

UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE TECNOLOGIA

Marina Balsamão Percebão

**Uso de Técnicas de Agrupamento K-Means e Fuzzy C-Means para
Análise de Tomografia Ultrassônica de Toras de Madeira**

Limeira, 2019

Marina Balsamão Percebão

Uso de Técnicas de Agrupamento K-Means e Fuzzy C-Means para Análise de Tomografia Ultrassônica de Toras de Madeira

Monografia apresentada à Faculdade de Tecnologia da Universidade Estadual de Campinas como parte dos requisitos para a obtenção do título de Bacharel em Sistemas de Informação.

Orientador: Prof. Dr. Marco Antonio Garcia de Carvalho

Este exemplar corresponde à versão final da Monografia defendida por Marina Balsamão Percebão e orientada pelo Prof. Dr. Marco Antonio Garcia de Carvalho.

Limeira, 2019

UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE TECNOLOGIA

Marina Balsamão Percebão

**Uso de Técnicas de Agrupamento K-Means e Fuzzy C-Means para
Análise de Tomografia Ultrassônica de Toras de Madeira**

Banca Examinadora:

- Prof. Dr. Marco Antonio Garcia de Carvalho
FT/UNICAMP
- Prof. Dr. Guilherme Palermo Coelho
FT/UNICAMP
- Prof Dr. Paulo Sérgio Martins Pedro
FT/UNICAMP

Limeira, 2019

Resumo

O objetivo deste trabalho é analisar qualitativamente o resultado dos algoritmos de agrupamento K-Means e Fuzzy C-Means na segmentação de imagens digitais. A contribuição acadêmica deste estudo servirá para detecção de deterioração em toras de madeira, fazendo com que não haja impactos físicos na estrutura dos objetos estudados, a fim de melhorar a visualização das anomalias encontradas. Esta monografia é uma continuação dos estudos de Strobel (2017) e Junior et al. (2018). O uso dos algoritmos de agrupamento, após a interpolação dos dados, tem como propósito a melhor visualização das imagens geradas pós interpolação, a fim de destacar anomalias, como ocos, no interior de toras madeira. No desenvolvimento desta monografia, são abordados os temas de ultrassom em madeira, interpolação espacial, algoritmos de agrupamento, lógica Fuzzy e segmentação de imagens. Por fim, este trabalho faz a análise comparativa dos algoritmos de agrupamento, avaliando qual destes obtiveram melhores resultados para atingir o objetivo proposto. Os resultados obtidos das imagens segmentadas nos mostram que os algoritmos se comportaram de maneira semelhante, porém não foram muito eficientes pois destacaram as anomalias apenas quando o número de clusters aumentou, resultando em imagens parecidas com as originais.

Palavras-chave: imagem, ultrassom, ensaios não-destrutivos, madeira, interpolação, agrupamento, segmentação de imagem, clusterização.

Abstract

The objective of this work is to qualitatively analyze the result of the K-Means and Fuzzy C-Means clustering algorithms in the segmentation of digital images. The academic contribution of this study will serve to detect deterioration in wood logs, causing no physical impacts on the structure of the studied objects, in order to improve the visualization of the anomalies. This monograph is a continuation of the studies by Strobel (2017) and Junior et al. (2018). The use of clustering algorithms, after data interpolation, has the purpose of better visualization of the images generated after interpolation, in order to highlight anomalies, such as hollow, inside wood logs. In the development of this monograph, the topics of wood ultrasound, spatial interpolation, clustering algorithms, Fuzzy logic and image segmentation are addressed. Finally, this work makes the comparative analysis of the clustering algorithms, evaluating which one obtained the best results to reach the proposed objective. The result obtained from the segmented images show us that the algorithms behaved similarly, but were not efficient because they highlighted the anomalies only when the number of clusters increased, resulting in similar images to the original ones.

Keywords: image, ultrasound, non-structural tests, wood, interpolation, clustering, image segmentation.

Lista de Figuras

2.1 Equipamento ultrassônico USLab	14
2.2 Transdutores de faces planas e transdutores de faces exponenciais	14
2.3 Esquema da propagação da onda e do aumento do tempo em função de uma anomalia na madeira.....	15
2.4 Seção transversal da madeira e os eixos radial e tangencial e ilustração do eixo longitudinal paralelo às fibras longitudinais	16
2.5 Exemplo de um ponto com velocidades conhecidas e desconhecidas em uma tora de madeira	17
2.6 Dendograma	19
2.7 Gráfico de função de pertinência triangular.....	21
2.8 Disco 7 pós segmentação utilizando algoritmo Fuzzy C-Means com 4 clusters	21
3.1 Imagens das toras de madeira	24
3.2 Imagem resultante da interpolação IDW para o Disco 7.	24
3.3 Imagem resultante da interpolação IDW para o Disco 10	25
3.4 Imagem resultante da interpolação IDW para o Disco Liquidambar Styraciflua	25
3.5 Imagem resultante da interpolação IDW para o Disco Platanus	25
3.6 Ensaio de ultrassom em tora de madeira. As linhas em branco representam possíveis percursos das ondas entre os transdutores	26
3.7 Gráfico obtido utilizando o Método de Elbow.	28
4.1 Segmentação da imagem pós interpolação IDW com $p=1$ do Disco 7 utilizando K-Means.....	34
4.2 Segmentação da imagem pós interpolação IDW com $p=2$ do Disco 7 utilizando K-Means.....	34
4.3 Segmentação da imagem pós interpolação IDW com $p=1$ do Disco 10 utilizando K-Means.....	35
4.4 Segmentação da imagem pós interpolação IDW com $p=2$ do Disco 10 utilizando K-Means.....	35
4.5 Segmentação da imagem pós interpolação IDW com $p=1$ do Disco Liquidambar Styraciflua utilizando K-Means	35

4.6 Segmentação da imagem pós interpolação IDW com $p=2$ do Disco Liquidambar Styraciflua utilizando K-Means	35
4.7 Segmentação da imagem pós interpolação IDW com $p=1$ do Disco Platanus utilizando K-Means.....	36
4.8 Segmentação da imagem pós interpolação IDW com $p=2$ do Disco Platanus utilizando K-Means.....	36
4.9 Segmentação da imagem pós interpolação IDW com $p=1$ do Disco 7 utilizando Fuzzy C-Means	37
4.10 Segmentação da imagem pós interpolação IDW com $p=2$ do Disco 7 utilizando Fuzzy C-Means	37
4.11 Segmentação da imagem pós interpolação IDW com $p=1$ do Disco 10 utilizando Fuzzy C-Means	37
4.12 Segmentação da imagem pós interpolação IDW com $p=2$ do Disco 10 utilizando Fuzzy C-Means	37
4.13 Segmentação da imagem pós interpolação IDW com $p=1$ do Disco Liquidambar Styraciflua utilizando Fuzzy C-Means	38
4.14 Segmentação da imagem pós interpolação IDW com $p=2$ do Disco Liquidambar Styraciflua utilizando Fuzzy C-Means	38
4.15 Segmentação da imagem pós interpolação IDW com $p=1$ do Disco Platanus utilizando Fuzzy C-Means	38
4.16 Segmentação da imagem pós interpolação IDW com $p=2$ do Disco Platanus utilizando Fuzzy C-Means	38

Lista de Tabelas

3.1 Descrição das toras de madeira.....	23
4.1 Número de clusters obtido através do método de Elbow para cada imagem do dataset.	33

Sumário

1 Introdução	11
1.1 Motivação e Justificativa	11
1.2 Objetivos da Pesquisa	11
1.3 Organização do texto	12
2 Fundamentação teórica e trabalhos relacionados	13
2.1 Tomografia ultrassônica	13
2.1.1 Propagação da Onda	14
2.2 Propriedades da madeira	16
2.3 Interpolação espacial	16
2.4 Algoritmos de agrupamento	18
2.4.1 Algoritmos Hierárquicos	18
2.4.2 Algoritmos Particionais	19
2.4.3 Medidas de Similaridade	19
2.5 Lógica Fuzzy	20
2.6 Segmentação de Imagem	21
3 Materiais e Métodos	23
3.1 Dataset	23
3.2 Ultrassom em madeira	26
3.3 Interpolação Espacial	27
3.4 Método Elbow	28
3.5 K-means	29
3.6 Fuzzy C-Means	29
4 Resultados	32
4.1 Leitura das imagens	32
4.2 Geração das imagens	32

4.3 K-Means	32
4.4 Fuzzy C-Means	33
4.5 Experimentos	33
4.5.1 Definindo o número de clusters pelo Método de Elbow	33
4.5.2 Agrupamento via técnica K-Means	34
4.5.3 Agrupamento via técnica Fuzzy C-Means	36
4.6 Análise dos Resultados	39
5 Conclusões	40
Referências Bibliográficas	42

1 Introdução

Este capítulo tem como objetivo introduzir a discussão desta monografia, apresentar as motivações e justificativas, objetivos do trabalho e organização do texto.

1.1 Motivação e Justificativa

A madeira é um material que, historicamente, é de grande relevância no uso do setor da engenharia civil, transportes, bélico, entre outros. Por isso, é importante descobrir as propriedades e a qualidade da madeira que será usada como matéria prima.

Segundo Strobel (2017), existiam métodos para análise das propriedades que eram muito invasivos e provocavam danos irreversíveis à madeira. Assim, de nada adiantaria a peça analisada para qualquer fim que ela poderia ter, já que a mesma estaria danificada fisicamente. Portanto, técnicas não-destrutivas foram criadas para analisar a qualidade dos materiais estudados, minimizando possíveis impactos em suas características. Com o objetivo de diminuir o problema de danos às árvores, as técnicas acústicas são usadas na inspeção de madeiras, como o ultrassom.

O ensaio de ultrassom é frequentemente utilizado para verificar a qualidade interna da madeira, descobrindo, assim, degenerações, ocos ou qualquer anomalia no material. Este trabalho propõe avaliar se algoritmos de agrupamento, como o K-means e o Fuzzy C-means, podem conceber um resultado melhor para a classificação dos materiais estudados. Esses algoritmos foram escolhidos por serem conhecidos no agrupamento de dados e ambos podem ser utilizados também como método para segmentação de imagens digitais.

1.2 Objetivos da Pesquisa

O objetivo desse trabalho de conclusão de curso é realizar a análise comparativa entre dois algoritmos de agrupamento, o K-Means e o Fuzzy C-Means, analisando qualitativamente qual destes obtém um melhor resultado na segmentação de imagens, resultadas a partir de tomografias ultrassônicas de toras

de madeira. Além disso, essa monografia dá continuidade aos estudos realizados por Jorge Strobel (2017) e Antonio Alberto Jr. et al (2018).

1.3 Organização do texto

Esta monografia está dividida em cinco capítulos: este é o primeiro, com a introdução do trabalho, motivação e justificativa e objetivos. O segundo refere-se a fundamentação teórica e trabalhos relacionados, descrevendo tópicos fundamentais como tomografia ultrassônica, propriedades da madeira, interpolação espacial, algoritmos de agrupamento, lógica Fuzzy e segmentação de imagens. O terceiro explora os materiais e métodos de todo o processo, desde os ensaios de ultrassom até a implementação dos algoritmos K-means e Fuzzy C-means. O quarto apresenta os resultados obtidos dos métodos aplicados, K-Means e Fuzzy C-Means, analisando qualitativamente os algoritmos em relação a sua performance na segmentação de imagens. Por fim, o último capítulo conclui as últimas contribuições sobre os assuntos abordados.

2 Fundamentação teórica e trabalhos relacionados

Este capítulo descreve um resumo do embasamento teórico deste trabalho, citando as principais referências, artigos e livros. Serão abordados temas como tomografia ultrassônica, interpolação espacial, algoritmos de agrupamento, lógica Fuzzy e segmentação de imagens.

2.1 Tomografia ultrassônica

As técnicas não-destrutivas foram desenvolvidas por pesquisadores para determinar a qualidade interna do material estudado. No caso de toras de madeira, a tomografia possibilita descobrir se há ocos em árvores vivas ou em peças de tronco. Segundo Ross (1998), a definição de técnica não-destrutiva é a “identificação das propriedades físicas e mecânicas de uma peça de determinado material, sem alterar suas capacidades de uso” (apud PUCCINI, 2002, p. 5). Dessa forma, as árvores não precisam ser danificadas para que seja possível analisar a qualidade de seu tronco e/ou se há ocos internamente. Uma dessas técnicas não-destrutivas é o ultrassom, técnica acústica que utiliza a propagação de ondas, formando imagens 2D. De acordo com Ramirez (2015), esse método possibilita mapear o objeto estudado internamente através de diversas projeções de ensaios para identificar suas propriedades físicas. Alguns parâmetros, como amplitude e a velocidade da onda, possibilitam a detecção de degradações no objeto estudado (RAMIREZ, 2015).

Para que seja possível realizar a tomografia ultrassônica, é necessário um equipamento de ultrassom e os transdutores (Figura 2.1 e 2.2), os quais transmitem e recebem ondas de comprimentos de 40mm (SECCO, 2011). Para este trabalho, será utilizado como referência um dos equipamentos ultrassônicos do estudo de Secco (2011). O equipamento de ultrassom utilizado é o USLab, representado na Figura 2.1, desenvolvido pela empresa AGRICEF® em parceria com o grupo de pesquisa do LabEND/UNICAMP e os transdutores de faces planas e exponenciais.



Figura 2.1: Equipamento ultrassônico USLab. Fonte: (CERRI, 2010)

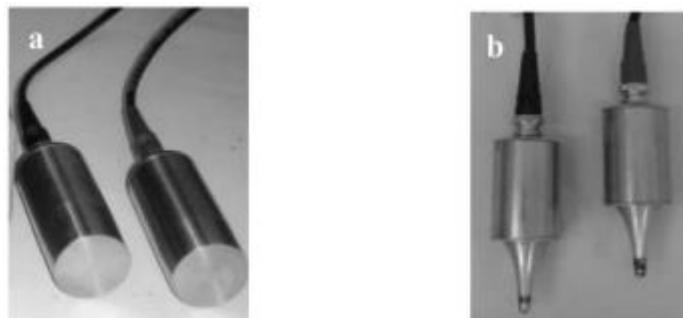


Figura 2.2: a) Transdutores de faces planas. b) Transdutores de faces exponenciais. Fonte: (SECCO, 2011).

De acordo com Strobel (2017), as imagens de ultrassom são construídas principalmente a partir do tempo de propagação, frequência da onda, entre outros parâmetros. Bucur (2005) diz que o aumento da frequência da onda amplia a definição e resolução da imagem, porém provoca a diminuição do sinal (apud STROBEL, 2017). Sendo assim, é importante encontrar a melhor combinação para que a resolução da imagem seja ideal sem comprometer o sinal da onda.

2.1.1 Propagação da Onda

Como as imagens são geradas a partir do tempo de propagação da onda, é importante entender como ela é calculada. Segundo Puccini (2002), o equipamento de ultrassom produz um sinal elétrico, pelo transdutor emissor, que atravessa a tora de madeira e é recebido pelo transdutor receptor, podendo ser visualizado em um osciloscópio, o qual permite medir os sinais e calcular o tempo de propagação da

onda, ou seja, quanto tempo a onda demora para chegar ao transmissor receptor. Os transdutores transformam esses sinais elétricos em pulso de ultra-som e, após a recepção pelos transdutores receptores, convertem esse pulso em sinais elétricos novamente para que seja possível realizar a medição pelo osciloscópio. Ainda conforme Puccini (2002), esses cálculos feitos pelo osciloscópio também levam a amplitude e a frequência da onda em consideração.

De acordo com Wang (2013), o tempo de propagação da onda entre os transdutores é definido como TOF, ou *time of flight*. A partir da tomografia realizada pelo equipamento USLab e a medição do TOF pelo osciloscópio, é possível calcular a velocidade acústica através da Equação 2.1.

$$C_T = \frac{S}{\Delta t} \quad (2.1)$$

Na Equação 2.1, C_T representa a velocidade acústica (m/s), S é o comprimento do percurso entre os transdutores emissor e receptor e Δt é o TOF. (WANG, 2013).

A propagação da onda está diretamente relacionada com a qualidade do objeto, nesse caso, a madeira. As madeiras de melhor qualidade em relação às deterioradas ou com menos rigidez, possuem propagação mais rápida, segundo Wang (2004, apud STROBEL, 2017). Isso acontece, pois, as ondas contornam as regiões defeituosas, fazendo com que haja um aumento no tempo da propagação entre os transdutores e diminuindo a velocidade. (STROBEL, 2017). A Figura 2.3 mostra a representação da propagação da onda em madeiras com oco.

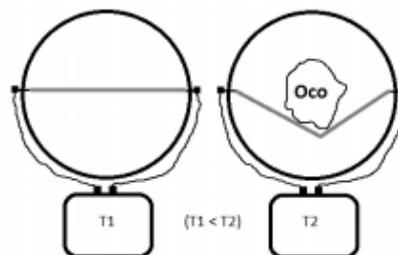


Figura 2.3: Esquema da propagação da onda e do aumento do tempo em função de uma anomalia na madeira. Fonte: (STROBEL, 2017)

Segundo Bucur e Böhnke (1994), as propriedades físicas da madeira, suas características geométricas, as condições do meio e o procedimento utilizado também são grandes influenciadores na propagação das ondas.

2.2 Propriedades da madeira

As propriedades da madeira podem variar de acordo com o sistema de referência, ou seja, de acordo com a direção (STROBEL, 2017). Assim, a propagação da onda se modifica dependendo do referencial utilizado no ensaio de ultrassom.

Conforme Strobel (2017), os três eixos principais na madeira são:

- a) longitudinal: paralelo às fibras longitudinais;
- b) radial: perpendicular às fibras longitudinais e perpendicular ao plano dos anéis de crescimento;
- c) tangencial: perpendicular às fibras longitudinais e tangente ao plano dos anéis de crescimento.

Estes eixos podem ser melhor visualizados na Figura 2.4, retirada do trabalho de Strobel (2017).

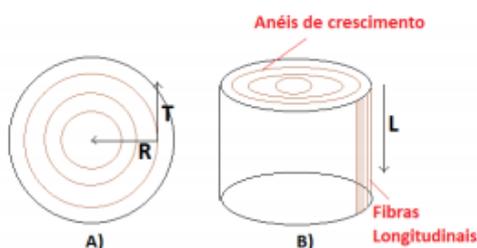


Figura 2.4: a) Seção transversal da madeira e os eixos radial e tangencial. b) Ilustração do eixo longitudinal paralelo às fibras longitudinais. Fonte: (STROBEL, 2017)

2.3 Interpolação espacial

No contexto deste trabalho, ultrassom em madeira, usa-se a interpolação espacial para estimar valores das velocidades acústicas em pontos desconhecidos de uma seção transversal plana (STROBEL, 2017).

Segundo Stark (1979), a interpolação espacial é um método matemático para encontrar valores intermediários entre valores discretos (apud Mazzini, Schettini, 2009). Em outras palavras, a interpolação serve para descobrir valores desconhecidos de um conjunto, levando em consideração os valores já conhecidos e calculados.

No contexto deste trabalho, a interpolação espacial foi utilizada através dos valores de tempo da propagação de ondas, gerados pós ultrassom em toras de madeira, para estimar os valores das velocidades acústicas em pontos desconhecidos. A Figura 2.5 exemplifica um ponto com velocidade desconhecida por entre pontos com velocidades conhecidas em uma seção transversal da madeira.

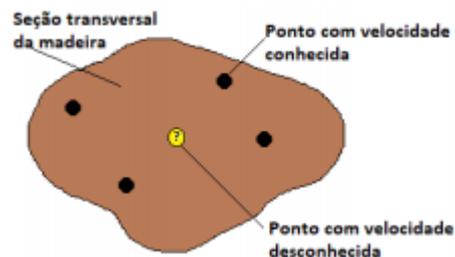


Figura 2.5: Exemplo de um ponto com velocidades conhecidas e desconhecidas em uma tora de madeira. Fonte: (STROBEL, 2017)

No trabalho realizado por Strobel (2017), foram utilizadas as técnicas de interpolação espacial inverso do quadrado da distância (IQD) e interpolação por zonas afetadas, após o ensaio de ultrassom nas toras de madeira. De acordo com Marcuzzo, Andrade & Melo (2011), o método de interpolação espacial IQD determina os valores desconhecidos dos pontos usando uma combinação linear ponderada. O peso de cada ponto é o inverso da distância (apud LUCAS, 2012). O segundo método utilizado, interpolação por zonas afetadas, conforme Strobel (2017), usa diretamente a velocidade das rotas do ultrassom. Assim, para cada rota, relaciona-se uma zona em forma elíptica. Deste modo, se um ponto com valor desconhecido estiver dentro da zona elíptica, este recebe o valor da rota associada.

2.4 Algoritmos de agrupamento

Para solucionar a dificuldade em encontrar a similaridade entre os elementos de um conjunto, foram utilizados os algoritmos de agrupamento ou clusterização, os quais buscam agrupar elementos de um conjunto, baseado nas similaridades e dissimilaridades entre os elementos (TAKAHASHI, BEDREGAL, LYRA, 2005). Conforme Tronco (2015), os métodos de agrupamento podem ser utilizados em várias aplicações, como reconhecimento de padrões e segmentação de imagens.

De acordo com Ferreira (2012), existem algumas abordagens de técnicas de agrupamento, e as mais conhecidas são as hierárquicas e as particionais. Essas duas técnicas são métodos não supervisionados que, segundo Borges (2010), “procuram determinar e identificar automaticamente como os dados estão organizados em um conjunto ou em uma base de dados” (p. 1).

2.4.1 Algoritmos Hierárquicos

Os algoritmos hierárquicos consistem em criar uma hierarquia dos grupos e, conforme Ferreira (2012), considera inicialmente que o conjunto inteiro de dados faz parte de apenas um grupo e, com o progresso do algoritmo, este grupo é quebrado em partes menores. O inverso também é considerado, na qual cada ponto é considerado um grupo e, com o progresso do algoritmo, os pontos são agrupados em grupos maiores (FERREIRA, 2012). Esses algoritmos podem ser representados por um dendograma que, segundo Ferreira (2012), “é um tipo de árvore que exhibe a ordem a qual os grupos foram fundidos, (...) ou a ordem em que foram divididos” (p. 13), dependendo do algoritmo utilizado. O exemplo de um dendograma pode ser visto na Figura 2.6.

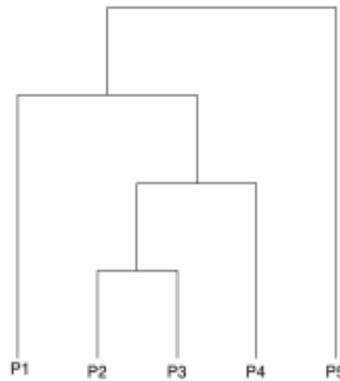


Figura 2.6: Dendograma. Fonte: (FERREIRA, 2012)

A Figura 2.6, retirada do trabalho de Ferreira (2012), mostra um exemplo de dendograma resultado de um algoritmo hierárquico usando um dataset que contém cinco pontos, ou seja, cinco elementos no conjunto.

2.4.2 Algoritmos Particionais

Nos algoritmos particionais, segundo Ferreira (2012), os dados são divididos em um número K de grupos. Os algoritmos K-Means e Fuzzy C-Means, utilizados neste trabalho, são exemplos de algoritmos particionais.

De acordo com Vale (2005), esse método se baseia em diminuir uma função de custo, assim, os dados são agrupados de acordo com o menor valor para esta função. A vantagem desses métodos em relação aos métodos Hierárquicos, citado na seção 2.4.1, é que os padrões podem mudar de agrupamento com o progresso do algoritmo, sem contar que é possível atuar utilizando um dataset maior. Por outro lado, alguns algoritmos dos métodos Particionais, como o K-Means, necessitam que a escolha do número de grupos seja feita inicialmente, podendo gerar diversas perspectivas sobre uma mesma base de dados (VALE, 2005).

2.4.3 Medidas de Similaridade

Para utilização dos algoritmos de agrupamento, é necessário usar uma medida de similaridade entre os elementos do conjunto a ser agrupado. As medidas de similaridade e dissimilaridade, também conhecidas como medidas de distância, são usadas nos algoritmos de agrupamento como forma de medir a distância entre

um ponto e o cluster ou entre os pontos. Segundo Tronco (2015), uma das medidas de similaridade mais utilizadas é a distância Euclidiana, que pode ser descrita de acordo com a Equação 2.2.

$$d = \sqrt{(X_A - X_B)^2 + (Y_A - Y_B)^2} \quad (2.2)$$

Neste trabalho, os algoritmos K-Means e Fuzzy C-Means foram implementados usando a distância Euclidiana como medida de dissimilaridade.

2.5 Lógica Fuzzy

Alguns algoritmos de agrupamento podem ser baseados na lógica Fuzzy, como é o caso do Fuzzy C-Means. A lógica Fuzzy é fundamentada nas teorias de conjuntos Fuzzy que, segundo Vargas (2012), permite resolver o problema dos paradoxos gerados com a classificação “verdadeiro ou falso”, em que uma premissa pode variar em grau de verdade, sendo parcialmente verdadeiro ou parcialmente falso ao mesmo tempo.

É importante entender o conceito de grau de pertinência nas técnicas Fuzzy. Vargas (2012) diz que a função de pertinência estabelece uma relação entre um elemento de um conjunto e um valor entre 0 e 1, indicando o grau de pertinência do elemento entre cada grupo. Assim, quanto mais próximo de 1 estiver o valor de pertinência, maior será o pertencimento deste elemento ao grupo relacionado. Além disso, as metodologias Fuzzy têm uma restrição em que a soma dos graus de pertinência de um elemento a cada grupo deve ser igual a 1 (VALE, 2005).

Conforme Tronco (2015), a função de pertinência de um conjunto A no Universo U, pode ser representada conforme a Equação 2.3.

$$\mu_A: U \rightarrow [0,1] \quad (2.3)$$

De acordo com Tronco (2015), as funções de pertinência mais utilizadas são as triangulares, gaussianas e trapezoidais. A Figura 2.7, retirada do trabalho de Tronco (2015), mostra um exemplo de função de pertinência triangular.

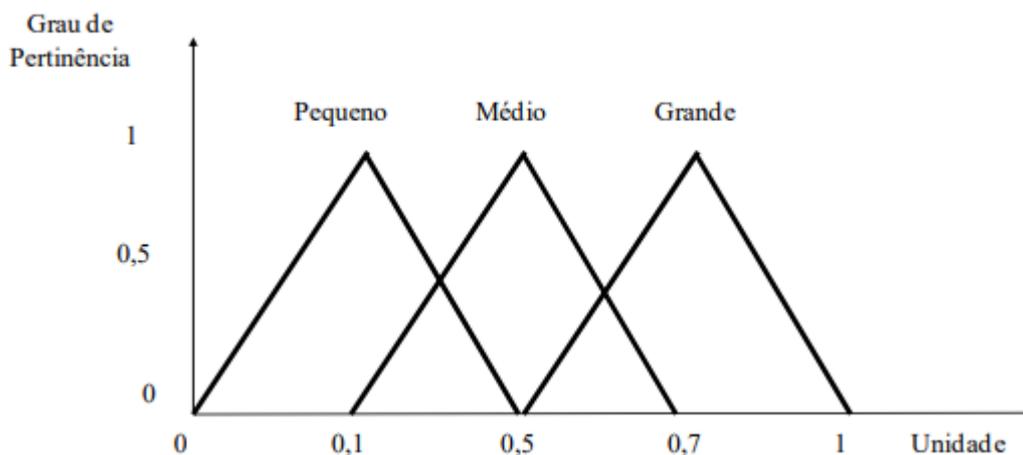


Figura 2.7: Gráfico de função de pertinência triangular. Fonte: (TRONCO, 2015)

Neste exemplo, é possível notar que, se a unidade está em 0,1, o grau de pertinência para “Pequeno” é 1, e para “Médio” é 0. Conforme a unidade aumenta, o grau de pertinência para “Pequeno” diminui e para “Médio” aumenta, diferenciando-se assim da lógica booleana, em que um elemento pertence ou não a um conjunto (TRONCO, 2015).

2.6 Segmentação de Imagem

Este trabalho irá utilizar os métodos de agrupamento, citados nas seções anteriores, para segmentação das imagens geradas após a interpolação espacial dos dados obtidos pelo ultrassom em toras de madeira.

Em diversos casos que necessitam de processamento de imagens, a segmentação serve como uma ferramenta utilizada para diversos fins, como reconhecimento facial, identificação de regiões, entre outras utilidades. (BORGES, 2010).

Segundo Borges (2010), a imagem é subdividida em regiões, com o objetivo de facilitar a análise, simplificando sua representação. No contexto desse trabalho, simplificar a representação das imagens é vantajoso para destacar possíveis anomalias no interior da madeira. Conforme a Figura 2.8, é possível visualizar o destaque de uma anomalia, usando como exemplo a imagem de uma das toras de madeira utilizadas nesse trabalho. A Figura 2.8 apresenta a imagem original (A) e a

imagem pós segmentação (B), utilizando o algoritmo Fuzzy C-Means com 4 clusters.

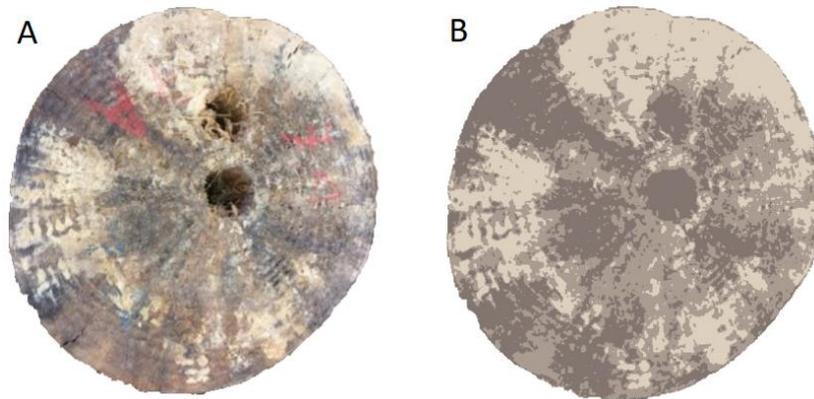


Figura 2.8: Disco 7 pós segmentação utilizando algoritmo Fuzzy C-Means com 4 clusters.

De acordo com Aly et al. (2011), para segmentar a imagem, os pixels devem apresentar alguma similaridade entre eles (apud SIEBRA, 2013). Nos algoritmos usados, podem ser utilizados diversos tipos de medidas de similaridades, como a distância entre os pixels, no caso deste trabalho.

3 Materiais e Métodos

Este capítulo irá utilizar a fundamentação teórica e trabalhos relacionados como embasamento para apresentar os materiais e métodos utilizados na construção desta monografia.

3.1 Dataset

Nesta seção serão apresentadas as toras de madeiras utilizadas na tomografia ultrassônica, sendo algumas delas com ocos artificiais, outras com deteriorações próprias.

A Tabela 3.1 faz uma breve descrição das toras de madeira utilizadas nesse estudo. Ela foi retirada do trabalho de Junior et al. (2018) e adaptada pela autora.

Tipo de madeira	Descrição
Disco 7	Dois orifícios artificiais de 50mm de diâmetro cada
Disco 10	Um orifício artificial de 50mm de diâmetro
Disco Liquidambar Styraciflua	Perto da medula existe uma pequena área com um estágio inicial de decomposição por fungos, além de rachaduras laterais da medula à casca
Disco Platanus	A maior parte da madeira mostra sinais de ataques de fungos e existem algumas áreas vazias causadas por cupins

Tabela 3.1: Descrição das toras de madeira. Fonte: (JUNIOR et al, 2018). Adaptado pela autora.

As imagens das toras de madeira originais estão representadas na Figura 3.1. A primeira tora de madeira na Figura 3.1 se refere ao Disco 7 (A), a segunda ao Disco 10 (B), a terceira ao Disco Liquidambar Styraciflua (C) e a última ao Disco Platanus (D).

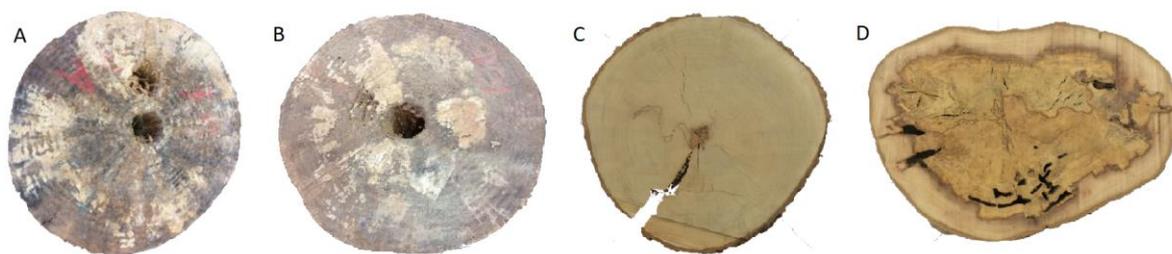


Figura 3.1: Imagens das toras de madeira. Fonte: (JUNIOR et al, 2018). Adaptado pela autora.

A partir das toras de madeira representadas na Figura 3.1, foi realizado o ensaio de ultrassom, que gerou resultados dos tempos de propagação da onda, ou TOF. Esses dados foram utilizados no método de Interpolação IDW, que estão representados nas Figuras 3.2, 3.3, 3.4 e 3.5.

Essas imagens serão utilizadas como entrada dos algoritmos K-Means e Fuzzy C-Means. Elas estão organizadas em: (A) $p=1$ e (B) $p=2$. Os valores de p serão melhor explicados posteriormente, na seção 3.3

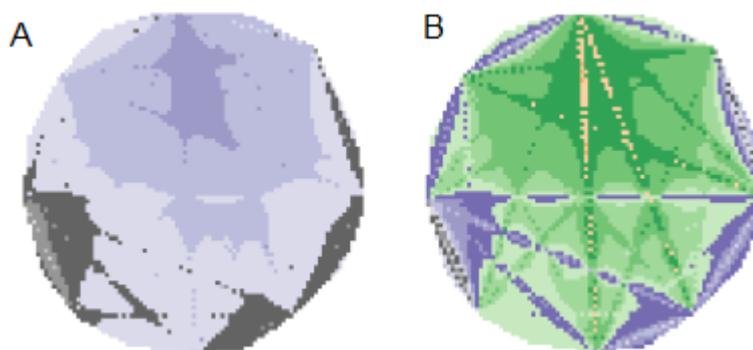


Figura 3.2: Imagem resultante da interpolação IDW para o Disco 7. Fonte: (JUNIOR et al, 2018). Adaptado pela autora.

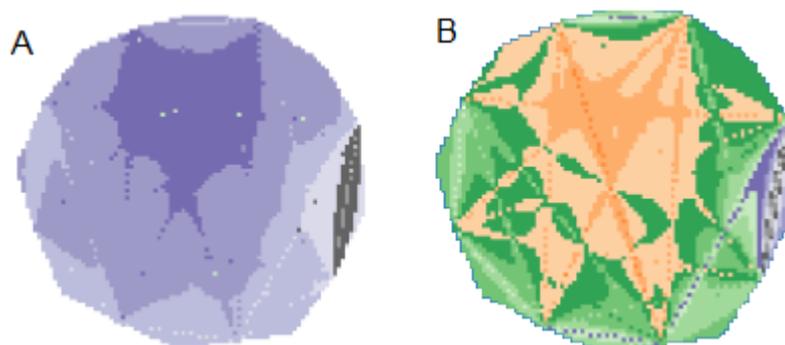


Figura 3.3: Imagem resultante da interpolação IDW para o Disco 10. Fonte: (JUNIOR et al, 2018). Adaptado pela autora.

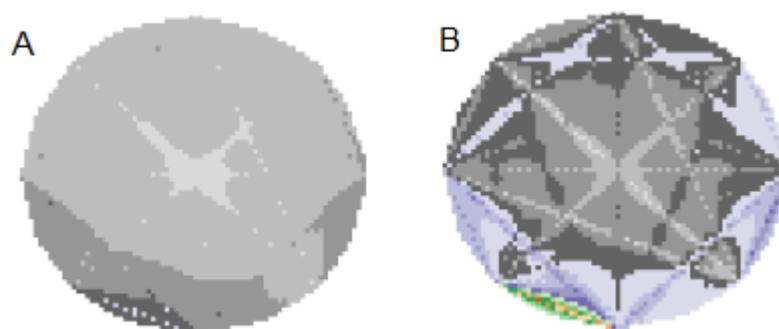


Figura 3.4: Imagem resultante da interpolação IDW para o Disco Liquidambar Styraciflua. Fonte: (JUNIOR et al, 2018). Adaptado pela autora.

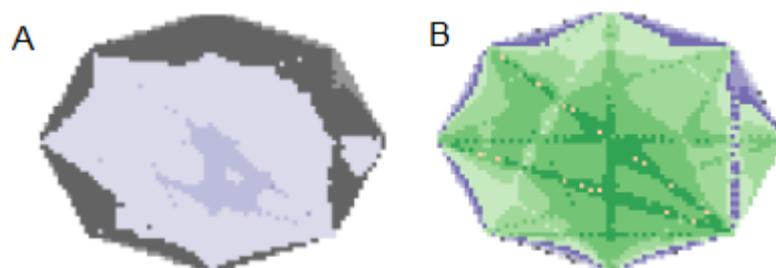


Figura 3.5: Imagem resultante da interpolação IDW para o Disco Platanus. Fonte: (JUNIOR et al, 2018). Adaptado pela autora.

Segundo Junior et al (2018), as cores das imagens representam a quantidade de faixas de velocidade, sendo assim, as cores mais escuras fazem a representação de uma quantidade maior de velocidade das ondas e, as mais claras, são faixas de velocidades mais lentas.

3.2 Ultrassom em madeira

Para realizar um ensaio de ultrassom, deve-se posicionar os transdutores nas bordas do objeto, medir o tempo de propagação das ondas, ou seja, quanto tempo a onda leva para percorrer o caminho entre o transdutor emissor e o transdutor receptor, determinando assim, a velocidade que a onda se propaga no objeto.

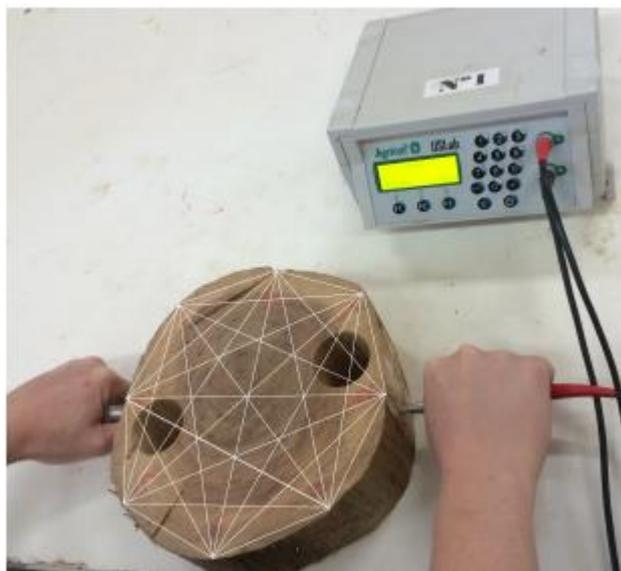


Figura 3.6: Ensaio de ultrassom em tora de madeira. As linhas em branco representam possíveis percursos das ondas entre os transdutores. Fonte: (STROBEL, 2017)

Nos experimentos realizados pelo LabEND/UNICAMP, oito transdutores foram utilizados em volta das toras de madeira, totalizando assim, oito pontos de medição. Com essa configuração e distribuição dos sensores nas toras, o experimento gerou 28 rotas de ondas ultrassônicas (STROBEL, 2017). A Figura 3.6 apresenta a imagem de um ultrassom realizado com as possíveis rotas das ondas, a partir do ponto de medição.

De acordo com Strobel (2017), após a medição, faz-se uma análise para verificar quais rotas terão maior significância para fazer as interpolações dos pontos desconhecidos. Os ensaios de ultrassom resultam em dados do tempo de propagação das ondas, nas quais serão utilizadas como entrada do algoritmo de interpolação espacial (JUNIOR et al, 2018).

3.3 Interpolação Espacial

Após realizada a tomografia ultrassônica nas toras de madeira, é utilizado o método de interpolação IDW, ou Inverso da Potência das Distâncias, para calcular os valores desconhecidos dos resultados do ultrassom. O método de interpolação espacial IDW, presume que os pontos mais próximos entre si possuem valores mais parecidos do que os pontos mais distantes, colocando então maior peso para os valores ao redor do ponto desconhecido. Assim, cada ponto possui uma influência que diminui à medida que a distância aumenta (JAKOB, YOUNG, 2006).

Segundo Junior et al (2018), o primeiro passo do método IDW é calcular a distância dos pontos usando a distância Euclidiana, que foi descrita na Seção 2.4.3, na Equação 2.2.

Após isso, é feito o cálculo do inverso da potência da distância, utilizando a Equação 3.1.

$$f(X) = \frac{\sum_{i=1}^N (d_i)^p \cdot Z_i}{\sum_{i=1}^N (d_i)^p}; d_i > 0, p > 0$$

$$Z_i; \quad d_i = 0$$
(3.1)

Na Equação 3.1, d_i é um ponto com valor calculado, Z_i é o valor do ponto, X é o ponto que será interpolado, d_i é a distância cartesiana entre X e d_i e N é o número de pontos que têm seus valores conhecidos. (JUNIOR et al, 2018). Neste trabalho, o valor de p será 1 e 2. Ainda de acordo com Junior et al. (2018), esta função garante maior influência aos pontos que estão mais perto do ponto com valor desconhecido.

A grande diferença entre o IDW e o IQD, citado na Seção 2.3, é que o valor de p pode variar de acordo com a escolha de quem utilizará o método. No método IQD, o valor de p sempre será 2.

As imagens utilizadas neste trabalho foram geradas no trabalho de Junior et al (2018), como resultado das interpolações do método IDW. Posteriormente, foram usadas como entrada nos algoritmos de agrupamento, K-Means e Fuzzy C-Means.

3.4 Método Elbow

O método de Elbow foi utilizado nesta monografia para encontrar o número ideal de clusters para cada base de dados. Segundo Kodinariya (2013), o método Elbow é a técnica mais antiga para determinar o número de clusters em algoritmos de agrupamento. É um método visual, em que o número de clusters k começa como 2 e vai aumentando em uma unidade a cada iteração. O método tem como objetivo analisar a porcentagem de variância da soma dos quadrados entre o número de clusters e em um gráfico, mostrar a diferença na variância entre os diferentes números de clusters. (KODINARIYA, 2013). Com esse gráfico, é possível notar que, em alguns casos, adicionar mais clusters não interfere na variância dos dados. É importante ressaltar que o método foi implementado usando o K-means como função principal para a escolha dos clusters.

A Figura 3.7 mostra o resultado gerado com o método de Elbow, utilizando um dos datasets deste trabalho. O método de Elbow recebe esse nome pois, no gráfico, onde a linha forma um “cotovelo”, é o lugar que a variância estabiliza. Assim, adicionando mais clusters ao agrupamento, a diferença nos dados será pouca.

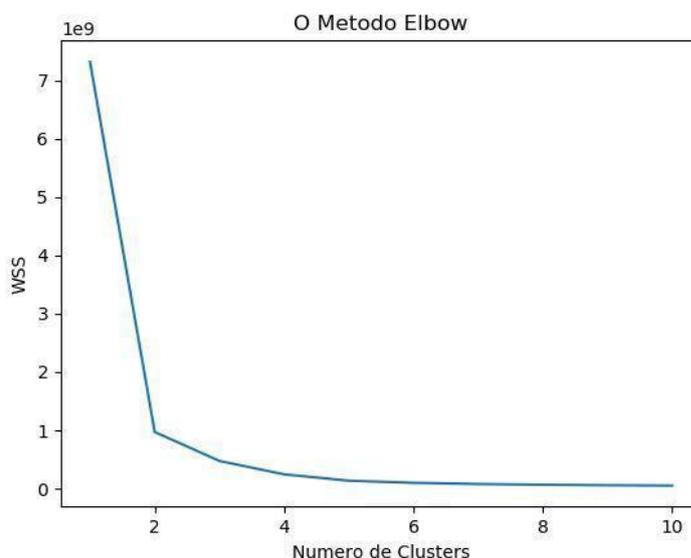


Figura 3.7: Gráfico obtido utilizando o Método de Elbow.

É possível notar na Figura 3.7 que o método inicia com dois clusters para verificar a variância dos dados utilizando até 10 clusters. Considerando este gráfico, o método retorna que, após 3 clusters, a variância dos dados começa a estabilizar.

Assim, não há diferenças significativas com o aumento do número de clusters. No gráfico da Figura 3.7, WCSS significa *Within-Cluster-Sum-of-Squares*, ou soma de quadrados dos clusters.

3.5 K-means

O algoritmo K-means, de acordo com Ferreira (2012), é um dos métodos de agrupamento mais conhecidos. Segundo Vale (2005), este é um algoritmo que minimiza a soma das distâncias de cada ponto ao centróide que, de acordo com Coletta (2011), “são os pontos médios de cada grupo” (p. 7). No caso deste trabalho, foi usado a distância Euclidiana para calcular as distâncias entre os pontos. Segundo Takahashi (2005), a distância entre o ponto e o centróide é o que define a qual grupo o ponto irá pertencer.

De acordo com Coletta (2011), o pseudocódigo do algoritmo K-Means pode ser definido como:

1. Usuário escolhe o número k de grupos
2. Algoritmo seleciona os centróides iniciais aleatoriamente
3. A distância entre os elementos e os centroides é calculada
4. É atribuído cada elemento ao grupo de centróide mais próximo
5. Centróides são recalculados de acordo com seus elementos
6. Algoritmo é finalizado após atingir o número de iterações máximo ou se algum critério de parada, determinado pelo usuário, for cumprido. Caso contrário, o algoritmo volta ao passo 3 e é iniciado novamente.

No contexto de segmentação de imagens, as imagens são particionadas por pixels agrupados, que possui similaridades entre eles. Estas relações podem ser feitas a partir do nível de cinza de cada pixel, textura, distância entre os pixels, entre outras (TAKAHASHI, BEDREGAL, LYRA, 2005).

3.6 Fuzzy C-Means

O algoritmo Fuzzy C-Means é uma extensão do algoritmo descrito na seção anterior, o K-Means. (COLETTA, 2011). Além disso, ele é resultante da combinação da lógica Fuzzy com a teoria de agrupamento de dados e é muito utilizado para clusterização e reconhecimento de padrões (TRONCO, 2015).

Sabendo que a lógica Fuzzy permite valores intermediários entre 0 e 1 (verdadeiro e falso), ela diferencia-se dos métodos de agrupamento tradicionais, como o K-means, em que o pertencimento do elemento é exclusiva a um grupo (verdadeiro ou falso). Assim, o Fuzzy C-Means mostra-se mais flexível em relação aos algoritmos de agrupamento tradicionais. (TRONCO, 2015).

Para encontrar os grupos para um conjunto de dados, o algoritmo FCM busca minimizar as distâncias entre os dados e os centros dos grupos aos quais os dados pertencem, levando em consideração o grau de pertinência (ROCHA *et al.*, 2012).

Nas Equações 3.2, 3.3 e 3.4, n é o número de dados, p é o número de clusters inicialmente escolhido pelo usuário, $m > 1$ é o parâmetro de fuzzificação, que usualmente está no intervalo $[1,25;2]$, segundo COX (2005, apud Vargas, 2012), x_i é um vetor de dados que representa um atributo do dado, c_j é o centro do agrupamento ($j = 1, 2, \dots, p$) e $d(x_i; c_j)$ é a distância entre x_i e c_j (VARGAS, 2012).

Segundo Vargas (2012), o método FCM busca minimizar a Equação 3.2.

$$J = \sum_{i=1}^n \sum_{j=1}^p \mu_{ij}^m d(x_i; c_j)^2 \quad (3.2)$$

Diferentemente do método K-Means, em que a posição inicial dos centroides é escolhida aleatoriamente, segundo Vargas (2012), no FCM os centros dos clusters são calculados de acordo com a Equação 3.3.

$$c_j = \frac{\sum_{i=1}^n \mu_{ij}^m x_i}{\sum_{i=1}^n \mu_{ij}^m} \quad (3.3)$$

Ainda de acordo com Vargas (2012), a função de pertinência é calculada através da Equação 3.4.

$$\mu_{ij} = \frac{\left(\frac{1}{d(x_i; c_j)}\right)^{\frac{2}{m-1}}}{\sum_{k=1}^p \left(\frac{1}{d(x_i; c_k)}\right)^{\frac{2}{m-1}}} \quad (3.4)$$

De acordo com Coletta (2011), o pseudocódigo do algoritmo Fuzzy C-Means pode ser descrito como:

1. Usuário escolhe o número c de grupos
2. Algoritmo seleciona os centros iniciais através da Equação 3.2
3. As distâncias entre os pontos e os centros são calculadas
4. É calculada a matriz de pertinência
5. Os centros são recalculados de acordo com as pertinências
6. Algoritmo é finalizado após atingir o número de iterações máximo ou se algum critério de parada, determinado pelo usuário, for cumprido. Caso contrário, o algoritmo volta ao passo 3 e é iniciado novamente.

4 Resultados

Este capítulo apresentará os resultados obtidos neste trabalho, com a aplicação dos algoritmos de agrupamento citados anteriormente, o K-Means e Fuzzy C-Means. Ao final do capítulo, será feita a análise comparativa entre os algoritmos de agrupamento, com o objetivo de identificar qual dos métodos teve um desempenho melhor na representação das anomalias.

4.1 Leitura das imagens

Os algoritmos utilizados neste trabalho foram desenvolvidos em Python, com auxílio de algumas bibliotecas. A primeira delas foi a *imageio*, usada para fazer a leitura de imagens digitais com formato PNG. De acordo com a documentação da biblioteca, a função *imread* recebe como parâmetro a imagem escolhida pelo usuário e a transforma em um *array*. Se a imagem for em escala de cinza, a função retorna um *array* 2D. Caso seja uma imagem colorida, a função retorna um *array* 3D, com as cores RGB de cada pixel da imagem. Neste trabalho, como as imagens são coloridas, a função retornou um *array* 3D.

4.2 Geração das imagens

Para a geração das imagens, foi utilizada a biblioteca *matplotlib*. A documentação da biblioteca mostra que é permitido inserir, como parâmetros da plotagem dos gráficos, a paleta de cores escolhida pelo usuário. Neste trabalho, a plotagem dos gráficos utilizará as próprias cores das imagens originais.

4.3 K-Means

A implantação do algoritmo K-Means foi feita utilizando a biblioteca *sklearn*, na qual possui a classe *KMeans*. Segundo a documentação da biblioteca, a classe *KMeans* pode receber como parâmetro o número de clusters, o método de inicialização (*random*, *number of array* ou *kmeans++*), o número de iterações, entre outros. No caso deste trabalho, o único parâmetro passado pelo usuário foi o número de clusters. O método de inicialização foi utilizado o default (*random*), onde o algoritmo escolhe aleatoriamente a localização dos clusters iniciais. O número de

iterações também foi utilizado o default, que é 300 iterações, ou seja, o processo é repetido 300 vezes antes de se encerrar.

4.4 Fuzzy C-Means

A implantação do algoritmo Fuzzy C-Means foi feita utilizando a biblioteca *scikit*. Esta biblioteca possui uma toolbox com funções implementadas para o uso do Fuzzy C-Means.

A função utilizada recebeu como parâmetro os dados a serem clusterizados, o número de clusters, o número expoente aplicado a função de pertinência, que nesse caso foi 2 e o número máximo de iterações, que é de 300 iterações.

4.5 Experimentos

4.5.1 Definindo o número de clusters pelo Método de Elbow

O primeiro experimento realizado foi a obtenção do número de clusters pelo método de Elbow. Todas as imagens foram inseridas como entrada do algoritmo de Elbow, para encontrar o melhor número de clusters para cada uma delas. Como o resultado foi variado e o objetivo é obter o melhor resultado, este trabalho usará 3, 4 e 5 clusters em ambos os algoritmos e em todas as imagens do dataset. De qualquer forma, a Tabela 4.1 mostra o número de clusters obtido para cada imagem.

Imagem	Nº de Clusters
Disco 7 com $p=1$	3
Disco 7 com $p=2$	5
Disco 10 com $p=1$	4
Disco 10 com $p=2$	4
Disco Liquidambar Styraciflua com $p=1$	3
Disco Liquidambar Styraciflua com $p=2$	6
Disco Platanus com $p=1$	3
Disco Platanus com $p=2$	5

Tabela 4.1: Número de clusters obtido através do método de Elbow para cada imagem do dataset.

4.5.2 Agrupamento via técnica K-Means

O segundo experimento consiste em passar todas as imagens do dataset no algoritmo K-Means.

Neste experimento, as imagens serão passadas usando o método de interpolação com $p=1$ e $p=2$. No método K-Means, os números de clusters serão $k=3$, $k=4$ e $k=5$. Apesar do método de Elbow ter escolhido apenas um valor para k , com o objetivo de melhor análise do resultado, foi feito o mesmo experimento com diversos valores de k .

As Figuras 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7 e 4.8 estão organizadas em: imagem original, (A) 3 clusters, (B) 4 clusters e (C) 5 clusters.

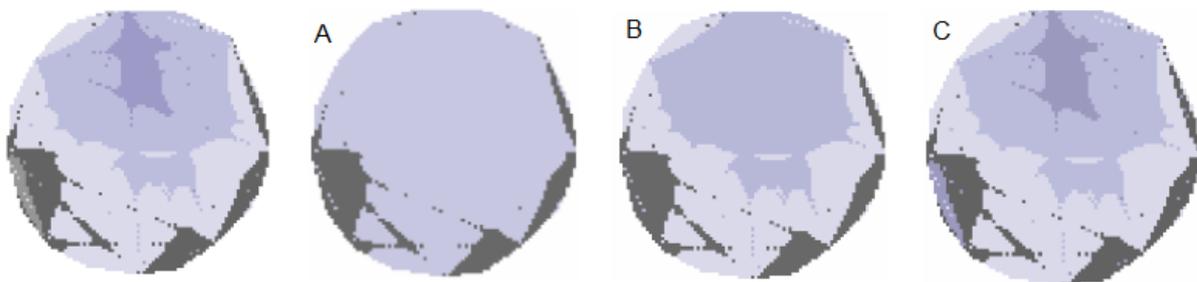


Figura 4.1: Segmentação da imagem pós interpolação IDW com $p=1$ do Disco 7 utilizando K-Means

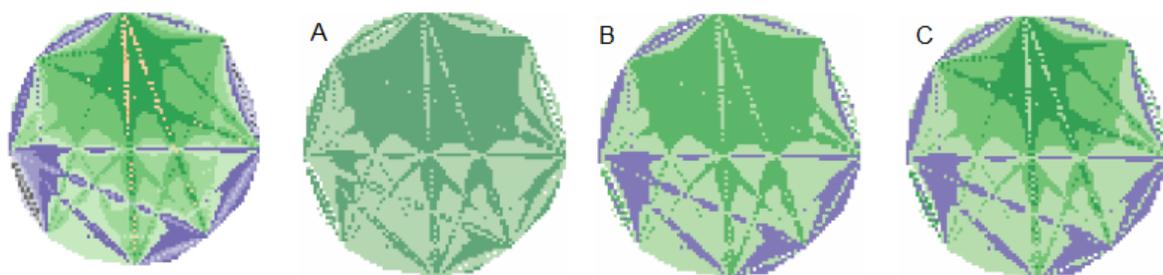


Figura 4.2: Segmentação da imagem pós interpolação IDW com $p=2$ do Disco 7 utilizando K-Means

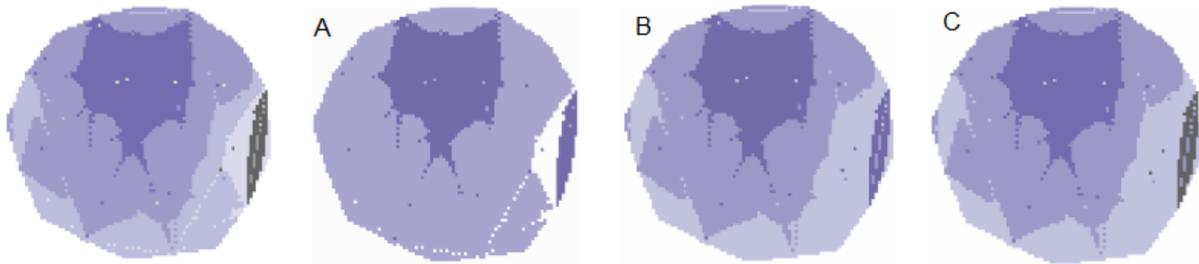


Figura 4.3: Segmentação da imagem pós interpolação IDW com $p=1$ do Disco 10 utilizando K-Means

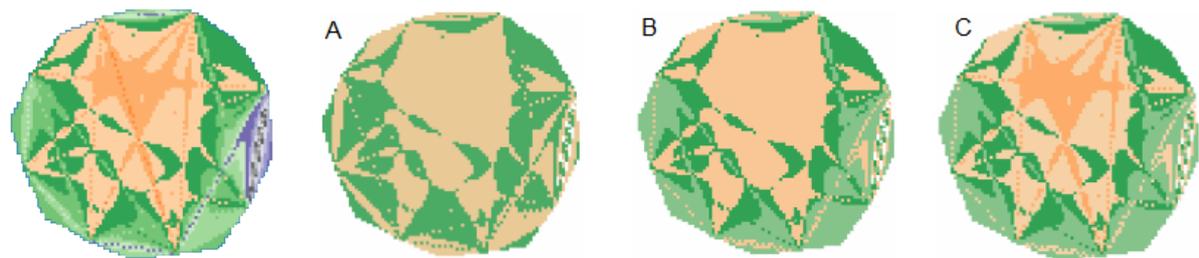


Figura 4.4: Segmentação da imagem pós interpolação IDW com $p=2$ do Disco 10 utilizando K-Means

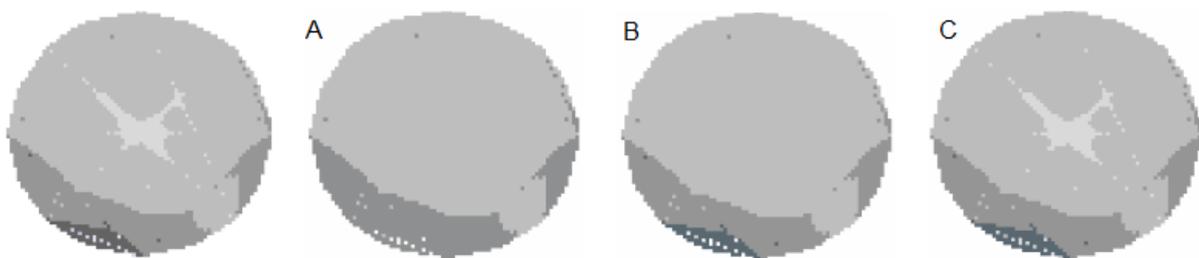


Figura 4.5: Segmentação da imagem pós interpolação IDW com $p=1$ do Disco Liquidambar Styraciflua utilizando K-Means

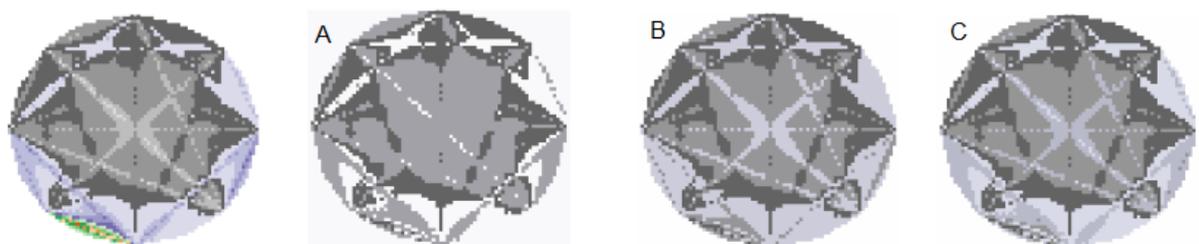


Figura 4.6: Segmentação da imagem pós interpolação IDW com $p=2$ do Disco Liquidambar Styraciflua utilizando K-Means

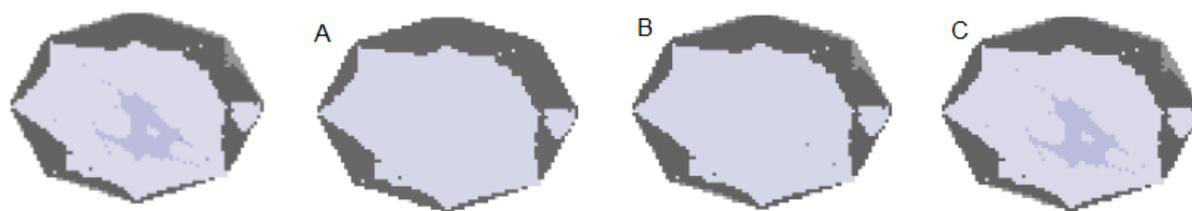


Figura 4.7: Segmentação da imagem pós interpolação IDW com $p=1$ do Disco Platanus utilizando K-Means

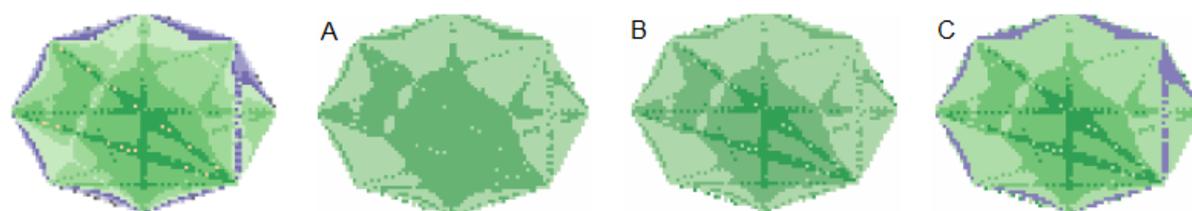


Figura 4.8: Segmentação da imagem pós interpolação IDW com $p=2$ do Disco Platanus utilizando K-Means

Após o uso do algoritmo K-Means nas imagens, a maioria dos resultados só conseguiu evidenciar qualquer anomalia utilizando 5 clusters. Na Figura 4.3, o algoritmo teve dificuldades de encontrar similaridades entre as cores dos pixels, fazendo com que o fundo branco da imagem fosse considerado e manchas brancas apareceram dentro da imagem segmentada com 3 clusters. Conclui-se que o algoritmo se mostrou eficiente apenas com o uso de 5 clusters, trazendo um resultado mais parecido com a imagem original.

4.5.3 Agrupamento via técnica Fuzzy C-Means

O terceiro experimento consiste em passar todas as imagens do dataset no algoritmo Fuzzy C-Means. Como no experimento anterior, as imagens também serão agrupadas usando $c = 3$, $c = 4$ e $c = 5$.

As Figuras 4.9, 4.11, 4.12, 4.13, 4.14, 4.15 e 4.16 estão organizadas em: imagem original, (A) 3 clusters, (B) 4 clusters e (C) 5 clusters.

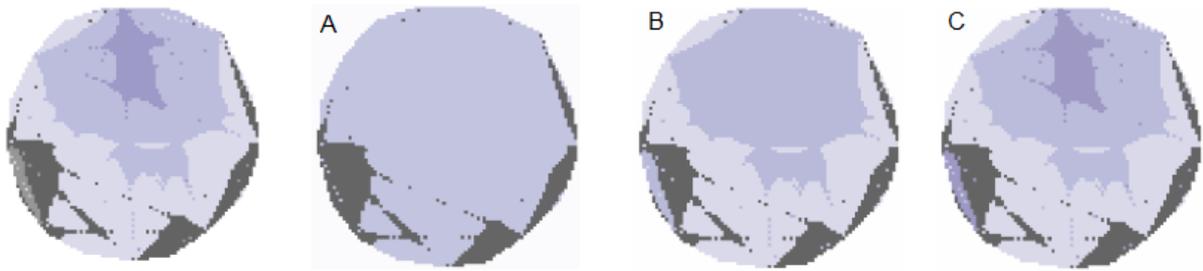


Figura 4.9: Segmentação da imagem pós interpolação IDW com $p=1$ do Disco 7 utilizando Fuzzy C-Means

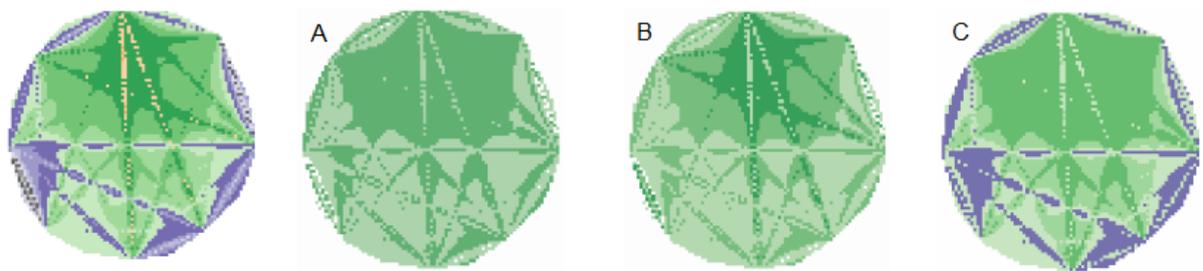


Figura 4.10: Segmentação da imagem pós interpolação IDW com $p=2$ do Disco 7 utilizando Fuzzy C-Means

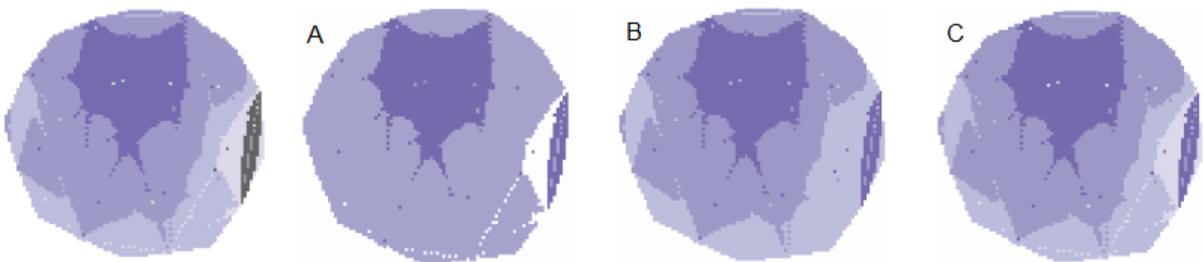


Figura 4.11: Segmentação da imagem pós interpolação IDW com $p=1$ do Disco 10 utilizando Fuzzy C-Means

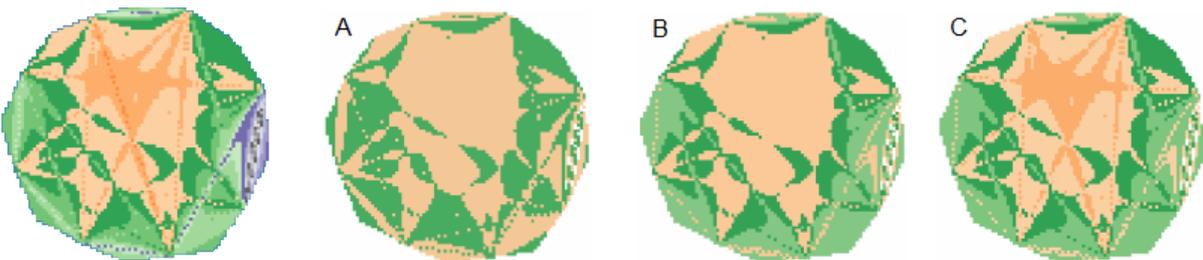


Figura 4.12: Segmentação da imagem pós interpolação IDW com $p=2$ do Disco 10 utilizando Fuzzy C-Means

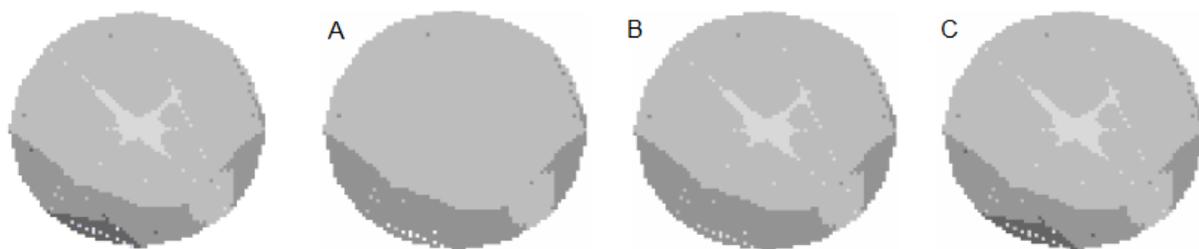


Figura 4.13: Segmentação da imagem pós interpolação IDW com $p=1$ do Disco Liquidambar Styraciflua utilizando Fuzzy C-Means

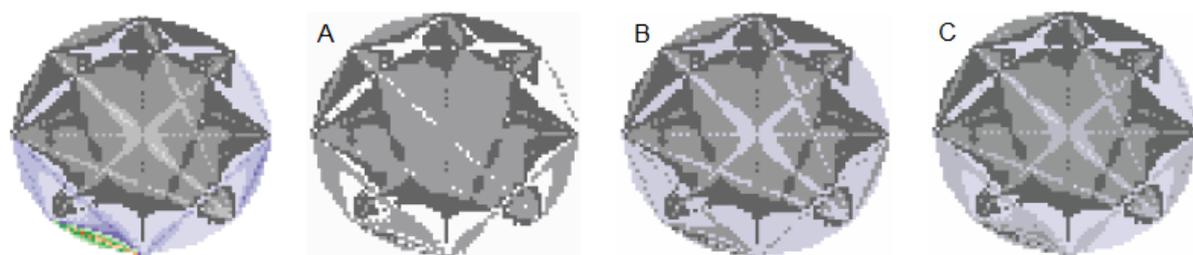


Figura 4.14: Segmentação da imagem pós interpolação IDW com $p=2$ do Disco Liquidambar Styraciflua utilizando Fuzzy C-Means

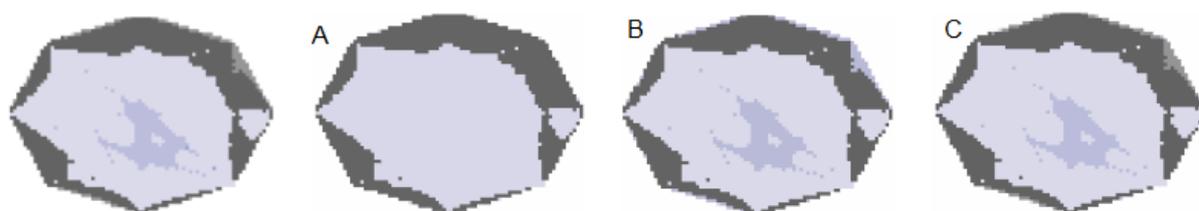


Figura 4.15: Segmentação da imagem pós interpolação IDW com $p=1$ do Disco Platanus utilizando Fuzzy C-Means

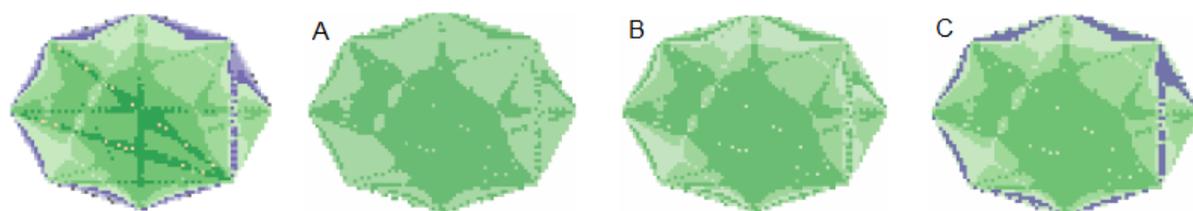


Figura 4.16: Segmentação da imagem pós interpolação IDW com $p=2$ do Disco Platanus utilizando Fuzzy C-Means

Após o uso do algoritmo Fuzzy C-Means nas imagens, a maioria dos resultados só conseguiu evidenciar qualquer anomalia utilizando 5 clusters. Nas

Figuras 4.11 e 4.14, o algoritmo teve dificuldades de encontrar similaridades entre as cores dos pixels, fazendo com que o fundo branco da imagem fosse considerado e manchas brancas apareceram dentro da imagem segmentada com 3 clusters. Nas Figuras 4.13 e 4.15 o algoritmo conseguiu identificar as anomalias utilizando 4 clusters. Conclui-se que o algoritmo se mostrou eficiente, na maioria das vezes, com 5 clusters e em poucos casos, com 4 clusters.

4.6 Análise dos Resultados

Conforme os resultados apresentados nos experimentos da seção 4.5, é possível notar que ambos os algoritmos levaram em consideração os pixels em branco ao fundo das imagens, fazendo que em alguns casos, pixels brancos aparecessem por entre os coloridos no meio das imagens.

Na maioria dos experimentos, foi difícil detectar mudanças consideráveis entre os dois métodos. Com a utilização de 3 clusters, os detalhes se perdem por entre as cores, sendo assim, não é possível identificar nenhuma anomalia nas imagens. Conforme o número de clusters aumenta, é possível perceber os detalhes das cores, mas é importante notar que, nas imagens originais, também são utilizadas poucas cores, todas sólidas, fazendo com que o resultado dos métodos de agrupamento fique mais fiel a imagem original. No método Fuzzy C-Means, o algoritmo encontrou similaridades significantes utilizando 4 clusters, diferentemente do K-Means que, na maioria das vezes, evidenciou as anomalias apenas com 5 clusters. As imagens utilizando 4 e 5 clusters não tiveram diferença significativa na maioria dos casos. Isso se dá pelo fato das imagens originais não ter muitas cores, trazendo assim resultados parecidos conforme o aumento do número de clusters.

É possível dizer que os dois algoritmos, K-Means e Fuzzy C-Means, se comportaram de maneira semelhante, sem diferenças muito evidentes.

5 Conclusões

Este trabalho abordou tópicos importantes para o estudo das propriedades internas de toras de madeira. Para isto, foi feito um levantamento de algumas técnicas computacionais, utilizando métodos de interpolação e métodos de agrupamento, com o propósito de conceber os melhores resultados para visualização de anomalias e deteriorações no interior das madeiras.

O ensaio de ultrassom é um método acústico importante deste segmento, pois consegue extrair informações do interior das toras sem danificar fisicamente o objeto estudado. Após o ensaio de ultrassom nas toras de madeira, que gerou os dados do tempo de propagação das ondas, foi utilizado o método de interpolação IDW, com a finalidade de descobrir os pontos que tiveram seus valores desconhecidos após o ensaio de ultrassom. Feito isso, foi utilizado os métodos de agrupamento não supervisionados e particionais, K-Means e Fuzzy C-Means. O método K-Means é um dos mais antigos e utilizados no contexto de agrupamento de dados. Este método gerou resultados satisfatórios com a utilização de 5 clusters em cada uma das imagens. O método Fuzzy C-Means, baseado na Lógica Fuzzy, é um método que utiliza a função de pertinência para encontrar os melhores grupos de dados, onde cada elemento do conjunto tem um grau de pertinência para cada grupo. Este método obteve resultados significativos, em alguns casos utilizando 4 clusters e, em outros, utilizando 5 clusters.

Como o objetivo deste trabalho é destacar as anomalias usando os métodos de agrupamento, é possível afirmar que ambos os métodos conseguiram demonstrar essas anomalias conforme o número de clusters aumentou, deixando a imagem clusterizada mais fiel às imagens originais. No entanto, não é possível afirmar que os métodos foram decisivos para a visualização das anomalias das toras de madeira. Também é necessário considerar que, tanto as imagens originais, quanto as clusterizadas, não deixam evidente as formas geométricas dos ocos e deteriorações.

O algoritmo K-Medoids foi cogitado para uso na segmentação das imagens, porém, esse método tem muitas semelhanças ao K-Means e, para que houvesse uma diferença considerável entre os resultados, foi descartado.

Por fim, é possível sugerir para trabalhos posteriores, a escolha de outros algoritmos de agrupamentos menos semelhantes no funcionamento ou algoritmos específicos apenas para segmentação de imagens.

Referências Bibliográficas

BORGES, V. R. P. **Comparação entre as técnicas de agrupamento K-Means e Fuzzy C-Means para Segmentação de Imagens Coloridas.** In: Encontro Anual de Computação, 2010, Catalão, Goiás. Anais On-Line do ENACOMP 2010, 2010. v. 1. p. 1-7.

BUCUR, V.; BÖHNKE, I. **Factors affecting ultrasonic measurements in solid wood.** Ultrasonics, v. 32, n. 5, p. 385-390, 1994.

CERRI, G. P., GRAY, G. R., GONÇALVES, R. **Equipamento de ultrassom para materiais não metálicos.** 2010. Faculdade de Engenharia Agrícola, Laboratório de Ensaio Não Destrutivos — LabEND, Universidade Estadual de Campinas, Campinas.

COLETTA, L. F. S. **Agrupamento de dados fuzzy colaborativo.** 2011. Dissertação (Mestrado em Ciências de Computação e Matemática Computacional) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2011

FERREIRA, L. N. **Técnica de agrupamento de dados baseada em redes complexas para o posicionamento de cluster heads em rede de sensores sem fio.** 2012. Dissertação (Mestrado em Ciências de Computação e Matemática Computacional) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2012.

JAKOB, A.A.E., YOUNG A.F. **O uso de métodos de interpolação espacial de dados nas análises sociodemográficas.** Trabalho apresentado no XV Encontro Nacional de Estudos Populacionais, ABEP, realizado em Caxambu, MG, Brasil, 2006. Disponível em: <www.abep.nepo.unicamp.br/encontro2006qdocspdf/ABEP2006_388.pdf>.

JUNIOR, A. A. P. et al. **Um estudo comparativo de técnicas de interpolação espacial de dados para geração de imagens do interior de toras de madeira.** SODEBRAS, 08/2018.

KODINARIYA, T. M., MAKWANA, R. P. **Review on determining number of Cluster in K-Means Clustering.** 2013. International Journal of Advance Research in Computer Science and Management Studies. Volume 1, Issue 6, November 2013 pg. 90-95.

LUCAS, T. de P. B. et al. **O efeito de borda da interpolação IQD em mapas de chuva: uma comparação entre malhas de postos pluviométricos.** REVISTA GEONORTE, [S.l.], v. 3, n. 9, p. 1364 – 1376, out. 2012. ISSN 2237-1419.

MARCUZZO, F. F. N., ANDRADE, L. R., MELO, D. C. R. **Métodos de interpolação matemática no mapeamento de chuvas no estado do Mato Grosso.** Revista Brasileira de Geografia Física, v.4, p.793-804, 2011.

MAZZINI, P. L. F., SCHETTINI, C. A. F. **Avaliação de metodologias de interpolação espacial aplicadas a dados hidrográficos costeiros quase sinóticos.** Brazilian Journal of Aquatic Science and Technology, v.13, n.1, p.53-64, 2009.

OLIVEIRA, F. G. R. de et al. **Efeito do comprimento do corpo-de-prova na velocidade ultra-sônica em madeiras.** Rev. Árvore, Viçosa, v. 30, n. 1, p. 141-145, Feb. 2006.

PERLIN, L. P.; PINTO, R. C. de A. **Tomografia Ultrassônica em Concreto.** RIEM - IBRACON Structures and Materials Journal, [S.l.], v. 6, n. 2, feb. 2013. ISSN 1983-4195.

PUCCINI, C. T. **Avaliação de aspectos de qualidade da madeira utilizando o ultra-som.** 2002. Tese (Doutorado em Engenharia Agrícola) — Faculdade de Engenharia Agrícola, Universidade Estadual de Campinas, Campinas.

RAMIREZ, F. C. **Detecção de danos em estruturas de concreto por meio de tomografia ultrassônica**. 2015. Dissertação (Mestrado em Engenharia de Estruturas) — Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos.

ROCHA, T.; PERES, S. M.; BÍSCARO, H. H.; MADEO, R. C. B; BOSCARIOLI, C. **Tutorial sobre Fuzzy-c-Means e Fuzzy Learning Vector Quantization: abordagens híbridas para tarefas de agrupamento e classificação**. Revista de Informática Teórica e Aplicada, Porto Alegre, v. 19, n. 1, p. 120-163, 2012.

SCIKIT-LEARN. Disponível em: <https://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html>. Acesso em: 25/10/2019.

SECCO, C. B. **Detecção de ocos em toras utilizando métodos de propagação de ondas ultrassônicas**. 2011. Dissertação (Mestrado em Engenharia Agrícola) — Faculdade de Engenharia Agrícola, Universidade Estadual de Campinas, Campinas.

SKFUZZY: 2.0 docs. Disponível em: <https://pythonhosted.org/scikit-fuzzy/api/skfuzzy.html>. Acesso em: 25/10/2019.

STROBEL, J. R. A. **Método de Interpolação Baseado em Elipses Associado à Análise Contextual de Rotas Para Geração de Tomografias Ultrassônicas em Toras de Madeira**. 2017. Dissertação (Mestrado em Tecnologia) — Faculdade de Tecnologia, Universidade Estadual de Campinas, Limeira.

TAKAHASHI, A.; BEDREGAL, B.R.C; LYRA, A. **Uma Versão Intervalar do Método de Segmentação de Imagens Utilizando o K-means**. TEMA - Tendências em Matemática Aplicada e Computacional, [S.l.], v. 6, n. 2, p. 315-324, june 2005. ISSN 2179-8451.

TRONCO, T. R. **Algoritmo de agrupamento Fuzzy C-Means para aprendizado e tomada de decisão em redes ópticas de próxima geração**. 2015. Tese

(Doutorado em Telecomunicações) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2015.

VALE, M. N. **Agrupamentos de Dados: Avaliação de Métodos e Desenvolvimento de Aplicativo para Análise de Grupos**. 2005. Dissertação (Mestrado em Engenharia Elétrica) — PUC-Rio.

VARGAS, R. R. de. **Uma nova forma de calcular os centros dos Clusters em algoritmos de agrupamento tipo fuzzy c-means**. 2012. 98 f. Tese (Doutorado em Ciência da Computação) - Universidade Federal do Rio Grande do Norte, Natal, 2012.

WANG, X. **Acoustic measurements on trees and logs: a review and analysis**. Wood Science and Technology, Springer Berlin Heidelberg, v. 47, n. 5, p. 965–975, 2013.