAR IMAGERY , USING LOCAL TEXTURAL AND REVERBERATIONAL CHARACTERISTICS.
- Fayad, M. and Schmidt, D. (1997). Object-oriented application frameworks. *Communications of the ACM*, 40(10):32–38.
- Forsyth, D. A. and Ponce, J. (2011). *Computer Vision A Modern Approach*. Pearson. ISBN 978-0136085928.
- Geladi, P., Isaksson, H., Lindqvist, L., Wold, S., and Esbensen, K. (1987). Principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, 2(1-3):37–52.

- Geladi, P., Isaksson, H., Lindqvist, L., Wold, S., and Esbensen, K. (1989). Principal component analysis of multivariate images. *Chemometr. Intell. Lab. Syst.*, 5(3):209.
- Gonzales, R. and Woods, R. E. (2007). *Digital Image Processing*. Prentice Hall; 3 edition (August 31, 2007). ISBN 978-0131687288.
- Hair, J., Anderson, R. E., Tathan, R., and Black, W. C. (1995). *Multivariate Data Analysis*. Prentice Hall, New Jersey. ISBN 978-0138132637.
- Haralick, R. M. and Shanmugan, K. (1973). Textural Features for Image Classification. pages 610—621.
- Haralick, R. M. and Shapiro, L. G. (1992). *Computer an Robot Vision, Vol I*. Addison-Wesley. ISBN 978-0201108774.
- Honorato, L. B. and Milano, D. D. (2010). Visão computacional. pages 1–7.
- Imen, K., Fablet, R., Boucher, J. M., and Augustin, J. M. (2005). Statistical discrimination of seabed textures in sonar images using co-occurrence statistics. In *Oceans 2005 - Europe*, volume 1, pages 605–610.
- Jain, A. K., Duin, R., and MAO, J. (2000). Statistical Pattern Recognition: A Review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 4–37.
- Kunzendorf, H. (1986). *Marine Mineral Exploration*. Elsevier Science. ISBN 978-0-444-42627-7.
- Linnett, L. M., Clarke, S. J., and Tress, A. D. (1993). Texture Analysis of Side Scan Sonar Data. *IEE, Savoy Place, London WCPR OBL. UK*.
- Lurton, X. (2002). *An Introduction to Underwater Acoustics: Principles and Applications*. Springer; 2nd ed. 2010 edition (September 7, 2010).

- Marbach, D., Costello, J. C., Küffner, R. M., Vega, N. M., Prill, D. M., Kyle, R., Aderhold, and Stolovitzky, G. (2012). Wisdom of crowds for robust gene network inference. *Nature Methods*, 9(8):796–804.
- Mendoza, M. R. (2014). Exploring ensemble learning techniques to optimize the reverse engineering of gene regulatory networks. (March):189.
- Michaelis (2005). *Michaelis Língua Portuguesa*. Melhoramentos.
- Minsky, M. (1974). A Framework to Representing Knowledge. *MIT-AI Laboratory Memo 306*.
- Oliveira Junior, A. M. (2012). *Classificação Remota de Sedimentos do Fundo Oceânico: Correlação Estatística entre Mapas e Amostras*. PhD thesis, Universidade Federal Fluminense, UFF.
- Ozkan, D. and Morency, L.-p. (2011). Modeling Wisdom of Crowds Using Latent Mixture of Discriminative Experts. *Computational Linguistics*, pages 335–340.
- Pace, N. and Dyer, C. (1979). Machine Classification of Sedimentary Sea Bottoms. *IEEE Transactions on Geoscience Electronics*, 17.
- Pedrini, H. (2008). *Análise de imagens digitais: princípios, algoritmos e aplicações*. Thomson Learning, 1a edição edition. ISBN 978-85-221-0595-3.
- Polikar, R. (2008). Ensemble Learning. 24:56—72.
- Pressman, R. S. (2006). *Engenharia de Software*. Bookman.
- Sharkey, A. and Sharkey, N. (1997). Combining diversses neural networks. 12:231—247.
- Sommerville, I. (2011). *Engenharia de Software*. Pearson Prentice Hall.

- Stewart, W. and Marra, M. (1994). A neural network approach to classification of sidescan sonar imagery from a midocean ridge area. *IEEE Journal of Oceanic Engineering*, 19:214–224.
- Subramaniam, S., Barad, H., Martinez, A., and Bourgeois, B. (1993). Seafloor characterization using texture. *Proceedings of Southeastcon '93*.
- Surowiecki, J. (2005). *The Wisdom of Crowds*. Anchor; Reprint edition (August 16, 2005).
- Szaliski, R. (2010). *Computer Vision Algorithms and Applications*. Springer. ISBN 978-1-1-84882-934-3.
- Tęgowski, J. (2005). Acoustical classification of the bottom sediments in the southern Baltic Sea. *Quaternary International*, 130(1):153–161.
- Theodoridis, S. and Koutroumbas, K. (2008). *Pattern Recognition*. Academic Press. ISBN 978-1597492720.
- Tomaso, P. (2013). Mouse Behavior Recognition with The Wisdom of Crowd. (2010).
- Velic, M., Grzinic, T., and Padavic, I. (2013). Wisdom of Crowds Algorithm for Stock Market Predictions. page 7.
- Watt, A. H. and Policarpo, F. (1998). *The Computer Image*. Addison-Wesley Pub Co (Net). ISBN 020-1-42298-0.
- Welinder, P., Branson, S., Belongie, S., and Perona, P. (2010). The Multidimensional Wisdom of Crowds. *Most*, 6:1–9.
- Welsby, V. (1971). *Underwater acoustics*, volume 9.
- Winston, P. H. (1976). *The psychology of computer vision*, volume 8.

Zortea, M. and Haertel, V. (2004). *Investigação sobre métodos para redução da dimensionalidade dos dados em imagens hiperspectrais*. PhD thesis, UFRGS.