



UNIVERSIDADE FEDERAL DO CEARÁ

CAMPUS SOBRAL

**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA E DE
COMPUTAÇÃO**

MESTRADO ACADÊMICO EM ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO

JONIEL BASTOS BARRETO

**USO DA DIVERGÊNCIA GENERALIZADA NA CLUSTERIZAÇÃO APLICADA À
SEGMENTAÇÃO DE IMAGENS**

SOBRAL, CEARÁ

2021

JONIEL BASTOS BARRETO

USO DA DIVERGÊNCIA GENERALIZADA NA CLUSTERIZAÇÃO APLICADA À
SEGMENTAÇÃO DE IMAGENS

Dissertação apresentada ao Curso de Mestrado Acadêmico em Engenharia Elétrica e de Computação do Programa de Pós-Graduação em Engenharia Elétrica e de Computação do *Campus* Sobral da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Engenharia Elétrica e de Computação. Área de Concentração: Sistema de Informação

Orientador: Prof. Dr. Rui Facundo Vigellis

SOBRAL, CEARÁ

2021

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

B263u Barreto, Joniel Bastos.

Uso da Divergência Generalizada na Clusterização Aplicada à Segmentação de Imagens / Joniel Bastos Barreto. – 2021.

71 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Campus de Sobral, Programa de Pós-Graduação em Engenharia Elétrica e de Computação, Sobral, 2021.

Orientação: Prof. Dr. Rui Facundo Vigelis.

1. Divergência Generalizada. 2. Clusterização. 3. Segmentação de Imagens. I. Título.

CDD 621.3

JONIEL BASTOS BARRETO

USO DA DIVERGÊNCIA GENERALIZADA NA CLUSTERIZAÇÃO APLICADA À
SEGMENTAÇÃO DE IMAGENS

Dissertação apresentada ao Curso de Mestrado Acadêmico em Engenharia Elétrica e de Computação do Programa de Pós-Graduação em Engenharia Elétrica e de Computação do *Campus* Sobral da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Engenharia Elétrica e de Computação. Área de Concentração: Sistema de Informação

Aprovada em:

BANCA EXAMINADORA

Prof. Dr. Rui Facundo Vigelis (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Iális Cavalcante de Paula Júnior
Universidade Federal do Ceará (UFC)

Prof. Dr. Francisco Alixandre Ávila Rodrigues
Universidade Federal do Cariri (UFCA)

Dedico este trabalho à minha família, por sempre me apoiarem e investirem em mim, em especial à minha mãe Iracilda, por toda sua dedicação, cuidado e amor que foram essenciais para minha jornada.

AGRADECIMENTOS

Diante da conclusão de mais uma etapa da minha vida acadêmica, minha gratidão principalmente a Deus por dar-me forças e temperança.

Ao meu orientador, Prof. Dr. Rui Facundo Vigelis, pela paciência e incentivo na orientação que foram necessários para a realização deste trabalho e para meu crescimento com a pesquisa científica.

Aos professores participantes da banca examinadora, Prof. Dr. Iális Cavalcante de Paula Júnior, membro interno, e o Prof. Dr. Francisco Alixandre Ávila Rodrigues, membro externo, pela disposição de tempo e contribuições.

Aos meus colegas de laboratório, em principal pela participação na avaliação dos resultados necessária para o desenvolvimento desse trabalho, pela disponibilidade para ajudar quando preciso e pelo tempo que passamos juntos.

Ao meu colega e amigo, Danilo Oliveira, por ter me ajudado disponibilizando um sistema que permitisse a avaliação das segmentações online, assim otimizando a coleta dos resultados.

A todos professores, colegas e amigos que direta ou indiretamente contribuíram para a realização deste trabalho.

“Para realizar grandes conquistas, devemos não apenas agir, mas também sonhar; não apenas planejar, mas também acreditar.”

(Anatole France)

RESUMO

A segmentação de imagens é um processo fundamental em aplicações envolvendo análise de imagens, pois é responsável por definir as regiões de interesse para o problema. Uma técnica bastante utilizada para segmentação de imagens é a clusterização. A clusterização baseia-se em uma medida de similaridade entre objetos para agrupá-los em regiões. A utilização de medidas de dissimilaridade no espaço métrico é bem frequente em sistemas de reconhecimento de padrões, principalmente em algoritmos de clusterização. Contudo, alguns estudos tem observado limitações com funções de distância convencionais na etapa de busca por similaridade. Em vista disso, pesquisas voltadas para o uso de divergências não métricas na busca por semelhança entre objetos em algoritmos de clusterização tem atraído a atenção de pesquisadores. Este trabalho apresenta um caminho alternativo para segmentação de imagens utilizando uma divergência generalizada como medida de dissimilaridade no algoritmo de clusterização *hard* de Bregman. Essa divergência é obtida pela generalização das entropias relativas de Shannon e Tsallis, e definida em relação a uma exponencial deformada φ . O trabalho é desenvolvido utilizando três funções exponenciais deformadas, denominadas φ_1 , φ_2 e φ_3 , e comparando seus resultados com a técnica de clusterização *K-means* clássica e para algumas imagens com o método de Otsu. Os testes foram realizados para 5 grupos de imagens, em que 2 grupos são de imagens em escala de cinza divididos em segmentação binária e segmentação em mais de duas regiões, e 3 grupos de imagens coloridas, para testes no espaço de cor RGB, HSV e com realce dos níveis RGB. Para analisar os resultados das segmentações são utilizados, além das métricas acurácia, precisão e sensibilidade, o coeficiente de similaridade Sørensen–Dice (CSD), e uma avaliação qualitativa obtida através da média de notas atribuídas empiricamente foram utilizadas. Os resultados apresentaram uma boa performance do método proposto com acurácia acima de 0,9 em todas as imagens utilizando a exponencial deformada φ_1 , CSD > 0,7 para 23 das 25 imagens testadas utilizando as exponenciais deformadas φ_1 e φ_2 e nota média acima de 7 para 24 das 25 imagens para as exponenciais deformadas φ_1 e φ_2 . Diante dos resultados obtidos é possível concluir que o método proposto não só se mostrou aplicável como também, dependendo da aplicação, pode apresentar um melhor desempenho que métodos já estabelecidos na literatura.

Palavras-chave: Divergência generalizada. Clusterização. Segmentação de imagens.

ABSTRACT

Image segmentation is a fundamental process in applications involving image analysis, as it is responsible for defining the regions of interest for the problem. A widely used technique for image segmentation is clustering. Clustering is based on a measure of similarity between objects to group them into regions. The use of dissimilarity measures in the metric space is quite common in pattern recognition systems, mainly in clustering algorithms. However, some studies have observed limitations with conventional distance functions in the search for similarity. In view of this, research focused on the use of non-metric divergences in the search for similarity between objects in clustering algorithms has attracted the attention of researchers. This work presents an alternative path for image segmentation using a generalized divergence as a measure of dissimilarity in Bregman hard clustering algorithm. This divergence is obtained by generalizing the relative entropies of Shannon and Tsallis, and defined in relation to a deformed exponential φ . The work is developed using three deformed exponential functions, called φ_1 , φ_2 and φ_3 , and comparing their results with the classic K- textit means clustering technique and for some images with the Otsu's method. The tests were performed for 5 groups of images, in which 2 groups are grayscale images divided into binary segmentation and segmentation in more than two regions, and 3 groups of color images, for tests in the RGB, HSV color space and with enhancement of the RGB levels. In order to analyze the segmentation results, in addition to the metrics accuracy, precision and sensitivity, the Sørensen – Dice similarity coefficient (CSD), and a qualitative evaluation obtained through the average of grades empirically attributed were used. The results showed a good performance of the proposed method with accuracy above 0.9 in all images using the φ_1 deformed exponential, $CSD > 0.7$ for 23 of the 25 images tested using the φ_1 and φ_2 deformed exponentials and average grade above 7 for 24 of the 25 images for the φ_1 and φ_2 deformed exponentials. In view of the results obtained, it is possible to conclude that the proposed method was not only shown to be applicable but also, depending on the application, it may present a better performance than methods already established in the literature.

Keywords: Generalized Divergence. Clustering. Image segmentation.

LISTA DE FIGURAS

Figura 1 – Limiarização pelo método de Otsu	25
Figura 2 – Clusterização de dados	26
Figura 3 – Exemplo de uma segmentação ideal do texto de uma imagem com ruídos.	29
Figura 4 – Avaliação do Coeficiente de Similaridade de Sørensen–Dice (CSD) na segmentação do texto de uma imagem com ruídos.	32
Figura 5 – Interpretação geométrica da divergência de Bregman no \mathbb{R}^2	37
Figura 6 – Exemplos de ruídos simulados do banco de imagens NoiseOffice.	43
Figura 7 – Exemplo de segmentação ideal para 3 regiões em uma imagem	43
Figura 8 – Exemplo de segmentação de uma região específica	44
Figura 9 – Resultado da segmentação da Imagem 2 do grupo 1 do Experimento I	53
Figura 10 – Resultado da segmentação da Imagem 1 do grupo 1 do Experimento I	54
Figura 11 – Histograma da imagem e sua classe positiva	55
Figura 12 – Histograma da imagem e sua classe positiva	57
Figura 13 – Histograma dos níveis RGB das imagens	60
Figura 14 – Histograma dos níveis HSV das imagens	62
Figura 15 – Histograma dos níveis RGB das imagens depois do pré-processamento	64
Figura 16 – Imagens do grupo 1 do Experimento I	74
Figura 17 – Imagens do grupo 2 do Experimento I	75
Figura 18 – Imagens do grupo 1 do Experimento II	76
Figura 19 – Imagens do grupo 2 do Experimento II	77
Figura 20 – Imagens do grupo 3 do Experimento II	78

LISTA DE TABELAS

Tabela 1 – Matriz de confusão	30
Tabela 2 – Divergências de Bregman geradas a partir de algumas funções convexas . .	37
Tabela 3 – Tabela dos resultados da acurácia para cada método no grupo 1 do Experimento I	50
Tabela 4 – Tabela dos resultados da precisão para cada método no grupo 1 do Experimento I	50
Tabela 5 – Tabela dos resultados da sensibilidade para cada método no grupo 1 do Experimento I	51
Tabela 6 – Tabela dos resultados do CSD para cada método no grupo 1 do Experimento I	51
Tabela 7 – Tabela da média das notas para cada método no grupo 1 do Experimento I .	52
Tabela 8 – Tabela dos resultados da acurácia para cada método no grupo 2 do Experimento I	56
Tabela 9 – Tabela dos resultados da precisão para cada método no grupo 2 do Experimento I	56
Tabela 10 – Tabela dos resultados da sensibilidade para cada método no grupo 2 do Experimento I	56
Tabela 11 – Tabela dos resultados do CSD para cada método no grupo 2 do Experimento I	56
Tabela 12 – Tabela da média das notas para cada método no grupo 2 do Experimento I .	57
Tabela 13 – Tabela dos resultados da acurácia para cada método no grupo 1 do Experimento II	58
Tabela 14 – Tabela dos resultados da precisão para cada método no grupo 1 do Experimento II	58
Tabela 15 – Tabela dos resultados da sensibilidade para cada método no grupo 1 do Experimento II	58
Tabela 16 – Tabela dos resultados do CSD para cada método no grupo 1 do Experimento II	58
Tabela 17 – Tabela da média das notas para cada método no grupo 1 do Experimento II .	59
Tabela 18 – Tabela dos resultados da acurácia para cada método no grupo 2 do Experimento II	60
Tabela 19 – Tabela dos resultados da precisão para cada método no grupo 2 do Experimento II	60

Tabela 20 – Tabela dos resultados da sensibilidade para cada método no grupo 2 do Experimento II	60
Tabela 21 – Tabela dos resultados do CSD para cada método no grupo 2 do Experimento II .	61
Tabela 22 – Tabela da média das notas para cada método no grupo 2 do Experimento II .	61
Tabela 23 – Tabela dos resultados da acurácia para cada método no grupo 3 do Experimento II	62
Tabela 24 – Tabela dos resultados da precisão para cada método no grupo 3 do Experimento II	62
Tabela 25 – Tabela dos resultados da sensibilidade para cada método no grupo 3 do Experimento II	63
Tabela 26 – Tabela dos resultados do CSD para cada método no grupo 3 do Experimento II	63
Tabela 27 – Tabela da média das notas para cada método no grupo 3 do Experimento II .	63

LISTA DE ABREVIATURAS E SIGLAS

CBIR	<i>Content-based Image Retrieval</i>
CSD	Coeficiente de Similaridade de Sørensen–Dice
DBs	divergências de Bregman
DTB	divergência total de Bregman
FN	Falso Negativo
FP	Falso Positivo
GT	<i>Ground Truth</i>
KL	Kullback-Leibler
LBG	Linde-Buzo-Gray
NN	Nearest Neighbor
PDI	Processamento Digital de Imagens
VN	Verdadeiro Negativo
VP	Verdadeiro Positivo

LISTA DE SÍMBOLOS

Ac	Acurácia
Pr	Precisão
Se	Sensibilidade
\mathbb{R}	Conjunto de números reais
\mathbb{R}^d	Espaço vetorial real d-dimensional
\mathbb{R}_+	Conjunto de números reais não negativos
\mathbb{R}_{++}	Conjunto de números reais estritamente positivos
$\mathbb{R}_{>}$	Conjunto de números reais
ri	Interior relativo
RGB	<i>Red, Green, Blue</i>
HSV	<i>Hue, Saturation, Value</i>

SUMÁRIO

1	INTRODUÇÃO	16
1.1	Motivação	18
1.2	Estado da Arte	19
1.3	Objetivos	21
<i>1.3.1</i>	<i>Objetivos Específicos</i>	<i>21</i>
1.4	Produção Científica	21
1.5	Estrutura da Dissertação	21
2	FUNDAMENTAÇÃO TEÓRICA	23
2.1	Segmentação	23
<i>2.1.1</i>	<i>Limiarização</i>	<i>24</i>
<i>2.1.1.1</i>	<i>Limiarização de Otsu</i>	<i>25</i>
<i>2.1.2</i>	<i>Clusterização</i>	<i>26</i>
<i>2.1.2.1</i>	<i>K-means</i>	<i>27</i>
<i>2.1.3</i>	<i>Avaliação da Segmentação</i>	<i>28</i>
<i>2.1.3.1</i>	<i>Matriz de Confusão</i>	<i>30</i>
<i>2.1.3.2</i>	<i>Coefficiente de Similaridade de Dice</i>	<i>31</i>
<i>2.1.3.3</i>	<i>Análise Qualitativa</i>	<i>32</i>
2.2	Entropia	33
<i>2.2.1</i>	<i>Entropia Relativa</i>	<i>34</i>
<i>2.2.2</i>	<i>Entropia Relativa Generalizada</i>	<i>35</i>
2.3	Divergências de Bregman	36
<i>2.3.1</i>	<i>Clusterização Hard de Bregman</i>	<i>39</i>
3	METODOLOGIA	42
3.1	Banco de imagens	42
3.2	Funções exponenciais deformadas	44
3.3	Clusterização baseada na divergência generalizada	45
4	EXPERIMENTOS E ANÁLISE DOS RESULTADOS	49
4.1	Experimento I	50
<i>4.1.1</i>	<i>Grupo 1</i>	<i>50</i>
<i>4.1.2</i>	<i>Grupo 2</i>	<i>55</i>

4.2	Experimento II	57
4.2.1	Grupo 1	57
4.2.2	Grupo 2	59
4.2.3	Grupo 3	61
4.3	Conclusão dos Resultados dos Experimentos	63
5	CONCLUSÕES E TRABALHOS FUTUROS	65
5.1	Considerações Finais	65
5.2	Trabalhos Futuros	66
	REFERÊNCIAS	67
	APÊNDICES	74
	ANEXOS	74
	ANEXO A – Imagens utilizadas para os testes	74

1 INTRODUÇÃO

Sistemas de análise de imagens geralmente envolvem três componentes principais: (1) a segmentação que separa as regiões de interesse da imagem, (2) a extração de atributos que diferenciam um objeto de outro, e (3) a análise dos atributos que visa identificar os objetos segmentados. Para que esses sistemas apresentem bons resultados e sejam eficientes em identificar diferentes objetos é preciso que os atributos extraídos apresentem características importantes para a descrição dos objetos e para isso é necessário que as regiões da imagem tenham sido bem segmentadas.

A segmentação é responsável por definir da imagem original quais objetos são importantes para o problema em questão, o que a torna uma etapa crucial para o desenvolvimento de um algoritmo de reconhecimento de imagens, pois um erro na segmentação resultaria em descritores ruins e conseqüentemente em uma má classificação. Para Garcia-Lamont *et al.* (2018), uma boa segmentação de imagens consiste em agrupar *pixels* de forma uniforme e homogênea em relação a algumas características, como tom, textura, ou nível de cinza. Contudo, imagens reais nem sempre possuem regiões bem definidas, o que dificulta o processo de segmentação, e avaliar a qualidade da segmentação pode ser uma tarefa desafiadora.

Estudos voltados para análise da performance de técnicas de segmentação já estão sendo realizados ao longo dos anos. Wang *et al.* (2020) apresentaram um estudo sobre técnicas de avaliação de segmentação de imagens utilizando métodos supervisionados e não-supervisionados. Chabrier *et al.* (2004) compararam cinco critérios de avaliação supervisionada testando em nove algoritmos de segmentação de imagens diferentes. Zhang *et al.* (2008) revisaram métodos de avaliação objetiva não-supervisionada propostos na literatura apresentando suas vantagens e desvantagens. Chen *et al.* (2018) apresentaram um resumo de métodos de avaliação amplamente usados no campo de sensoriamento remoto e analisaram suas vantagens e desvantagens. Os autores também afirmam que estudar o impacto do método de avaliação no resultado da segmentação é tão importante quanto estudar o algoritmo de segmentação em si.

Nas últimas décadas, técnicas de segmentação têm sido utilizadas na solução de diversos problemas bem como em algumas aplicações tais como reconhecimento facial (KAMENCAY *et al.*, 2013), reconhecimento de caracteres (DUNN; WANG, 1992), imagens médicas (HAMEED *et al.*, 2013), entre outros (KHAN, 2014). A literatura fornece uma variedade de algoritmos de segmentação de imagens, contudo cada imagem possui suas peculiaridades e não há uma técnica que possa funcionar eficientemente em todos os tipos de imagens. Das técnicas

baseadas em similaridade entre os *pixels*, a limiarização, crescimento de região e clusterização são as mais comuns em pesquisas (SONG; YAN, 2017).

A técnica clássica de limiarização consiste em separar os níveis de cinza de uma imagem a partir de um limiar específico, onde uma região é definida por *pixels* com intensidade igual ou abaixo do limiar e a outra pelos *pixels* com intensidade acima. A partir de um conjunto de pontos pré-definidos, chamados de sementes, o crescimento de região consiste em combinar os *pixels* vizinhos da semente com propriedades semelhantes para formar a região. A segmentação por clusterização baseia-se na similaridade de objetos para dividir a imagem em regiões de forma que os objetos de uma região sejam mais similares entre si do que em relação aos objetos de grupos distintos. Diferentemente da limiarização ou do crescimento de região, na clusterização não se faz necessário especificar um limiar ou definir sementes.

Para que o algoritmo de clusterização realize sua tarefa é necessário o uso de uma medida de similaridade, ou dissimilaridade, para quantificar a semelhança entre os objetos. A distância Euclidiana é uma métrica comumente utilizada como medida de dissimilaridade em algoritmos de clusterização, como na técnica *K-means*. Nesta técnica, primeiramente é definido o número de grupos em que os dados serão divididos, em seguida são localizados os centroides desses grupos, e, por fim, cada ponto é atribuído ao grupo que apresentar menor distância, ou seja, maior similaridade. Em várias pesquisas utiliza-se o *K-means* em aplicações com segmentação de imagens, como Wanjari *et al.* (2015) que usam a técnica de clusterização *K-means* para segmentar imagens de documentos, e Katkar e Baraskar (2015) que utilizam PCA em conjunto com o *K-means* para segmentar imagens médicas.

A medida de dissimilaridade pode ser definida por diferentes funções de divergências, que podem não ser, necessariamente, uma métrica. Bregman (1967) definiu divergências obtidas a partir de funções convexas que apresentam propriedades importantes para as aplicações na área de processamento de imagens. Segundo Banerjee *et al.* (2005) o algoritmo *K-means* é um caso especial do algoritmo de clusterização *hard* de Bregman para a distância Euclidiana quadrática. Outros exemplos são o algoritmo Linde-Buzo-Gray (LBG) (BUZO *et al.*, 1980) e o algoritmo de *information theoretic clustering* (DHILLON *et al.*, 2003) que utilizam a distância Itakura-Saito e a divergência Kullback-Leibler (KL), respectivamente.

Este trabalho tem como foco apresentar um método de segmentação de imagens por meio do algoritmo de clusterização *hard* de Bregman, proposto por Banerjee *et al.* (2005), utilizando uma divergência não métrica como medida de dissimilaridade.

1.1 Motivação

Medidas de dissimilaridades são utilizadas para o desenvolvimento de diversas estruturas e algoritmos como em sistemas envolvendo reconhecimento de padrões, agrupamento ou análise de imagens. Na maior parte dos trabalhos que dependem da consulta por similaridade são utilizados espaços métricos (ZHANG *et al.*, 2009). Nos espaços métricos, as funções de proximidades são estritamente simétricas, não negativas e satisfazem a desigualdade triangular. Contudo, existem alguns problemas reais que abordam matrizes, comparação de distribuições de probabilidades, imagens, entre outros, em que a aplicação de uma distância métrica pode não ser adequada.

Embora muitos métodos desenvolvidos para análise de imagens e visão computacional tenham gerados bons resultados, alguns estudos tem observados que existem aplicações em que as funções de dissimilaridade convencionais, tais como a euclidiana e a cosseno, têm apresentado limitações na busca por semelhança entre objetos (SAKJI-NSIBI; BENAZZA-BENYAHIA, 2010), (LIU, 2011), (XU *et al.*, 2012). A partir disso, as divergências de Bregman (DBs) têm surgido como alternativa para medir a dissimilaridade entre dados, pois apresentam propriedades importantes na busca por similaridades que são válidas para uma família de funções. Diferente das métricas convencionais, as DBs podem ser não simétricas e nem sempre satisfazem a desigualdade triangular (ZEZULA *et al.*, 2006).

Na literatura científica já é possível encontrar diversos trabalhos que generalizam estruturas e algoritmos utilizando DBs em aplicações envolvendo análise de dados e clusterização. Banerjee *et al.* (2005) propuseram e analisaram algoritmos paramétricos de agrupamento *hard* e *soft* baseados nas DBs. Cayton (2008) apresenta uma estrutura de dados utilizando DBs que permite encontrar de forma mais eficiente o objeto vizinho mais próximo no algoritmo Nearest Neighbor (NN). Essa generalização é importante pois possibilita a escolha da função de dissimilaridade, o que torna a aplicação desses métodos de busca por similaridade mais flexível.

Ferreira (2017) aponta que uma das principais vantagens na generalização propiciada pelas DBs é a possibilidade de desenvolvimento de métodos mais eficientes e aplicáveis. Isso porque é possível considerar um conjunto de funções de dissimilaridades ao invés de uma única medida e assim poder escolher uma medida adequada ao problema em questão. Devido a flexibilidade das DBs em relação a outras medidas como euclidiana e cosseno, as divergências não métricas têm se tornado um importante material de estudo para o cálculo da similaridade entre diferentes objetos. Há trabalhos em que a medida mais adequada não é clara e a escolha dessa

medida é apresentada como material de aprendizagem (XING *et al.*, 2002), (GOLDBERGER *et al.*, 2004).

Considerando os estudos de divergências não métricas e a importância em aplicações envolvendo medidas de dissimilaridade, esse trabalho propõe o emprego da entropia relativa generalizada, definida em relação a uma exponencial deformada φ , para o cálculo da dissimilaridade na clusterização *hard* de Bregman aplicada a segmentação de imagens.

1.2 Estado da Arte

As DBs são medidas de dissimilaridade utilizadas em diversos problemas de otimização e análise numérica, mas que também já apresenta contribuições em aplicações envolvendo aprendizagem estatística (GUO *et al.*, 2017), problemas inversos (BURGER, 2016) e clusterização de dados (LIU; BELKIN, 2016). Alguns trabalhos apresentam resultados ao substituir a distância euclidiana pelas DBs em aplicações como problemas de otimização (KIWIEL, 1997), problemas de inequações variacionais (SEMENOV, 2018). Recentemente, as DBs também estão sendo utilizadas em pesquisas da área de processamento de imagens como segmentação (WU; ZHANG, 2020) e classificação (TAKENOUCI; ISHII, 2018).

Em sistemas de clusterização, uma das divergências mais conhecidas é a distância euclidiana quadrática, utilizada como medida de similaridade no algoritmo *K-means* (MACQUEEN, 1967). Telgarsky e Dasgupta (2012) desenvolveram uma teoria de clusterização aglomerativo com DBs a partir do conceito apresentado por Banerjee *et al.* (2005) de que o *K-means* é uma consequência natural do algoritmo de clusterização *hard*. Outros exemplos de DBs conhecidas são a distância Itakura-Saito e a divergência KL, utilizadas nos algoritmos LBG (LINDE *et al.*, 1980), (BUZO *et al.*, 1980) e *information theoretic clustering* (DHILLON *et al.*, 2003), respectivamente.

Algumas pesquisas voltadas para a recuperação de imagens também fazem uso das DBs, como Xu *et al.* (2012) que desenvolveram um algoritmo de ranqueamento de variedades como um problema de otimização utilizando DBs, aplicam a problemas de recuperação de imagens baseada em conteúdo, *Content-based Image Retrieval* (CBIR). Sakji-Nsibi e Benazza-Benyahia (2010) descreveram o uso da divergência KL simétrica como medida de similaridade em sistemas CBIR, enquanto Rocha (2015) utilizou a divergência KL e a divergência generalizada para mesma aplicação. Rocha (2015) também realizou um estudo comparativo sobre a utilização das DBs em relação às funções euclidiana e cosseno como medida de similaridade e averígua

suas vantagens e desvantagens. Ennekhli *et al.* (2018) apresentaram um novo sistema de CBIR usando a divergência KL em conjunto do *Parzen Relevance Feedback* para reduzir a lacuna semântica entre imagens.

Ferreira (2017) propôs novos métodos de análise de imagens por meio de DBs e divergência total de Bregman (DTB). Já Liu *et al.* (2012) fizeram uso da DTB como uma medida eficiente e robusta de distância para quantificar a diferença entre formas, com aplicação a recuperação de formas de objetos. Também são encontrados trabalhos com o uso da DTB em análise de imagens de Tensor de Difusão (LIU, 2011), (VEMURI *et al.*, 2011) e em técnicas de rastreamento de objetos (ROMERO *et al.*, 2013). Wu e Zhang (2020) desenvolveram uma nova clusterização utilizando DTB para melhorar a precisão e robustez do algoritmo *fuzzy local information C-means* para imagens corrompidas por alto ruído.

Gupta *et al.* (2015) apresentaram uma abordagem hierárquica de clusterização aglomerativa baseada na divergência KL simétrica, com aplicação a segmentação de múltiplos níveis de despigmentação em imagens de vitiligo. Também compararam com outros métodos já estabelecidos na literatura, provando que a divergência KL simétrica é um limite superior para distribuições gaussianas unimodais. Zhao *et al.* (2016) aplicaram a divergência KL bidimensional em conjunto a otimização de enxame de partículas modificada em um algoritmo de limiarização multinível. Nesse algoritmo é usado a formulação da divergência KL bidimensional como uma função objetivo da segmentação das imagens.

Wei *et al.* (2018) propuseram um novo método de clusterização para segmentação de imagens, *Fuzzy Cluster Ensembler* baseado na divergência KL. Este método diferencia-se do *fuzzy C-means* apenas na substituição da distância euclidiana pela divergência KL para realizar as associações. Zhao *et al.* (2015) também apresentaram um novo algoritmo de clusterização *fuzzy* utilizando a distância de Mahalanobis como medida de dissimilaridade aplicado a segmentação de imagens. Já Gauri e Sharma (2018) propuseram como melhorar o desempenho da técnica *K-means* usando DBs para medir a similaridade entre *pixels*. Com base em seus resultados, concluíram que com as DBs as bordas são identificadas de maneira mais refinada do que utilizando a distância euclidiana no *K-means*, mostrando ser capaz de identificar facilmente vários objetos com base na semelhança de cores na imagem.

1.3 Objetivos

O presente trabalho tem como objetivo geral aplicar a divergência generalizada, obtida a partir da generalização das entropias relativas de Shannon e Tsallis, como medida de dissimilaridade no algoritmo de clusterização *hard* de Bregman e apresentar sua eficiência na segmentação de imagens.

1.3.1 *Objetivos Específicos*

- Apresentar um método eficaz para segmentação de imagens;
- Aplicar a entropia relativa generalizada como a divergência do algoritmo clusterização *hard* de Bregman no agrupamento de dados;
- Avaliar o desempenho da divergência generalizada como medida de dissimilaridade em segmentação de imagens por clusterização;
- Comparar o desempenho do método proposto com técnicas de segmentação conhecidas.

1.4 Produção Científica

A partir desta pesquisa foi desenvolvido um trabalho utilizando a divergência generalizada no algoritmo de clusterização *hard* de Bregman. A metodologia foi aplicada para segmentação de imagens e testada com banco de imagens NoisyOffice (MARTINEZ *et al.*, 2007). O trabalho foi publicado no **XXXVII Simpósio Brasileiro de Telecomunicações e Processamento de Sinais** (BARRETO; VIGELIS, 2019).

Um trabalho também foi apresentado no **XXXVIII Simpósio Brasileiro de Telecomunicações e Processamento de Sinais** onde é aplicada a clusterização baseada na divergência generalizada como técnica de segmentação para um sistema de identificação de lesões em imagens de radiografias periapicais (BARRETO *et al.*, 2020).

1.5 Estrutura da Dissertação

Para a exposição do estudo e visando uma melhor compreensão das técnicas utilizadas, metodologia, experimentos e resultados, este trabalho está organizado em 5 capítulos. A seguir é descrito o conteúdo apresentado nos capítulos seguintes.

- **Capítulo 2:** Este capítulo aborda conceitos básicos e algumas técnicas conhecidas de

segmentação de imagens, assim como métodos de avaliação de segmentação. Além disso, o capítulo abrange o desenvolvimento de estudos da entropia, entropia relativa e entropia relativa generalizada no contexto de teoria da informação. Também são apresentadas definições e propriedades das divergências de Bregman e suas aplicações em algoritmos de clusterização.

- **Capítulo 3:** Neste capítulo são apresentadas as etapas de execução da metodologia proposta de clusterização *hard* para segmentação de imagens e as funções exponenciais utilizadas na divergência generalizada, assim como o banco de imagem utilizado para os testes.
- **Capítulo 4:** São apresentados neste capítulo os experimentos realizados e os resultados obtidos pelo método proposto. Também é analisado e discutido a comparação do método proposto em relação a outras técnicas consolidadas na literatura.
- **Capítulo 5:** Por fim, as considerações finais, principais contribuições e propostas para pesquisas futuras são relatadas neste capítulo.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo primeiramente é apresentado com detalhes o conceito e algumas técnicas de segmentação de imagens, assim como métodos de avaliar seu desempenho. Em seguida, é abordado o princípio da entropia na teoria da informação e estudos no desenvolvimento da entropia relativa generalizada. Por fim, é relatado sobre divergências e o seu uso em algoritmos de clusterização.

2.1 Segmentação

A segmentação é uma das etapas fundamentais em sistemas de visão computacional, que consiste em agrupar objetos semelhantes em uma ou mais características, dividindo uma imagem em diferentes regiões e extraíndo as regiões de interesse. O processo de definir padrões e diferenciá-los, assim como tomar decisões com base na informação visual obtida, ocorre naturalmente no sistema visual-cognitivo humano. Tornar o computador capaz de realizar essa separação automaticamente é a principal ideia da segmentação de imagens (DE *et al.*, 2016).

A imagem segmentada deve apresentar partições homogêneas onde os *pixels* sejam similares com os do mesmo grupo e diferente dos *pixels* de outros grupos. Fu e Mui (1981) descrevem este processo da seguinte forma: considerando uma imagem R , a segmentação é a decomposição de R em n regiões R_1, R_2, \dots, R_n , de forma que

$$\bigcup_{i=1}^n R_i = R, \quad (2.1)$$

$$R_i \text{ é um conjunto conectado, } i = 1, 2, \dots, n \quad (2.2)$$

$$R_i \cap R_j = \emptyset, \forall i \neq j, \quad (2.3)$$

$$P(R_i) = \text{VERDADEIRO para } i = 1, 2, \dots, n, \quad (2.4)$$

$$P(R_i \cup R_j) = \text{FALSO para regiões adjacentes } R_i \text{ e } R_j, \quad (2.5)$$

onde $P(R_k)$ representa as propriedades de homogeneidade definidas dos *pixels* sobre a região R_k . Essa operação pode ser realizada de modo manual, baseado na interpretação visual do operador, ou automático, com um algoritmo responsável para diferenciar as regiões.

Segundo Gonzalez e Woods (2011) a técnica de segmentação de imagens depende da aplicação do problema e das propriedades da imagem, podendo se basear em critérios de similaridade ou descontinuidade. Portanto não segue um modelo padrão. Características como níveis de cinza, cor, brilho, contraste, textura, dentre outros, podem ser usados para definir as

semelhanças entre os objetos e diferenças entre as regiões, o que torna essa técnica um processo empírico e adaptativo. Gonzalez e Woods (2011) dividem a segmentação em duas principais categorias: segmentação por descontinuidade e segmentação por similaridade.

As segmentações por descontinuidade consistem em diferenciar um objeto do plano de fundo considerando a mudança abrupta da intensidade dos *pixels* para identificar os limites desse objeto. Para determinar a variação entre um *pixel* e seus vizinhos são utilizadas filtros derivativos como operador Gradiente (primeira derivada) ou operador Laplaciano (segunda derivada). Nessa abordagem a segmentação é realizada por meio de uma convolução que pode ser expressa, para uma imagem bidimensional, pela seguinte equação:

$$g(x,y) = M_{(N \times N)} * f(x,y), \quad (2.6)$$

sendo $M_{(N \times N)}$ uma matriz com os coeficientes relativos à derivada, $f(x,y)$ a imagem de entrada e $g(x,y)$ a imagem segmentada. De acordo com Gonzalez e Woods (2011) essas técnicas são aplicadas a imagens em níveis de cinza, para a detecção de linhas, pontos e bordas.

Abordagens baseadas na segmentação por similaridade fundamentam-se em agrupar *pixels* que satisfaçam certos critérios de semelhança predefinidos. Esses métodos comparam propriedade como intensidade, cor e/ou textura para definir a similaridade entre diferentes *pixels*. A limiarização, o crescimento de regiões e a clusterização são métodos de segmentação por similaridade comumente encontrados na literatura.

2.1.1 Limiarização

Segundo Gonzalez e Woods (2011) uma das principais abordagens em aplicações de segmentação de imagens é a limiarização devido às suas propriedades intuitivas, simplicidade de implementação e velocidade computacional. Também chamada de binarização, esta técnica parte do princípio de que, sendo uma imagem composta por objetos claros em um fundo escuro, com valores de intensidade agrupados em dois grupos dominantes, a maneira óbvia de separar os objetos do fundo é selecionar um limiar para separar esses modos. Ou seja, para uma imagem $f(x,y)$ é possível rotular cada *pixel* da imagem como objeto, caso o valor do *pixel* seja acima do limiar T , ou como fundo, valor igual ou abaixo do limiar. A imagem limiarizada $g(x,y)$ pode ser obtida por

$$g(x,y) = \begin{cases} 1, & \text{se } f(x,y) > T \\ 0, & \text{se } f(x,y) \leq T. \end{cases} \quad (2.7)$$

Esta operação também pode ser realizada para isolar *pixels* do restante da imagem. Desta forma, os *pixels* com valores acima do limiar são definidos como 1 e os *pixels* restantes são mantidos inalterados.

2.1.1.1 Limiarização de Otsu

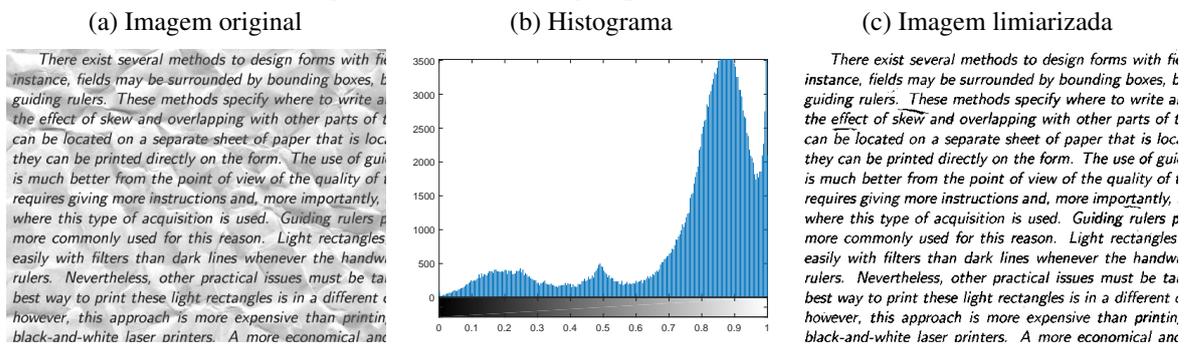
Definir o valor de um limiar T é mais simples em imagens que possuem um vale bem definido entre os dois picos do histograma. Contudo, na maioria das imagens reais, é difícil detectar com precisão o fundo do vale, principalmente em casos com vale plano e amplo ou quando os picos possuem alturas bem desiguais. Visando maximizar a variância entre classes, Otsu (1979) apresentou um método que minimiza o erro médio ao atribuir *pixels* a um ou mais grupos. O método de Otsu é uma das técnicas baseadas na limiarização mais conhecidas na literatura.

A limiarização de Otsu baseia-se inteiramente em cálculos realizados a partir do histograma de uma imagem buscando um valor para o limiar T que minimize a soma da variância intraclases da imagem. O cálculo dessa variância é dado por

$$\sigma_W^2 = W_b \sigma_b^2 + W_f \sigma_f^2, \quad (2.8)$$

sendo W_b e W_f o peso para cada classe, correspondente à probabilidade de um pixel pertencer à classe b (*background*) ou à f (*foreground*), respectivamente. O resultado desse método pode ser observado na Figura 1.

Figura 1 – Limiarização pelo método de Otsu

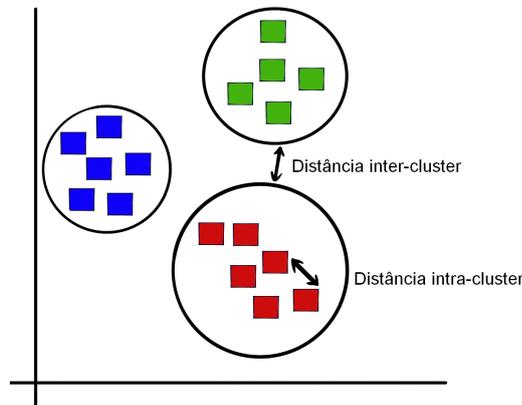


Fonte: Adaptado de Martinez *et al.* (2007)

2.1.2 Clusterização

A clusterização de dados, também chamada de análise de agrupamento, é um método fundamental de aprendizado não supervisionado que vem sendo estudado em diversos trabalhos ao longo de vários anos (JAIN; DUBES, 1988). Esta abordagem consiste em agrupar objetos em clusters de forma que os pontos de dados de um clusters estejam próximos entre si e distantes dos pontos de outros clusters, sendo os critérios que definem essa proximidade preestabelecidos. Ou seja, a clusterização é realizada com o objetivo de maximizar a homogeneidade interna (intra-cluster) e maximizar a heterogeneidade entre clusters distintos, como observado na Figura 2.

Figura 2 – Clusterização de dados



Hruschka e Ebecken (2003) descrevem a clusterização considerando um conjunto de n objetos $X = \{x_1, x_2, \dots, x_n\}$, em que $x_i \in \mathbb{R}_+$ objetos devem ser agrupados em k clusters disjuntos $C = \{C_1, C_2, \dots, C_k\}$ de forma que:

$$C_1 \cup C_2 \cup \dots \cup C_k = X; \quad (2.9)$$

$$C_i \neq \emptyset, \text{ para } 1 \leq i \leq k; \quad (2.10)$$

$$C_i \cap C_j = \emptyset, \forall i \neq j, \text{ para } 1 \leq i, j \leq k. \quad (2.11)$$

Com isso, é possível observar que a segmentação de imagens se assemelha com um processo de clusterização, uma vez que se considera as regiões das imagens como clusters e os *pixels* como pontos de dados. Diferentemente da limiarização que comumente é utilizada para segmentar duas regiões, na clusterização é possível definir o número de regiões a ser segmentada. Também apresenta certa vantagem em relação ao crescimento de regiões, pois não é preciso definir sementes.

Para o desenvolvimento de um algoritmo de clusterização são necessários parâmetros que variam dependendo da técnica utilizada. Contudo na maioria dessas técnicas é exigido um parâmetro que determine a quantidade de grupos, seja diretamente, como definir um número k de clusters, ou indiretamente, como limiar de erro (SALVADOR; CHAN, 2004). Assim sendo, é importante que o desenvolvedor tenha uma noção de quantas regiões o algoritmo deve segmentar a imagem.

Os métodos de clusterização são divididos em quatro principais categorias: particionais, hierárquicos, baseados em densidade dos *pixels* e baseados em grade (SALVADOR; CHAN, 2004). Os algoritmos por particionamento são os métodos mais comuns de clusterização. Consistem em dividir o conjunto de dados em k partições a partir de um critério de semelhança, sendo k um valor definido pelo usuário. Os hierárquicos estruturam o conjunto de dados hierarquicamente, de acordo com a proximidade entre os objetos, como um dendrograma. Nessa abordagem, a clusterização é realizada dividindo o dendrograma em k clusters. Os métodos baseados em densidade partem do princípio de que dentro de cada cluster há uma densidade de objetos típica que é consideravelmente maior que fora dos clusters. São adequados para descobrir regiões de forma arbitrária e lidar com *outliers*. Já os algoritmos baseados em grade dividem os dados em células formando uma estrutura de grade, onde são realizadas as operações de clusterização, permitindo o agrupamento eficiente de conjuntos de dados muito grandes. Dentre os métodos de clusterização, o algoritmo mais popular e simples é o *K-means*.

2.1.2.1 *K-means*

O algoritmo *K-means* (Algoritmo 1) é um método iterativo e heurístico não supervisionado de clusterização por particionamento, apresentado por MacQueen (1967), que particiona um conjunto de n objetos em k clusters visando ter uma similaridade intracluster alta e uma similaridade intercluster baixa. Mesmo tendo sido proposto a mais de 50 anos, o *K-means* ainda é um dos algoritmos de clusterização mais utilizados. Isto se deve a sua simplicidade, facilidade em implementação, eficiência e sucesso empírico, e por ser utilizado como base para outras técnicas (JAIN, 2010).

Para a implementação do algoritmo *K-means* primeiramente é necessário definir o número k de regiões em que a imagem deve ser segmentada. Dada a quantidade de clusters, são inicializados k valores para representar os centroides de cada cluster C_j , podendo estes valores serem aleatórios. Em seguida é realizada a etapa de atribuição, em que o *pixel* é atribuído ao

cluster que possui seu centro mais próximo do *pixel*. Nessa etapa é calculado a distância de cada *pixel* a cada centroide utilizando a distância euclidiana, que é dada por

$$d(x_i, \mu_j) = \|x_i - \mu_j\|, \quad (2.12)$$

sendo x_i o i -ésimo *pixel* da imagem, μ_j o centroide do cluster C_j e $\|\cdot\|$ denota a norma. Posteriormente são calculados novos centros para cada cluster, em que o centroide μ_j é obtido a partir da média dos *pixels* pertencentes ao cluster C_j . As etapas de atribuição e re-estimação são repetidas até que o algoritmo alcance algum critério de parada, que pode ser um limite de iterações pré-definido ou minimizar uma função objetivo. A função objetivo mais utilizada é o erro quadrático (GUHA *et al.*, 1999), dada por

$$E = \sum_{j=1}^k \sum_{i=1}^m \|x_i - \mu_j\|^2, \quad (2.13)$$

em que $m = |C_j|$ e $x_i \in C_j$.

Algoritmo 1: Algoritmo K-means

Input: Conjunto dos *pixels* da imagem de entrada $X = \{x_i\}_{i=1}^n$, número de clusters k , número máximo de iterações t_{\max}

Output: Imagem segmentada

Método:

Inicializar os centroides $\{\mu_j\}_{j=1}^k$ dos clusters com valores aleatórios

repetir

{Etapa de atribuição}

para $i = 1$ até n **fazer**

 Calcular distâncias $d(x_i, \mu_j) = \|x_i - \mu_j\|$

 Atribuir para cada $x_i \in X$ um $x_i \in C_j$, onde $j = \operatorname{argmin} d(x_i, \mu_j)$

fim do para

{Etapa de re-estimação}

 Calcular os novos centros dos clusters

$t = t + 1$

até convergir ou $t = t_{\max}$

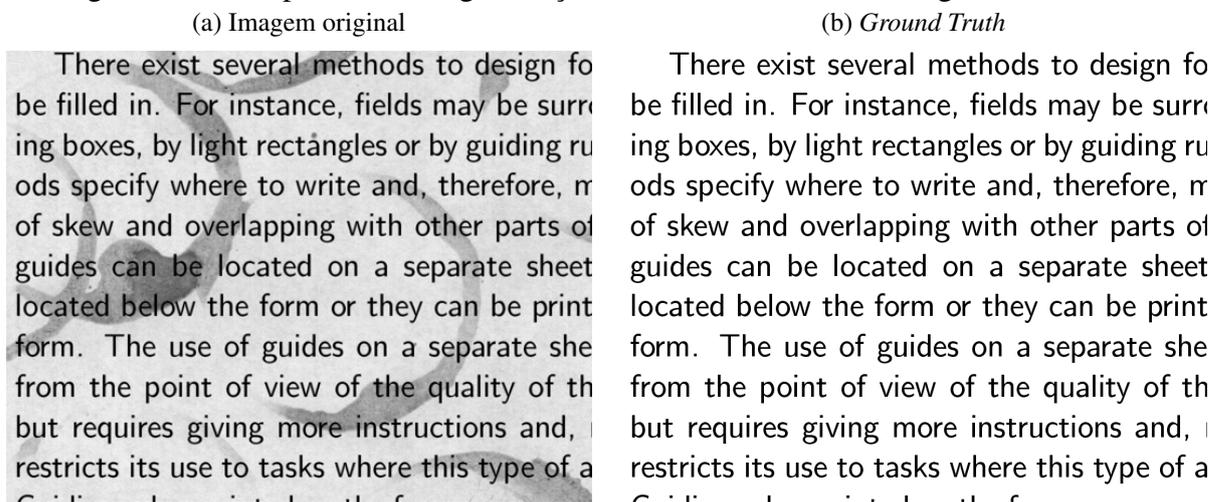
2.1.3 Avaliação da Segmentação

Em sistemas de Processamento Digital de Imagens (PDI), a segmentação, dependendo da aplicação, pode ser uma das operações mais difíceis e determinante no sucesso ou falha na etapa de reconhecimento da imagem. Esquef *et al.* (2003) afirmam que dentre os passos fundamentais do processamento de imagens a segmentação é a mais crítica do tratamento de

informações por ser a etapa em que são definidas as regiões de interesses que serão analisadas nas etapas posteriores. É importante que a segmentação seja realizada de forma rigorosa, para que não haja erro refletindo nos processos seguintes.

As métricas de avaliação da segmentação de imagens podem ser divididas em dois principais métodos: supervisionado, no qual é avaliado o resultado com um *Ground Truth* (GT), que consiste em um padrão de referência para segmentação, e não-supervisionado, quando não é necessário o uso de um GT (ZHANG *et al.*, 2008). O GT é uma imagem, geralmente segmentada manualmente, das regiões que representam uma segmentação ideal. Na Figura 3 há um exemplo de uma imagem não segmentada e de um GT referente ao texto da imagem.

Figura 3 – Exemplo de uma segmentação ideal do texto de uma imagem com ruídos.



Fonte: (MARTINEZ *et al.*, 2007)

Métodos de avaliação supervisionado, também conhecido como métodos de discrepância empírica, medem a diferença entre a saída de um algoritmo de segmentação e o seu GT. Uma das medições mais simples é a probabilidade do erro, que para a segmentação de uma imagem em objeto-*background* é dada por:

$$P(\text{error}) = P(O)P(B|O) + P(B)P(O|B), \quad (2.14)$$

em que $P(B|O)$ é a probabilidade em classificar um objeto como *background* e $P(O|B)$ a de classificar *background* como objeto. $P(O)$ e $P(B)$ são a probabilidade a priori para as classes objeto e *background*, respectivamente, obtidas a partir do GT.

Outra maneira de avaliar o resultado da segmentação é analisá-lo como um problema de classificação e considerar as regiões como classes. Neste caso, um dos recursos que pode ser

utilizado para comparar a predição com a referência é uma matriz de confusão (SOKOLOVA; LAPALME, 2009).

Também conhecidos como métodos empíricos de qualidade, os métodos de avaliação não-supervisionados permitem analisar uma imagem considerando o quão bem ela corresponde a um conjunto de características de imagens segmentadas sem a necessidade de uma segmentação ideal (CHABRIER *et al.*, 2006). Consiste em pontuar e classificar a segmentação por meio de medidas de qualidades estabelecidas a partir da percepção humana do que seria uma boa segmentação.

Para o desenvolvimento deste trabalho foram utilizadas métricas de avaliação supervisionadas obtidas a partir da matriz de confusão: acurácia, precisão, sensibilidade e o índice Sørensen-Dice para comparar o desempenho do método proposto em imagens que possuem GT em relação a outros métodos. Também é feita uma análise qualitativa, onde é avaliado visualmente a qualidade da segmentação.

2.1.3.1 Matriz de Confusão

A matriz de confusão é uma matriz que apresenta a correspondência entre as classes reais e as preditas pelo algoritmo de classificação (BURNEY; TARIQ, 2014). É definida como na Tabela 1 para um problema binário, no qual as linhas indicam a classe obtida pelo algoritmo e as colunas a saída original. Com isso é possível obter informações como o número de predições corretas ao dizer que uma amostra pertence a uma classe e de não pertencer a essa classe, chamadas de Verdadeiro Positivo (VP) e Verdadeiro Negativo (VN), respectivamente. Conseqüentemente, também o número de predições incorretas ao classificar uma amostra como 1 ou como 0, chamadas de Falso Positivo (FP) e Falso Negativo (FN), respectivamente.

Tabela 1 – Matriz de confusão

		Classe Real	
		1	0
Classe Predita	1	VP	FP
	0	FN	VN

Fonte: o autor.

Diversas métricas podem ser utilizadas para avaliar algoritmos de classificação a partir da matriz de confusão. Dentre elas, seja para problemas binários ou multiclases, a mais utilizada é a acurácia (HOSSIN; SULAIMAN, 2015). Esta métrica mede a proporção de

previsões corretas em relação ao total de amostras, como se observa na Equação 2.15. Outras medidas bastante utilizadas também são a precisão e a sensibilidade ou *Recall*. Enquanto a precisão mede a proporção das amostras preditas como positivas classificadas corretamente, a sensibilidade representa a taxa de amostras positivas do GT foram identificadas corretamente. Essa métricas são definidas nas Equações 2.16 e 2.17, respectivamente.

$$Ac = \frac{VP + VN}{VP + VN + FP + FN} \quad (2.15)$$

$$Pr = \frac{VP}{VP + FP} \quad (2.16)$$

$$Se = \frac{VP}{VP + FN} \quad (2.17)$$

Em um problema multiclasse as métricas são calculadas para cada classe separada. Considerando uma classe como a classe positiva, as outras classes são consideradas como a negativa e assim é tratado como um problema binário. Em segmentação, essas métricas podem ser utilizadas para avaliar a qualidade de uma imagem segmentada, considerando cada região como uma classe e cada *pixel* como um elemento de entrada do classificador.

2.1.3.2 Coeficiente de Similaridade de Dice

Na literatura é possível encontrar diversos trabalhos que desenvolveram coeficientes baseados em regiões para medidas de sobreposição espacial, utilizados para calcular a similaridade entre duas amostras (ZAKANI *et al.*, 2016). Um dos índices mais comum para esta comparação é o CSD (AFROZ, 2014). Proposto pelos botânicos Sørensen (1948) e por Dice (1945), o CSD, também encontrado na literatura como Índice Sørensen-Dice, é uma estatística que mede a associação entre diferentes amostras. É definida por

$$CSD(V_1, V_2) = \frac{2|V_1 \cap V_2|}{|V_1| + |V_2|}, \quad (2.18)$$

em que V_1 e V_2 são conjuntos de dados, $|V|$ denota a cardinalidade do conjunto e $0 \leq CSD(V_1, V_2) \leq 1$.

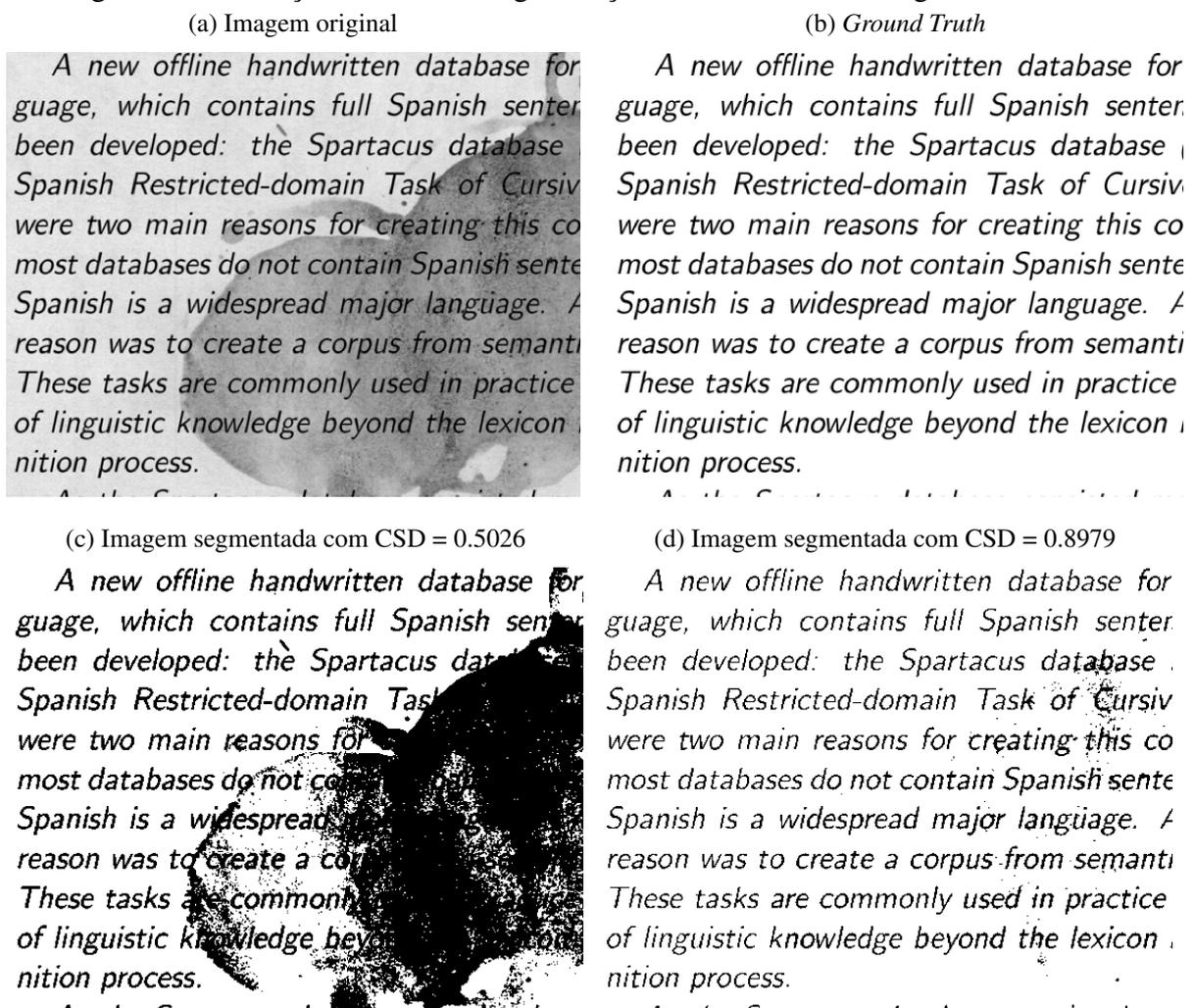
Podendo ser interpretada como uma medida de similaridade, a métrica CSD já vem sendo usada amplamente para avaliação do desempenho de segmentação, como em análise de lesões de substância branca no cérebro humano (ZIJDENBOS *et al.*, 1994), doenças renais em imagens de tomografia abdominal (GOMALAVALLI *et al.*, 2016) e similaridade entre palavras (KONDRAK *et al.*, 2003). É possível estimar a performance de um algoritmo de segmentação

considerando o CSD como uma medida para a sobreposição espacial entre a imagem segmentada e o GT. Considerando a matriz de confusão, a Equação 2.18 pode ser reescrita como

$$CSD(V_1, V_2) = \frac{2VP}{2VP + FN + FP}. \quad (2.19)$$

Zijdenbos *et al.* (1994) afirma que um CSD acima de 0,7 indica uma boa medida para segmentação. Na Figura 4 é possível comparar a avaliação do CSD para diferentes resultados da segmentação de uma mesma imagem.

Figura 4 – Avaliação do CSD na segmentação do texto de uma imagem com ruídos.



Fonte: Adaptado de Martinez *et al.* (2007)

2.1.3.3 Análise Qualitativa

Embora haja várias técnicas para avaliar o desempenho da segmentação, ainda é difícil de estimar, seja para imagens específicas ou um conjunto de imagens genéricas, qual

algoritmo apresenta segmentações mais precisas que outro. O método mais comum para determinar o melhor resultado é a avaliação subjetiva, onde é comparado visualmente por um humano as imagens obtidas para cada técnica de segmentação (ZHANG *et al.*, 2008). Os padrões necessários para que o resultado de uma segmentação seja considerado um bom resultado podem variar dependendo da pessoa que avalia, contudo a diferença entre uma boa imagem segmentada e uma inferior é perceptível.

A capacidade de comparar resultados de segmentação sem a necessidade de um algoritmo é uma tarefa importante. Zhang *et al.* (2008) enumera três motivos: (1) selecionar entre duas segmentações, seja para parâmetros ou técnicas diferentes; (2) propor uma nova técnica de segmentação ou apresentar um estudo de uma já existente; e (3) garantir o desempenho da segmentação monitorando os resultados em tempo real. Para muitos problemas, apenas o desenvolvedor do sistema de segmentação não é o suficiente para analisar os resultados e se faz necessário a avaliação de especialistas.

2.2 Entropia

A entropia é uma palavra que abrange diversos conceitos que se distinguem dependendo de sua aplicação (THURNER *et al.*, 2017). Devido a sua complexidade, a entropia não possui uma definição padronizada, podendo ser encontrado na literatura científica clássica em diferentes formas como taxa, quantidade e também como unidade de medida (PAVIOTTI, 2019). Originalmente introduzida na termodinâmica, pode ser encontrada em estudos na física estatística, teoria da informação e interferência estatística. Na teoria da informação, contexto deste trabalho, pode ser interpretada como uma medida-chave de informação em sistemas dinâmicos.

Shannon (1948) foi o primeiro a deduzir que a informação está ligada a probabilidade, ao definir a entropia como uma medida de uma distribuição de probabilidade $p = p_i$, em que

$$\sum_{i=1}^n p_i = 1, \quad p_i \geq 0 \forall i, \quad (2.20)$$

é possível obter uma medida de incerteza média associada a uma variável aleatória por meio da quantidade, expressa por

$$H(p) = - \sum_i p_i \ln p_i, \quad (2.21)$$

que passou a ser chamada de entropia de Shannon. A entropia como medida de incerteza desempenha um papel crucial no estudo da informação, sendo tão fundamental para a ciência da informação quanto o teorema de Pitágoras é para a geometria (TRIBUS; MCIRVINE, 1971).

Para mesma distribuição de probabilidade, Tsallis (1988) definiu

$$H_q(p) = \frac{1}{q-1} \left(1 - \sum_{i=1}^n p_i^q \right), \text{ para } q \in [0, \infty), \quad (2.22)$$

conhecida como entropia de Tsallis. Essa expressão é considerada uma generalização da entropia de Shannon, visto que no limite de $q \rightarrow 1$ a entropia de Tsallis se aproxima da entropia de Shannon. Ambas as entropias podem ser relacionadas por expressões análogas, uma vez que é possível reescrever a Equação 2.21 como

$$H(p) = - \sum_{i=1}^n \frac{\ln(p_i)}{\ln'(p_i)}, \quad (2.23)$$

e que a entropia de Tsallis pode ser definida em termos de q -logaritmo de Tsallis (TSALLIS, 1994). A Equação 2.22 pode ser expressa como a Equação 2.23, apenas substituindo $\ln(\cdot)$ por $\ln_q(\cdot)$, o que resulta em

$$H(p) = - \sum_{i=1}^n \frac{\ln_q(p_i)}{\ln'_q(p_i)}. \quad (2.24)$$

A entropia de Tsallis desempenha um papel crucial nas estatísticas não extensivas ou nas estatísticas de Tsallis.

2.2.1 Entropia Relativa

Enquanto a entropia permite medir o grau de incerteza de uma variável aleatória, a entropia relativa diz o quão bem uma distribuição de probabilidade g se aproxima de uma distribuição p . A entropia relativa, também conhecida como divergência estatística, possui um papel fundamental na geometria da informação (ZHANG, 2004). Foi primeiramente introduzida no contexto da estatística por Kullback e Leibler (1951) ao definirem a entropia relativa de Shannon, também chamada de divergência KL, como uma medida de distância estatística entre duas distribuições de probabilidades.

Dado o conjunto de probabilidades estritamente positivas, como o conjunto

$$\Delta_n = \{(p_1, \dots, p_n) : \sum_{i=1}^n p_i = 1\}, \quad p_i > 0 \forall i, \quad (2.25)$$

Kullback e Leibler (1951) estabeleceram a expressão

$$D(p||g) = \sum_{i=1}^n p_i \ln \left(\frac{p_i}{g_i} \right). \quad (2.26)$$

Essa expressão pode ser interpretada como uma expressão análoga não-simétrica da distância euclidiana quadrática (CSISZÁR; SHIELDS, 2004). A entropia relativa de Tsallis (BORLAND

et al., 1998), generalização da entropia relativa de Shannon, é definida ao substituir o logaritmo natural pelo q -logaritmo na divergência KL. Para distribuições de probabilidade p e g pertencendo a Δ_n em 2.25 é dada por

$$D_q(p||g) = - \sum_{i=1}^n p_i \ln_q \left(\frac{g_i}{p_i} \right). \quad (2.27)$$

2.2.2 Entropia Relativa Generalizada

A entropia relativa generalizada é obtida a partir do uso de uma deformada exponencial φ . Uma exponencial deformada é uma função $\varphi : \mathbb{R} \rightarrow [0, \infty)$ que satisfaz as seguintes propriedades:

- $\varphi(\cdot)$ é convexa;
- $\lim_{u \rightarrow -\infty} \varphi(u) = 0$ e $\lim_{u \rightarrow \infty} \varphi(u) = \infty$;
- Existe um vetor $\mathbf{u}_0 = (u_{0,i})$, com $u_{0,i} \in (0, \infty)$, tal que

$$\sum_{i=1}^n \varphi(c_i + \lambda u_{0,i}) < \infty, \quad \forall \lambda > 0, \quad (2.28)$$

para todo vetor $\mathbf{c} = c_i$, para cada $c_i \in \mathbb{R}$ e $\sum_{i=1}^{\infty} \varphi(c_i) = 1$.

Considerando que a entropia relativa de Shannon, a Equação 2.26, pode ser reescrita como

$$D(p||g) = \sum_{i=1}^n \frac{\ln(p_i) - \ln(g_i)}{\ln'(p_i)}. \quad (2.29)$$

Uma vez que

$$-p_i \frac{(g_i/p_i)^{1-g} - 1}{1-g} = \frac{p_i - g_i^{1-g} p_i^g}{1-g} = \frac{1}{p_i^{-g}} \left(\frac{p_i^{1-g} - 1}{1-g} - \frac{g_i^{1-g} - 1}{1-g} \right), \quad (2.30)$$

a entropia relativa de Tsallis, dada na Equação 2.27, pode ser expressa por

$$D_q(p||g) = \sum_{i=1}^n \frac{\ln_q(p_i) - \ln_q(g_i)}{\ln'_q(p_i)}. \quad (2.31)$$

Observa-se que as entropias relativas de Shannon e Tsallis apresentam expressões similares diferenciando apenas pelo uso do $\ln(\cdot)$ e $\ln_q(\cdot)$.

A partir da similaridade entre as entropias relativas de Shannon e Tsallis é possível generalizá-las substituindo as funções q -logaritmo $\ln_q(\cdot)$ e logaritmo natural $\ln(\cdot)$ pelo inverso da exponencial deformada φ (Vigelis *et al.*, 2020) (SOUZA *et al.*, 2016) (VIGELIS; CAVALCANTE, 2015) (VIGELIS; CAVALCANTE, 2013). A entropia relativa generalizada entre duas

distribuição de probabilidades $p = p_i$ e $g = g_i$ em Δ_n , dada uma função exponencial deformada $\varphi : \mathbb{R} \rightarrow [0, \infty)$, é expressa por

$$D_\varphi(p||g) = \sum_{i=1}^n \frac{\varphi^{-1}(p_i) - \varphi^{-1}(g_i)}{(\varphi^{-1})'(p_i)}. \quad (2.32)$$

Em consequência da concavidade de $\varphi^{-1}(x)$ têm-se que

$$(y-x)(\varphi^{-1})'(y) \leq \varphi^{-1}(y) - \varphi^{-1}(x), \quad \forall x, y > 0. \quad (2.33)$$

Considerando $y = p_i$ e $x = g_i$, e usando a desigualdade dada na Equação 2.33, é possível escrever

$$D_\varphi(p||g) = \sum_{i=1}^n \frac{\varphi^{-1}(p_i) - \varphi^{-1}(g_i)}{(\varphi^{-1})'(p_i)} \geq \sum_{i=1}^n (p_i - g_i) = 0. \quad (2.34)$$

Sendo $\varphi(x)$ estritamente convexa, é claro que $D_\varphi(p||g) = 0$ se $p = g$.

2.3 Divergências de Bregman

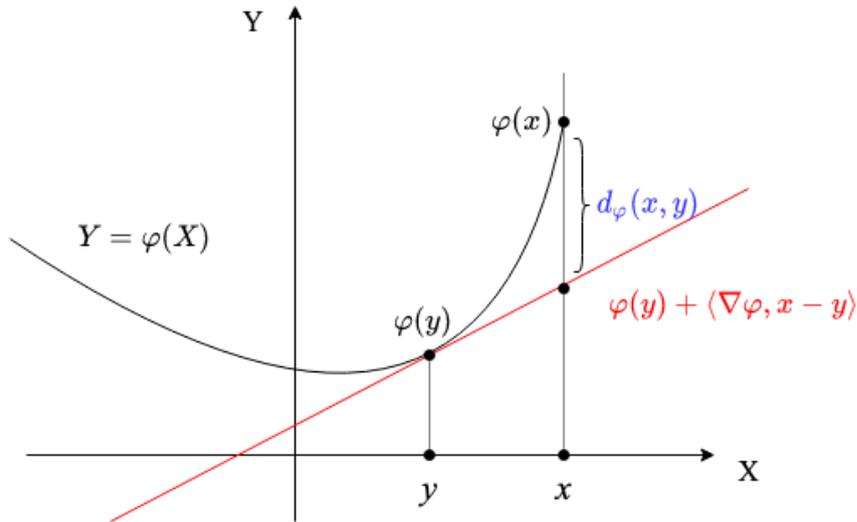
Divergências de Bregman, ou DBs, são medidas de dissimilaridade não métricas que consistem em determinar a similaridade entre dois pares de dados complexos, representados por números, vetores, matrizes ou funções densidade de probabilidades (FERREIRA, 2017). Dada uma função estritamente convexa, $\varphi : S \rightarrow \mathbb{R}$, definida em um conjunto convexo $S \subseteq \mathbb{R}^d$ em que $S = \text{dom}(\varphi)$, tal que φ é diferenciável em $\text{ri}(S)$ e não vazio, a divergência de Bregman $d_\varphi : S \times \text{ri}(S) \rightarrow [0, \infty)$ entre x e y é definida por

$$d_\varphi(x, y) = \varphi(x) - \varphi(y) - \langle x - y, \nabla \varphi(y) \rangle, \quad (2.35)$$

sendo $\nabla \varphi(y)$ o vetor gradiente de φ avaliado em y e $\text{ri}(S)$ o interior relativo de S (BREGMAN, 1967). Na Figura 5 é exemplificado essa definição no espaço 1-dimensional S .

É possível observar na Figura 5 que diferentes escolhas para a função φ leva a diferentes métricas. Dentre as DBs, a distância Euclidiana quadrática é talvez a mais simples e utilizada. Sendo a função base $\varphi(x) = \langle x, x \rangle$ estritamente convexa e diferenciável no \mathbb{R}^d tem-se

$$\begin{aligned} d_\varphi(x, y) &= \langle x, x \rangle - \langle y, y \rangle - \langle x - y, \nabla \varphi(y) \rangle \\ &= \langle x, x \rangle - \langle y, y \rangle - \langle x - y, 2y \rangle \\ &= \langle x - y, x - y \rangle = \|x - y\|^2. \end{aligned}$$

Figura 5 – Interpretação geométrica da divergência de Bregman no \mathbb{R}^2 

Fonte: (FERREIRA, 2017)

Outro exemplo bastante utilizado é a divergência KL. Para uma distribuição de probabilidade discreta p em que $\sum_{i=1}^d p_i = 1$, a entropia negativa $\varphi(p) = \sum_{i=1}^d p_i \log_2 p_i$ é uma função convexa. A divergência de Bregman correspondente é

$$\begin{aligned}
 d_\varphi(p, q) &= \sum_{i=1}^d p_i \log_2 p_i - \sum_{i=1}^d q_i \log_2 q_i - \langle p - q, \nabla \varphi(q) \rangle \\
 &= \sum_{i=1}^d p_i \log_2 p_i - \sum_{i=1}^d q_i \log_2 q_i - \sum_{i=1}^d (p_i - q_i) (\log_2 q_i + \log_2 e) \\
 &= \sum_{i=1}^d p_i \log_2 \left(\frac{p_i}{q_i} \right) - \log_2 e \sum_{i=1}^d (p_i - q_i) \\
 &= KL(p \| q),
 \end{aligned}$$

resultando na divergência KL entre as duas distribuições de probabilidade, visto que $\sum_{i=1}^d q_i = \sum_{i=1}^d p_i = 1$. Na Tabela 2 são listadas algumas funções convexas comuns e suas DBs.

Tabela 2 – Divergências de Bregman geradas a partir de algumas funções convexas

Domínios	$\varphi(x)$	$d_\varphi(x, y)$	Divergência
\mathbb{R}	x^2	$(x - y)^2$	Perda quadrática
\mathbb{R}_{++}	$-\log x$	$\frac{x}{y} - \log\left(\frac{x}{y}\right) - 1$	Distância Itakura-Saito
\mathbb{R}^d	$\ x\ ^2$	$\ x - y\ ^2$	Euclidiana quadrática
\mathbb{R}^d	$x^T A x$	$(x - y)^T A (x - y)$	Distância Mahalanobis
$d - \text{Simplex}$	$\sum_{j=1}^d x_j \log_2 x_j$	$\sum_{j=1}^d x_j \log_2 \frac{x_j}{y_j}$	Divergência KL
\mathbb{R}_+	$\sum_{j=1}^d x_j \log x_j$	$\sum_{j=1}^d x_j \log \frac{x_j}{y_j} - \sum_{j=1}^d (x_j - y_j)$	Generalized I-Divergence

As DBs são similares às métricas, mas não satisfazem a desigualdade triangular e

geralmente não são simétricas. Banerjee *et al.* (2005) apresenta algumas propriedades das DBs, em principal:

- **Não negatividade:** $d_\varphi(x, y) \geq 0$, $\forall x \in S$, $y \in \text{ri}(S)$, e $d_\varphi(x, y) = 0$ se e somente se $x = y$. Isso é em consequência da convexidade de φ .
- **Convexidade:** d_φ é sempre convexa no primeiro argumento, mas não necessariamente no segundo argumento. A divergência de Bregman correspondente à função estritamente convexa $\varphi(x) = x^3$, definida em \mathbb{R}_+ e dada por $d_\varphi(x, y) = x^3 - y^3 - 3(x - y)y^2$ é um exemplo de divergência que não é convexa em y . Já a distância Euclidiana quadrática e a divergência KL são exemplo de DBs que são convexas em seus dois argumentos.
- **Linearidade:** A divergência de Bregman é um operador linear, ou seja, $\forall x \in S$, $y \in \text{ri}(S)$,

$$d_{\varphi_1 + \varphi_2} = d_{\varphi_1}(x, y) + d_{\varphi_2}(x, y),$$

$$d_{c\varphi}(x, y) = cd_\varphi(x, y), \quad \text{para } c \geq 0.$$

- **Separação linear:** O conjunto de todos os pontos $x \in S$ que são equidistantes de dois pontos fixos $\mu_1, \mu_2 \in \text{ri}(S)$ em termos de uma divergência de Bregman é um hiperplano, ou seja, as partições induzidas pelas DBs têm separadores lineares dados por

$$d_\varphi(x, \mu_1) = d_\varphi(x, \mu_2)$$

$$\Rightarrow \varphi(x) - \varphi(\mu_1) - \langle x - \mu_1, \nabla\varphi(\mu_1) \rangle = \varphi(x) - \varphi(\mu_2) - \langle x - \mu_2, \nabla\varphi(\mu_2) \rangle$$

$$\Rightarrow \langle x, \nabla\varphi(\mu_2) - \nabla\varphi(\mu_1) \rangle - (\varphi(\mu_1) - \varphi(\mu_2)) - (\langle \mu_1, \nabla\varphi(\mu_1) \rangle - \langle \mu_2, \nabla\varphi(\mu_2) \rangle)$$

- **Teorema de Pitágoras generalizado:** Para qualquer $x_1 \in S$ e $x_2, x_3 \in \text{ri}(S)$ a seguinte propriedade se mantém:

$$d_\varphi(x_1, x_2) + d_\varphi(x_2, x_3) \leq d_\varphi(x_1, x_3).$$

Diversas DBs, como distância Euclidiana quadrática, distância de Mahalanobis, distância de Itakura-Saito e entropia relativa de Shannon, são utilizadas em algoritmos de clusterização como medida de dissimilaridade para o agrupamento de dados (MACQUEEN, 1967), (LINDE *et al.*, 1980), (BUZO *et al.*, 1980), (DHILLON *et al.*, 2003). Estes algoritmos são conhecidos como clusterização de Bregman. Os métodos de clusterização podem ser divididos em duas principais categorias: *hard* e *soft*. Os algoritmos de clusterização *hard* particionam os dados de forma que cada ponto pertença a apenas um cluster, já na clusterização *soft* cada ponto possui uma certa probabilidade de pertencer a cada um dos clusters.

2.3.1 Clusterização Hard de Bregman

Utilizando a divergência de Bregman para medir a taxa de distorção de Shannon, Banerjee *et al.* (2005) introduziu o conceito de informação de Bregman ao interpretar o problema de encontrar uma partição e seus representantes como um problema de perda da informação de Bregman. A informação de Bregman é a função ótima que visa minimizar a distorção entre uma variável aleatória de origem e uma variável aleatória decodificada. Seja X uma variável aleatória com valores em $\mathcal{X} = \{x_1, \dots, x_n\} \subset S \subseteq \mathbb{R}^d$ seguindo a medida de probabilidade discreta ν e a divergência de Bregman d_φ , a informação de Bregman de X com relação a φ é expressa por

$$I_\varphi(X) = \min_{s \in \text{ri}(S)} E_\nu[d_\varphi(X, s)] = \min_{s \in \text{ri}(S)} \sum_{i=1}^n \nu_i d_\varphi(x_i, s). \quad (2.36)$$

O vetor s responsável por atingir a distorção mínima entre as variáveis de origem e a decodificada é chamado de representante de Bregman de X , sendo este representante independente da escolha da divergência de Bregman (BANERJEE *et al.*, 2005). Portanto, assumindo

$$s = \mu = \sum_{i=1}^n \nu_i x_i, \quad (2.37)$$

é possível reescrever a Equação 2.36 como

$$I_\varphi(X) = \sum_{i=1}^n \nu_i d_\varphi(x_i, \mu). \quad (2.38)$$

Dado um conjunto de observação que assume valores em $\mathcal{X} = \{x_1, \dots, x_n\} \subset S \subseteq \mathbb{R}^d$ e com medida de probabilidade ν , a clusterização *hard* de Bregman é feita particionando o conjunto \mathcal{X} em k partições disjuntas $\{\mathcal{X}_h\}_{h=1}^k$, em que cada variável é atribuída a apenas uma partição e cada partição possua seu representante de Bregman. O conjunto de representantes é denotado por $\mathcal{M} = \{\mu_h\}_{h=1}^k$ e $\pi = \{\pi_h\}_{h=1}^k$ com $\pi_h = \sum_{x_i \in \mathcal{X}_h} \nu_i$ denotando a probabilidade induzida medida em \mathcal{M} .

Similar ao algoritmo do *K-means* clássico, para implementar a clusterização, primeiramente é definido o número de clusters k . Em seguida é inicializado o conjunto de representantes \mathcal{M} com valores aleatórios, sendo μ_j o representante do cluster \mathcal{X}_h . Posteriormente é realizada a etapa de atribuição, em que cada $x_i \in \mathcal{X}$ é associado a $x_i \in \mathcal{X}_h$ que atenda $\min_{\mu \in \mathcal{M}} d_\varphi(x_i, \mu_h)$. Por fim é executado a etapa de re-estimação, onde são definidos novos representantes para o conjunto \mathcal{M} , sendo μ_h definido pela média dos dados em \mathcal{X}_h . As etapas de atribuição e re-estimação são repetidas até que o algoritmo convirja.

Seja X uma variável aleatória que assume valores em \mathcal{X} , a clusterização é operada de modo que uma variável M , a partir dos representantes, sirva como uma quantização apropriada de X , assumindo valores em \mathcal{M} seguindo π . A qualidade da quantização M pode ser medida a partir da divergência de Bregman esperada entre X e M , $E_{X,M}[d_\varphi(X,M)]$. Como M é uma função determinística de X , a expectativa é apenas sobre a distribuição de X , tendo assim

$$E_X[d_\varphi(X,M)] = \sum_{h=1}^k \sum_{x_i \in \mathcal{X}_h} v_i d_\varphi(x_i, \mu_h) = \sum_{h=1}^k \pi_h \sum_{x_i \in \mathcal{X}_h} \frac{v_i}{\pi_h} d_\varphi(x_i, \mu_h) = E_\pi[I_\varphi(X_h)], \quad (2.39)$$

onde X_h é uma variável aleatória que assume valores em \mathcal{X}_h seguindo a distribuição de probabilidade $\frac{v_i}{\pi_h}$ e $I_\varphi(X_h)$ a informação de Bregman de X_h .

Dhillon *et al.* (2003) medem a qualidade do particionamento em termos da perda de informações mútuas resultante da quantização da variável aleatória original X . Estendendo essa formulação, Banerjee *et al.* (2005) apresenta uma maneira alternativa de medir a qualidade da quantização M pela perda da informação de Bregman, expressa por

$$L_\varphi(M) = I_\varphi(X) - I_\varphi(M) = \sum_{h=1}^k \sum_{i=1}^n v_i d_\varphi(x_i, \mu_h), \quad (2.40)$$

sendo $I_\varphi(X)$ a informação total de Bregman, $I_\varphi(M)$ a informação de Bregman inter-cluster, e a perda de informação de Bregman, $L_\varphi(M)$, representa a informação de Bregman intra-cluster. A partir disso, a clusterização *hard* de Bregman (Algoritmo 2) pode ser interpretada como um algoritmo que visa encontrar o particionamento de \mathcal{X} que minimize a perda da informação de Bregman. Comparando os algoritmos 1 e 2 é possível observar que a clusterização *hard* de Bregman e o *K-means* são métodos análogos.

Algoritmo 2: Clusterização *Hard* de Bregman

Input: Conjunto $\mathcal{X} = \{x_i\}_{i=1}^n \subset S \subseteq \mathbb{R}^d$; medida de probabilidade de v sobre \mathcal{X} ;
divergência de Bregman $d_\varphi : S \times \text{ri}(S) \rightarrow \mathbb{R}$ e o número de clusters k

Output: \mathcal{M} ; agrupamento *hard* $\{\mathcal{X}_h\}_{h=1}^k$ de \mathcal{X} ; minimizador local de $L_\varphi(\mathcal{M})$, onde
 $\mathcal{M} = \{\mu_h\}_{h=1}^k$ de \mathcal{X}

Método:

Inicializa $\{\mu_h\}_{h=1}^k$ com $\mu_h \in \text{ri}(S)$ com valores aleatórios

repetir

{Etapa de atribuição}

$\mathcal{X}_h \leftarrow \emptyset, 1 \leq h \leq k$

para $i = 1$ até n **fazer**

$\mathcal{X}_h \leftarrow \mathcal{X}_h \cup \{x_i\}$, onde $h = \text{argmin} d_\varphi(x_i, \mu_h)$

fim do para

{Etapa de re-estimação}

para $h = 1$ até k **fazer**

$\pi_h \leftarrow \sum_{x_i \in \mathcal{X}_h} v_i$

$\mu_h \leftarrow \frac{1}{\pi_h} \sum_{x_i \in \mathcal{X}_h} v_i$

fim do para

até convergir

$\mathcal{M} \leftarrow \{\mu_h\}_{h=1}^k$

3 METODOLOGIA

Baseado nas técnicas descritas anteriormente, neste capítulo são abordados os materiais necessários para o desenvolvimento do método proposto. Inicialmente, na seção 3.1 discute-se sobre as imagens utilizadas para os testes de segmentação. Considerando que para o uso da divergência generalizada é necessário especificar a função exponencial φ , em seguida, na seção 3.2, é descrito o equacionamento das funções exponenciais utilizadas. Finalmente, na seção 3.3 é discutida a implementação do algoritmo de clusterização proposto.

3.1 Banco de imagens

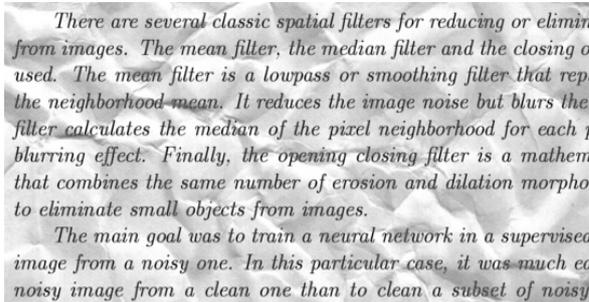
O primeiro passo para executar o projeto é definir as imagens para os testes. Como a segmentação por clusterização baseada na divergência generalizada não está sendo proposto para um conjunto de imagens específico, são utilizadas imagens de aplicações gerais. Visando apresentar uma melhor aplicabilidade, o método de segmentação proposto é testado em imagens em escala de cinza e coloridas.

Para os testes em escala de cinza é utilizado imagens do conjunto de dados NoisyOffice (MARTINEZ *et al.*, 2007) que fornece imagens de textos escaneados com ruídos reais e imagens com quatro tipos de ruídos simulados. O *dataset* também disponibiliza o GT das imagens com ruídos simulados. As simulações são referentes a folhas dobradas, folhas amassadas, manchas de copo e pegadas. Na Figura 6 é possível observar um exemplo para cada ruído. Além do banco NoisyOffice, em que a segmentação é realizada como um problema de binarização para separar duas regiões (texto e fundo da imagem), também são realizados testes com imagens de outros conjuntos de dados (ARBELAEZ *et al.*, 2011), (GONZALEZ; WOODS, 2011), em que há mais de duas regiões, visando aplicações gerais.

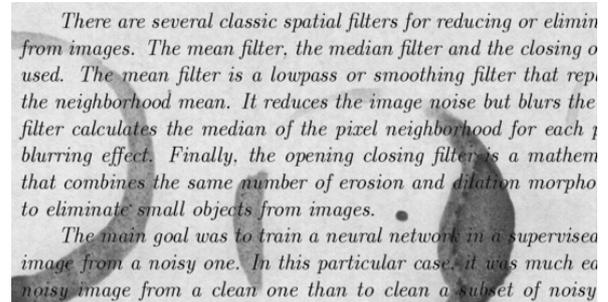
Os testes realizados em imagens coloridas utilizam amostras de *datasets* variados, Gonzalez e Woods (2011), Mamaev (2018) e Arbelaez *et al.* (2011). Como as imagens utilizadas não possuem um padrão de referência para as regiões, os GT são criados manualmente. A definição das regiões nas imagens é feita, de forma empírica, do que representaria uma segmentação ideal. Na Figura 7 é apresentado um exemplo de GT gerado manualmente, considerando a segmentação ideal da imagem para 3 regiões: as flores; os caules e folhas; e o céu. Outro exemplo é observado na Figura 8 para a segmentação ideal de uma região específica da imagem.

Figura 6 – Exemplos de ruídos simulados do banco de imagens NoiseOffice.

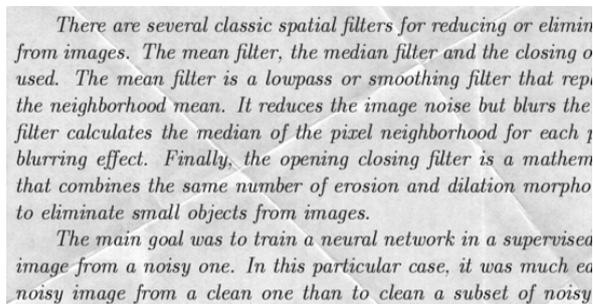
(a) Folha amassada



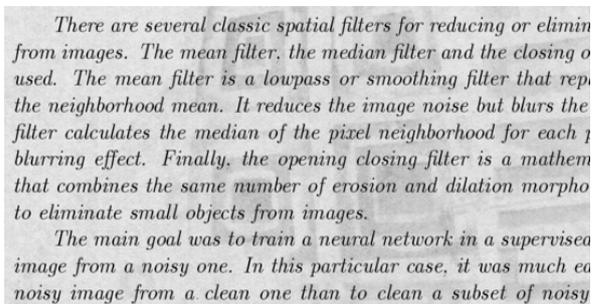
(b) Mancha de copo



(c) Folha dobrada



(d) Pegada



(e) Ground truth

There are several classic spatial filters for reducing or eliminating from images. The mean filter, the median filter and the closing operator are used. The mean filter is a lowpass or smoothing filter that replaces the neighborhood mean. It reduces the image noise but blurs the image. The median filter calculates the median of the pixel neighborhood for each pixel, which is effective for removing salt and pepper noise. Finally, the opening and closing filters are mathematical morphology operations that combine the same number of erosion and dilation morphological operations to eliminate small objects from images.

The main goal was to train a neural network in a supervised learning task to clean a noisy image from a clean one. In this particular case, it was much easier to clean a noisy image from a clean one than to clean a subset of noisy

Fonte: Martinez *et al.* (2007).

Figura 7 – Exemplo de segmentação ideal para 3 regiões em uma imagem

(a) Imagem original



(b) Ground truth para três regiões

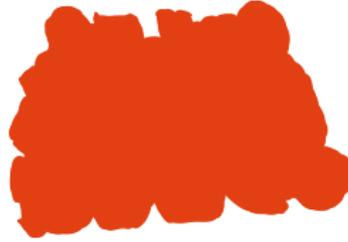


Fonte: Adaptado de Mamaev (2018).

Figura 8 – Exemplo de segmentação de uma região específica

(a) Imagem original

(b) GT para uma região específica



Fonte: Adaptado de Arbelaez *et al.* (2011)

3.2 Funções exponenciais deformadas

Como observado na Equação 2.32, a divergência generalizada é definida em termos de uma função exponencial deformada $\varphi(\cdot)$. Nesse trabalho é testado o uso de três funções, as quais denominamos φ_1 , φ_2 e φ_3 . A primeira função exponencial utilizada é dada por

$$\varphi_1(x) = \ln(\exp(x) + 1), \quad (3.1)$$

cuja a função inversa e a derivada da função inversa é dada, respectivamente, por

$$\varphi_1^{-1}(x) = \ln(\exp(x) - 1), \quad (3.2)$$

$$(\varphi_1^{-1})'(x) = \frac{\exp(x)}{\exp(x) - 1}. \quad (3.3)$$

A segunda função exponencial utilizada é a função κ -exponencial proposta por Kaniadakis (2001), sendo definida por

$$\varphi_2(x) = \exp_{\kappa}(x) = \exp\left(\int_0^x \frac{dt}{\sqrt{1 + \kappa^2 t^2}}\right), \text{ para cada } -1 < \kappa < 1. \quad (3.4)$$

Visto que essa exponencial deformada é reduzida à exponencial padrão quando o parâmetro de deformação real κ se aproxima de zero, tem-se

$$\varphi_2(x) = \exp_{\kappa}(x) = \begin{cases} (\kappa x + \sqrt{1 + \kappa^2 x^2})^{\frac{1}{\kappa}}, & \text{se } \kappa \neq 0, \\ \exp x, & \text{se } \kappa = 0. \end{cases} \quad (3.5)$$

Para cada κ dado, a função \exp_{κ} que mapeia \mathbb{R} para \mathbb{R}_{++} , estritamente crescente e estritamente convexa. Assim sendo, a inversa da função κ -exponencial é dada por

$$\varphi_2^{-1}(x) = \ln_{\kappa}(x) = \begin{cases} \frac{x^{\kappa} - x^{-\kappa}}{2\kappa}, & \text{se } \kappa \neq 0, \\ \ln x, & \text{se } \kappa = 0, \end{cases} \quad (3.6)$$

cuja derivada é igual a

$$(\varphi_2^{-1})'(x) = \ln'_{\kappa}(x) = \begin{cases} \frac{x^{\kappa} + x^{-\kappa}}{2} \frac{1}{x}, & \text{se } \kappa \neq 0, \\ \frac{1}{x}, & \text{se } \kappa = 0. \end{cases} \quad (3.7)$$

A terceira função exponencial utilizada é definida pela função q -exponencial de Tsallis, expressa por

$$\varphi_3(x) = \exp_q(x) = \begin{cases} [1 + (1 - q)x]_+^{1/(1-q)}, & \text{se } q \in (0, 1], \\ \exp x, & \text{se } q = 1, \end{cases} \quad (3.8)$$

e que $[x]_+ = x$ para $x \geq 0$, e $[x]_+ = 0$ para $x < 0$. A função inversa de $\exp_q(x)$ é definida pelo q -logaritmo de Tsallis, que para $q \in (0, 1]$ é dada por

$$\varphi_3^{-1}(x) = \ln_q(x) = \frac{1}{1 - q}(x^{1-q} - 1), \quad (3.9)$$

sendo a derivada da função inversa $(\varphi_3^{-1})'(x) = \ln'_q(x) = \frac{1}{x^q}$.

3.3 Clusterização baseada na divergência generalizada

A implementação do método de clusterização proposto para segmentação de imagens baseia-se na estrutura de clusterização *hard* de Bregman para realizar o agrupamento dos *pixels* em regiões definindo a divergência generalizada, Equação 2.32, como medida de dissimilaridade entre pontos de dados. Como observado no Algoritmo 2, a medida da dissimilaridade é dada por $d_{\varphi} = (x_i, \mu_j)$. Desta forma, tem-se $p = x_i$ e $q = \mu_j$. Assim, a divergência generalizada pode ser reescrita como

$$D_{\varphi}(x_i, \mu_j) = \sum_{h=1}^m \frac{\varphi^{-1}(x_{ih}) - \varphi^{-1}(\mu_{jh})}{(\varphi^{-1})'(x_{ih})}, \quad (3.10)$$

em que m é dado pelo número de canais de cores da imagem. Sendo a divergência generalizada no domínio \mathbb{R}_+^d , as variáveis p e q não estão no Δ_n , Equação 2.25, visto que para essa aplicação assumem os valores dos *pixels* da imagem.

O conjunto de entrada $\mathcal{X} = \{x_i\}_{i=1}^n$ do algoritmo de clusterização é definido pelos *pixels* como características da imagem, em que cada *pixel* é dado como uma variável aleatória x_i . Para imagens em nível de cinza x_i é dado pelo valor da intensidade I do brilho, onde $I \in [0,255]$. Já para imagens coloridas é dado pelos três canais do padrão de cores RGB, sendo $R, G, B \in [0,255]$, ou no espaço HSV, em que $H \in [0,360)$, $S \in [0,1]$ e $V \in [0,1]$. Sendo MAX e MIN os valores máximo e mínimo, respectivamente, dos valores (R, G, B) , a conversão de RGB para HSV é dada nas Equações 3.11, 3.12, 3.13. Para ambas aplicações, escala cinza e coloridas, esses valores são trabalhados numa escala de $[0,1]$.

$$H = \begin{cases} 0, & \text{se } MAX = MIN, \\ 60 * \frac{G-B}{MAX-MIN} + 0, & \text{se } MAX = R \text{ e } G \geq B, \\ 60 * \frac{G-B}{MAX-MIN} + 360, & \text{se } MAX = R \text{ e } G < B, \\ 60 * \frac{B-R}{MAX-MIN} + 120, & \text{se } MAX = G, \\ 60 * \frac{R-G}{MAX-MIN} + 240, & \text{se } MAX = B, \end{cases} \quad (3.11)$$

$$S = \begin{cases} \frac{MAX-MIN}{MIN}, & \text{se } MAX > 0, \\ 0, & \text{se } MAX = 0, \end{cases} \quad (3.12)$$

$$V = MAX. \quad (3.13)$$

Entretanto, as funções exponenciais deformadas utilizadas na divergência generalizada são definidas utilizando a função logarítmica cujo domínio é $x > 0$. A função inversa da exponencial φ é calculada para ambas variáveis, x_i e μ_j , e a derivada da função inversa apenas para x_i . Desta forma, a implementação exige um tratamento dos valores dos componentes do conjunto de entrada. Assim, supondo que se para algum i e j , $p \neq 0$ e $q = 0$, na parcela $\varphi^{-1}(q)$ e utilizando φ_1 tem $\ln(\exp(0) - 1) = -\infty$ e para $\varphi_2^{-1}(q)$ com $\kappa \neq 0$, $\frac{0^\kappa - 0^{-\kappa}}{2^\kappa} = -\infty$. De forma análoga, para $p = 0$ e $q \neq 0$, $\varphi_1^{-1}(p) = -\infty$ e $\varphi_1^{-2}(p) = -\infty$. Já para a parcela $(\varphi^{-1})'(0)$, tem-se $\frac{\exp(0)}{\exp(0)-1} = \infty$ e $\frac{0^\kappa + 0^{-\kappa}}{2} \frac{1}{0} = \infty$. Essa convenção se justifica por continuidade (COVER; THOMAS, 2006). Logo, para o uso das duas funções exponenciais é preciso definir os conjuntos $\mathcal{X} = \{x_i\}_{i=1}^n$ para $x_i \neq 0$ e $\mathcal{M} = \{\mu_j\}_{j=1}^k$ para $\mu_j \neq 0$.

Assim como os tratamentos para divergência KL sugeridos por Rocha (2015) é possível fazer os mesmos para a divergência generalizada. A primeira opção é para $\varphi^{-1}(0) = -\gamma$ ou $(\varphi^{-1})'(0) = \sigma$, onde γ e σ são um valores suficientemente grandes. Outra opção é para

$p = 0$ considerar $p \rightarrow 0$ e assumir $p = \varepsilon$, sendo ε um valor bem próximo de 0; fazer o mesmo para $q = 0$. Uma terceira forma é realizar um deslocamento das variáveis p e q somando um valor $\alpha \in \mathbb{R}$ em coordenada dos vetores evitando que assumam valores iguais a zero, ou seja, $\mathcal{X} = \{x_1 + \alpha, x_2 + \alpha, \dots, x_n + \alpha\}$ e $\mathcal{M} = \{\mu_1 + \alpha, \mu_2 + \alpha, \dots, \mu_k + \alpha\}$.

Definido a quantidade de clusters k e com o conjunto de entrada \mathcal{X} já tratada, é realizada a segmentação por agrupamento dos dados. Primeiramente é inicializado o conjunto de representantes $\mathcal{M} = \{\mu_j\}_{j=1}^k$ com valores aleatórios. A partir da medida de dissimilaridade é realizada a etapa de atribuição, em que é calculada a divergência entre os *pixels* e os representantes dos clusters, e então o *pixel* x_i é atribuído à região \mathcal{X}_j que apresentar menor dissimilaridade $D_\varphi(x_i, \mu_j)$.

Segundo Banerjee *et al.* (2005), os algoritmos de clusterização geralmente assumem uma medida uniforme, ou seja, $v_i = \frac{1}{n}$ para todo i . Da mesma forma, é adotado que todos os *pixels* possuem a mesma medida de probabilidade v . Assim sendo, na etapa de re-estimação os novos representantes do conjunto \mathcal{M} são dados pela média dos *pixels* da região, ou seja, $\mu_j = \frac{1}{\pi_j} \sum_{x_i \in \mathcal{X}_j} x_i$, onde π_j pode ser definido pela quantidade de *pixels* em \mathcal{X}_j .

As etapas de atribuição e re-estimação são repetidas até que a perda de informação de Bregman $L_\varphi(\mathcal{M})$, Equação 2.40, não possa mais ser diminuída nem pela etapa de atribuição e nem alterando as médias dos clusters existentes, ou tenha alcançado um número finito e pré-definido de iterações. Como o cálculo da função objetivo $L_\varphi(\mathcal{M})$ pode ser interpretado como a média das divergências entre os pontos x_i e os representantes μ_j , outra maneira de avaliar a convergência é quando não há mais alteração nos valores dos representantes. É possível observar o pseudocódigo do método de clusterização proposto no Algoritmo 3.

Algoritmo 3: Clusterização baseada na divergência generalizada

Input: Conjunto $\mathcal{X} = \{x_i\}_{i=1}^n \subset S \subseteq \mathbb{R}_+^d$ $x_i \neq 0 \forall i$; medida de probabilidade de $v_i = \frac{1}{n} \forall i$; divergência generalizada D_φ ; função exponencial φ ; o número de clusters k e o número máximo de iterações t_{\max}

Output: Imagem segmentada em regiões $\{\mathcal{X}_j\}_{j=1}^k$ distintas e disjuntas

Método:

Inicializa $\{\mu_j\}_{j=1}^k$ com $\mu_j \in \text{ri}(S)$ com valores aleatórios tal que $\mu_j \neq 0 \forall j$ e o número inicial de iterações $t = 0$

repetir

{Etapa de atribuição}

$\mathcal{X}_j \leftarrow \emptyset, 1 \leq j \leq k$

para $i = 1$ até n **fazer**

$\mathcal{X}_j \leftarrow \mathcal{X}_j \cup \{x_i\}$, onde $j = \text{argmin } D_\varphi(x_i, \mu_j)$

fim do para

{Etapa de re-estimação}

para $j = 1$ até k **fazer**

$\pi_j \leftarrow |\mathcal{X}_j|$

$\mu_j \leftarrow \frac{1}{\pi_j} \sum_{x_i \in \mathcal{X}_j} x_i$

fim do para

$t \leftarrow t + 1$

até $t = t_{\max}$ ou $\min L_\varphi(\mathcal{M})$

$\mathcal{M} \leftarrow \{\mu_j\}_{j=1}^k$

4 EXPERIMENTOS E ANÁLISE DOS RESULTADOS

Neste capítulo são apresentados os experimentos realizados com o intuito de avaliar a eficácia do método proposto de clusterização baseado na divergência generalizada aplicado a segmentação de imagens, conforme especificado no Capítulo 3. Também são analisados e discutidos os resultados obtidos para os diferentes testes e métodos de avaliação, sendo comparados com técnicas de segmentação já estabelecidas na literatura.

Os testes foram divididos em dois grupos: imagens em escala de cinza (experimento I) e imagens coloridas (experimento II). O experimento I é dividido em dois grupos de testes, no grupo 1 as imagens são segmentadas utilizando apenas 2 clusters, já no grupo 2 o número de clusters é definido para cada imagem individualmente. Já no experimento II as imagens são divididas em três grupos. As segmentações são realizadas usando os valores dos pixels no espaço RGB no primeiro grupo e no HSV no segundo grupo. Já no terceiro grupo de teste é efetuado uma normalização dos níveis RGB das imagens. Com exceção do grupo 3 do experimento II, os outros grupos de testes foram segmentados sem haver nenhum pré-processamento.

Os resultados apresentados para ambos os grupo foram analisados conforme dois métodos: avaliação métrica e avaliação qualitativa. Na avaliação métrica, após o resultado da segmentação, são calculadas as medidas de acurácia, precisão e sensibilidade em relação ao GT, assim como o coeficiente *Dice*. Já na avaliação qualitativa são analisados visualmente por 12 pessoas os resultados da segmentação, em que tendo em vista os critérios empíricos do que representa uma boa segmentação é dado uma nota de 0 a 10 para o resultado de cada de segmentação, sendo 0 (muito ruim) e 10 (muito bom). Essa avaliação é realizada através de um sistema que disponibiliza as imagens originais, seus respectivos GTs e os resultados das segmentações com um campo para adicionar as notas.

Considerando que na exponencial deformada φ_2 tem-se $-1 < \kappa < 1$ e na exponencial deformada φ_3 $q \in (0, 1]$ os valores de κ e q são escolhidos empiricamente, visando não somente bons resultados na avaliação métrica como também o quão bem as regiões são segmentadas sem a presença de muitos ruídos. Para todas as imagens do grupo 1 do experimento I é definido $k = 0,2$ e $q = 0,3$. Já para o grupo 2 do experimento I e todos os testes do experimento II, esses valores foram escolhidos individualmente para cada imagem.

4.1 Experimento I

4.1.1 Grupo 1

A execução do método proposto para imagens em escala de cinza é feita utilizando um conjunto de 5 imagens do banco de dados NoisyOffice com amostras dos quatro tipos de ruídos simulados, como descrito na seção 3.1. A clusterização é realizada visando dividir a imagem em duas regiões, texto e fundo da imagem (*background*), resultando em uma imagem binária. Ou seja, a clusterização é realizada definindo o número de clusters $k = 2$. Considerando os GTs fornecidos, o *background* é estabelecido na cor branca (*pixels* igual a 1) e a região segmentada, que corresponde a região do texto, na cor preta (*pixels* igual a 0). Para comparar com o método proposto é realizado a segmentação das imagens também pela técnica *K-means* utilizando a distância euclidiana e Limiar de Otsu.

Realizado as segmentações, é obtido a matriz de confusão para cada amostra comparando a imagem segmentada em relação ao seu GT. Embora o valor dos *pixels* referentes à região texto seja igual a 0, por ser a região principal da segmentação, ela é definida como a classe positiva (1) e o *background* como a classe negativa (0). Dada a matriz de confusão, são calculados as métricas para cada imagem de teste em cada técnica implementada. Os resultados da acurácia, precisão, sensibilidade e do CSD para cada método são mostrados nas Tabelas 3, 4, 5 e 6, respectivamente.

Tabela 3 – Tabela dos resultados da acurácia para cada método no grupo 1 do Experimento I

	Divergência D_{φ_1}	Divergência D_{φ_2}	Divergência D_{φ_3}	K-means	Otsu
Imagem 1	0,9577	0,9475	0,9475	0,7015	0,7068
Imagem 2	0,9651	0,9612	0,9580	0,9378	0,9378
Imagem 3	0,9569	0,9532	0,9511	0,9962	0,9962
Imagem 4	0,9743	0,9690	0,9681	0,9934	0,9934
Imagem 5	0,9791	0,9637	0,9622	0,9950	0,9950

Tabela 4 – Tabela dos resultados da precisão para cada método no grupo 1 do Experimento I

	Divergência D_{φ_1}	Divergência D_{φ_2}	Divergência D_{φ_3}	K-means	Otsu
Imagem 1	0,9738	0,9914	0,9936	0,3088	0,3126
Imagem 2	0,9056	0,9978	0,9997	0,6470	0,6470
Imagem 3	1	1	1	0,9860	0,9844
Imagem 4	0,9985	0,9998	0,9998	0,9705	0,9705
Imagem 5	0,9998	1	1	0,9775	0,9775

Tabela 5 – Tabela dos resultados da sensibilidade para cada método no grupo 1 do Experimento I

	Divergência D_{φ_1}	Divergência D_{φ_2}	Divergência D_{φ_3}	K-means	Otsu
Imagem 1	0,7014	0,6115	0,5877	1	0,9998
Imagem 2	0,7742	0,6612	0,6312	0,9999	1
Imagem 3	0,6326	0,6011	0,5832	0,9812	0,9835
Imagem 4	0,7865	0,7415	0,7339	0,9746	0,9746
Imagem 5	0,8416	0,7243	0,7135	0,9849	0,9849

Tabela 6 – Tabela dos resultados do CSD para cada método no grupo 1 do Experimento I

	Divergência D_{φ_1}	Divergência D_{φ_2}	Divergência D_{φ_3}	K-means	Otsu
Imagem 1	0,8155	0,7564	0,7386	0,4719	0,4763
Imagem 2	0,8347	0,7954	0,7738	0,7857	0,7857
Imagem 3	0,7749	0,7508	0,7366	0,9836	0,9840
Imagem 4	0,8799	0,8515	0,8464	0,9726	0,9726
Imagem 5	0,9139	0,8401	0,8328	0,9812	0,9812

Como observado na Tabela 3, para a maioria das amostras as técnicas apresentaram resultados bem próximos, sendo os maiores valores de acurácia atingidos pelas técnicas K-means e Limiar de Otsu. Contudo, o método proposto apresentou para as exponenciais φ_2 e φ_3 , em todas as imagens, valores acima de 90% de acurácia, mesmo para a amostra que as técnicas K-means e Limiar de Otsu ficam abaixo de 71%. Quando analisados os resultados da Tabela 6 de cada imagem, é possível observar um comportamento similar aos da acurácia. Nota-se que para muitas amostras o K-means e o Limiar de Otsu mostraram valores acima do método proposto, mas para a Imagem 1, ambos os métodos tiveram $CSD < 0,5$, enquanto em todos os testes do método proposto o CSD foi acima de 0,7.

Analisando os resultados na Tabela 4 e na Tabela 5 é possível notar que, em relação as técnica K-means e limiar de Otsu, os testes com a clusterização baseada na divergência apresentaram uma precisão superior e uma sensibilidade inferior. Ou seja, dentre todos os pixels referentes a classe positiva (texto) do GT o K-means e o limiar de Otsu se mostraram melhor que o método proposto para identificá-los, contudo dentre os pixels classificados como pertencentes ao texto o método proposto apresentou uma melhor proporção de acerto do que o K-means.

Os resultados da avaliação qualitativa para cada imagem são dados na Tabela 7. Nota-se que nas imagens 4 e 5, em que as técnicas K-means e limiar de Otsu apresentaram as melhores notas, o método proposto também recebeu notas quase tão boas quanto. Já para as imagens que as técnicas K-means e limiar de Otsu receberam notas médias abaixo de 6, o método proposto ainda apresentou notas boas, acima de 8 para as exponenciais deformadas φ_2 e φ_3 e acima de 7 para a exponencial φ_1 .

Tabela 7 – Tabela da média das notas para cada método no grupo 1 do Experimento I

	Divergência D_{φ_1}	Divergência D_{φ_2}	Divergência D_{φ_3}	K-means	Otsu
Imagem 1	7,5	8,6	8,6	4,3	4,5
Imagem 2	7,1	8,7	8,8	5,1	5,2
Imagem 3	9,1	8,9	9,0	9,7	9,8
Imagem 4	9,4	8,7	9,2	8,5	8,7
Imagem 5	9,5	8,8	9,0	9,7	9,8

Comparando a Tabela 6 com a Tabela 7 observa-se que a Imagem 2 tem um CSD acima de 0,7 para todos os métodos implementados, o que pode ser interpretado como todos os métodos tendo realizado uma boa segmentação. Contudo, quando analisadas as notas, o *K-means* e o limiar de Otsu tiveram notas abaixo de 6, enquanto o método proposto com a exponencial deformada φ_1 tem uma nota acima de 7, e com as exponenciais φ_2 e φ_3 acima de 8. É possível inferir que o resultado de uma segmentação mesmo com um CSD acima de 0,7 pode não ser considerada uma boa segmentação.

Observando os resultados da segmentação da Imagem 2 na Figura 9 e considerando como uma boa segmentação uma que não comprometeria o reconhecimento dos caracteres, conclui-se que o CSD acima de 0,7 para técnica *K-means* e Limiar de Otsu se dá pelo alto número de VP, e as notas baixas por grande parte do ruído ter sido segmentado como parte da região do texto. Já utilizando a exponencial deformada φ_1 parte do ruído é segmentado como texto, mas uma área consideravelmente menor com relação ao *K-means* e o Limiar de Otsu. Enquanto com as exponenciais φ_2 e φ_3 o ruído não é segmentado com o texto e parte dos *pixels* do texto são segmentados como fundo da imagem, o que explica a sensibilidade inferior do método proposto em comparação às outras técnicas, não sendo suficiente para comprometer a representação dos caracteres, o que explica as notas altas.

Com base nos resultados do *K-means* clássico e do Limiar de Otsu para todas as métricas é possível considerar que a Imagem 1 apresenta aspectos visuais que a tornam mais desafiadora em relação às outras imagens. Na Figura 20 observa-se que essa amostra possui um ruído consideravelmente grande e que essas técnicas tiveram dificuldade em diferenciá-lo do texto da imagem, sendo grande parte desse ruído segmentado como a classe positiva.

Considerando a métrica amplamente utilizada para avaliar o desempenho de segmentação, CSD, e as notas atribuídas por avaliação empírica, o método proposto apresentou melhores resultados do que as outras técnicas para as imagens 1 e 2; já o *K-means* e Limiar de Otsu tiveram melhores resultados para as imagens 3 e 5. Na Figura 11, são apresentados os histogramas das Imagens 2 e 5 e os histogramas dos pixels referentes a classe positiva. Observa-se que para

Figura 9 – Resultado da segmentação da Imagem 2 do grupo 1 do Experimento I

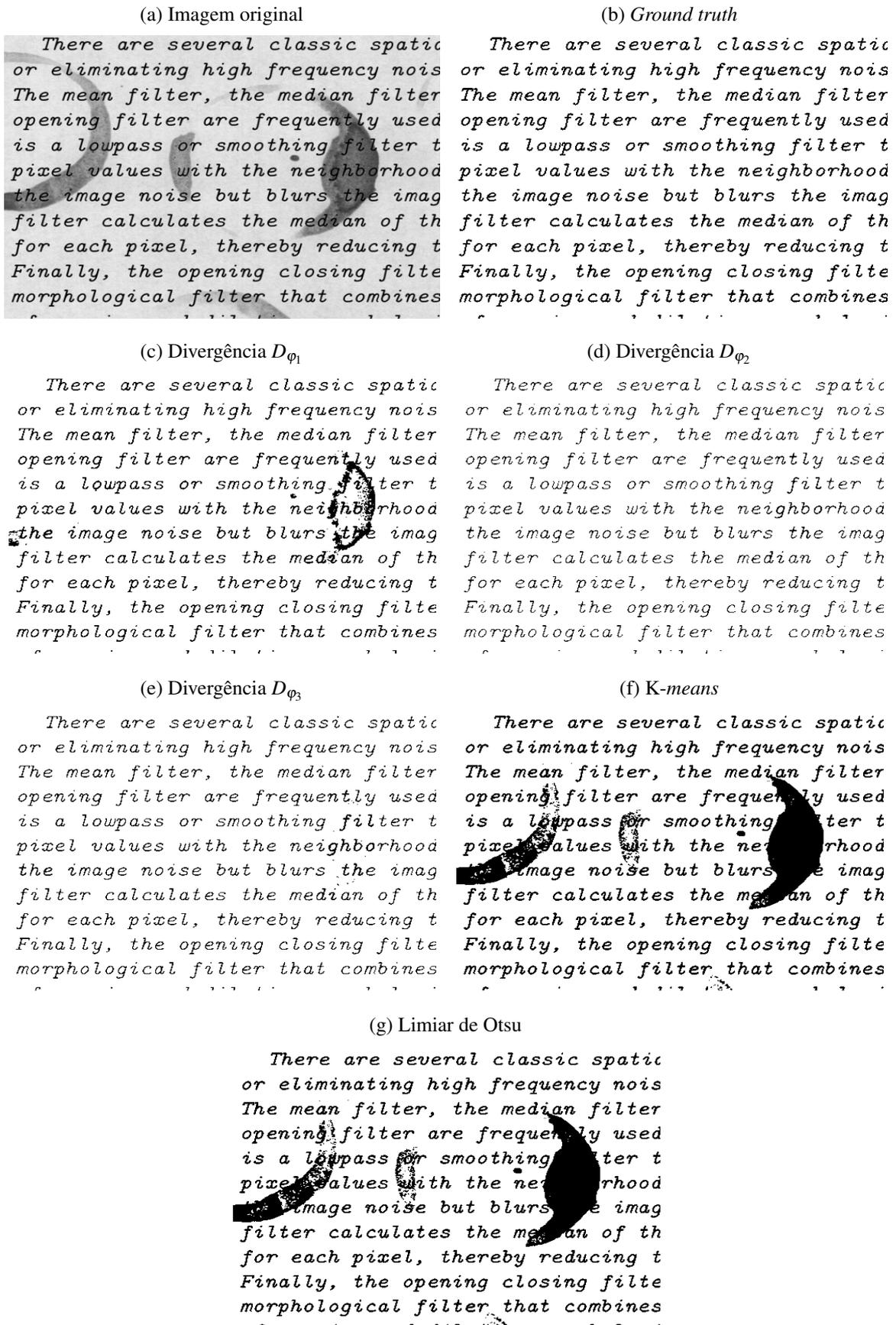
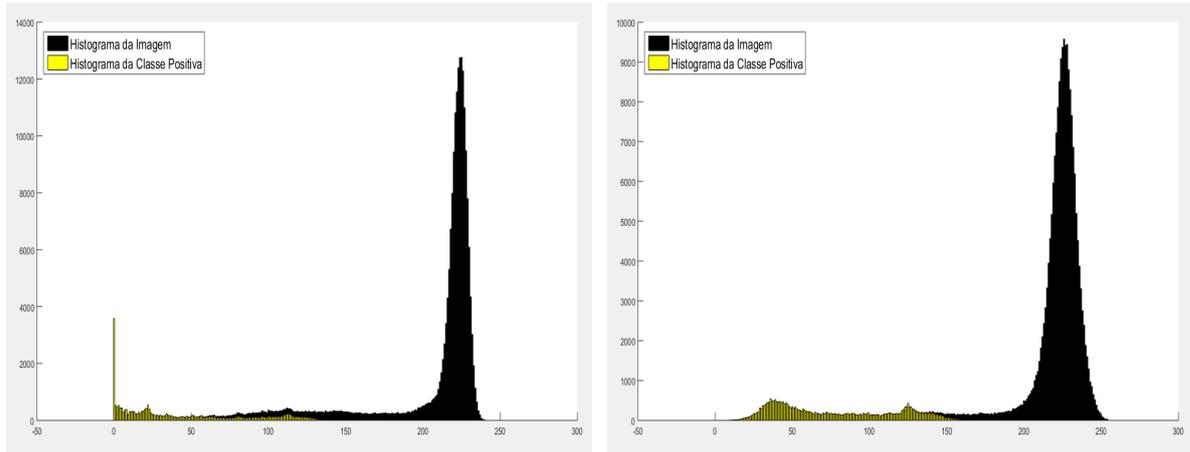


Figura 10 – Resultado da segmentação da Imagem 1 do grupo 1 do Experimento I



ambas as imagens o vale é plano e amplo e os picos possuem alturas bem desiguais. Contudo, na Imagem 2 entre as intensidades 70 e 130, aproximadamente, há *pixels* de mesma intensidade para as duas classes, o que dificulta ainda mais determinar o valor do limiar mais adequado.

Figura 11 – Histograma da imagem e sua classe positiva
(a) Histograma da Imagem 2 (b) Histograma da Imagem 5



4.1.2 Grupo 2

O segundo grupo de testes para imagens em escala cinza é realizado com um conjunto de 5 imagens em que o algoritmo de clusterização é operado com mais de 2 clusters. Diferente do grupo 1, os testes não são realizados visando uma aplicação específica e as imagens são de conjuntos de dados variados. Logo cada imagem é de uma aplicação diferente. Desta forma, o número de clusters é definido individualmente considerando empiricamente o número de regiões contidas na imagem. A segmentação é realizada objetivando selecionar uma região específica da imagem. Portanto, a matriz de confusão é obtida considerando a região de interesse como classe positiva e as outras regiões como classe negativa. Os resultados do método proposto são comparados com o resultado da segmentação pela técnica *K-means* utilizando a distância euclidiana. Nas Tabelas 8, 9, 10 e 11 são apresentados os valores das métricas acurácia, precisão, sensibilidade e CSD para cada método e cada imagem, respectivamente.

A partir dos resultados apresentados é possível observar nas Tabelas 9 e 10 que a precisão e a sensibilidade mostraram um comportamento similar ao do grupo 1 para a maioria das imagens, tendo o método proposto melhores resultados para precisão e inferiores na sensibilidade que o *K-means*. Já na Tabela 8 nota-se que em todas as imagens o uso da clusterização baseada na divergência generalizada resultou em maiores valores em pelo menos uma das exponenciais

Tabela 8 – Tabela dos resultados da acurácia para cada método no grupo 2 do Experimento I

	Divergência D_{φ_1}	Divergência D_{φ_2}	Divergência D_{φ_3}	K-means
Imagem 1	0,9462	0,9462	0,8995	0,8975
Imagem 2	0,9697	0,9669	0,9960	0,9949
Imagem 3	0,9928	0,9922	0,9930	0,9209
Imagem 4	0,9825	0,9775	0,9710	0,9574
Imagem 5	0,9592	0,9571	0,9448	0,9448

Tabela 9 – Tabela dos resultados da precisão para cada método no grupo 2 do Experimento I

	Divergência D_{φ_1}	Divergência D_{φ_2}	Divergência D_{φ_3}	K-means
Imagem 1	0,4811	0,4791	0,2876	0,2836
Imagem 2	0,9971	0,9975	0,9878	0,9843
Imagem 3	0,9716	0,9759	0,9661	0,2994
Imagem 4	0,3988	0,3396	0,2850	0,2132
Imagem 5	0,8881	0,8557	0,9592	0,9592

Tabela 10 – Tabela dos resultados da sensibilidade para cada método no grupo 2 do Experimento I

	Divergência D_{φ_1}	Divergência D_{φ_2}	Divergência D_{φ_3}	K-means
Imagem 1	0,3295	0,3008	0,6207	0,6261
Imagem 2	0,7742	0,6612	0,6312	0,9995
Imagem 3	0,8155	0,7958	0,8283	0,9530
Imagem 4	0,9611	0,9668	0,9713	0,9765
Imagem 5	0,8761	0,9063	0,7127	0,7127

Tabela 11 – Tabela dos resultados do CSD para cada método no grupo 2 do Experimento I

	Divergência D_{φ_1}	Divergência D_{φ_2}	Divergência D_{φ_3}	K-means
Imagem 1	0,3911	0,3695	0,3930	0,3904
Imagem 2	0,9492	0,9442	0,9936	0,9919
Imagem 3	0,8867	0,8767	0,8919	0,4556
Imagem 4	0,5637	0,5026	0,4407	0,3499
Imagem 5	0,8821	0,8803	0,8178	0,8178

deformadas utilizadas.

As médias das notas atribuídas pela avaliação qualitativa são apresentadas na Tabela 12. Comparando esses resultados com o da Tabela 11 observa-se que o método proposto resultou em melhores valores do que com o K-means para ambos os métodos de avaliação das segmentações, com exceção da Imagem 5, que a técnica K-means teve uma nota média igual à exponencial φ_3 e superior às outras duas. Na Figura 12 é possível observar os histogramas das Imagens 3 e 5 e da classe positiva em que para ambas a segmentação foi realizada definindo o número de regiões igual a 3. Nota-se no histograma da Imagem 3 que, como na Imagem 2 do grupo 1, o método proposto se apresenta melhor do que a técnica K-means para segmentar regiões em que maior parte dos pixels da classe positiva possui intensidade abaixo de 50.

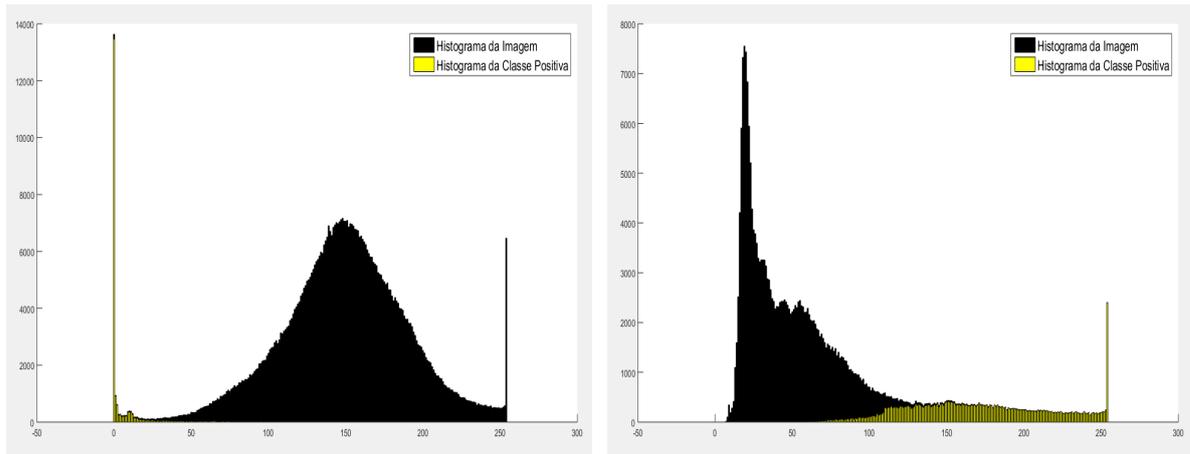
Tabela 12 – Tabela da média das notas para cada método no grupo 2 do Experimento I

	Divergência D_{φ_1}	Divergência D_{φ_2}	Divergência D_{φ_3}	K-means
Imagem 1	6,3	6,1	6,1	5,7
Imagem 2	9,0	9,1	8,8	8,4
Imagem 3	7,9	7,8	7,6	6,0
Imagem 4	7,8	7,6	7,8	7,3
Imagem 5	8,1	7,7	8,5	8,5

Figura 12 – Histograma da imagem e sua classe positiva

(a) Histograma da Imagem 3

(b) Histograma da Imagem 5



4.2 Experimento II

O segundo experimento foi realizado utilizando 15 imagens coloridas de aplicações gerais, sendo elas divididas em três grupos, 5 imagens para os testes da segmentação no espaço de cor RGB, 5 para os testes em HSV e 5 para os níveis RGB realçados. Também é aplicada a técnica *K-means* utilizando a distância euclidiana para comparar os resultados obtidos pelo método proposto. Diante das particularidades de cada imagem, o número de clusters é definido empiricamente para cada amostra considerando quantas regiões há na imagem ou o necessário para separar uma região específica, conforme descrito na Seção 3.1.

4.2.1 Grupo 1

No primeiro grupo a clusterização é realizada com os níveis de cor RGB sem nenhum pré-processamento como atributos de cada *pixel*. As Tabelas 13, 14, 15 e 16 apresentam os resultados da acurácia, precisão, sensibilidade e CSD, respectivamente, obtidos para cada método implementado. Para imagens em que a segmentação é testada agrupando a imagem em mais de uma região específica, dividindo a imagem em mais do que duas regiões, as métricas são dadas

pela média dos resultados de cada região.

Tabela 13 – Tabela dos resultados da acurácia para cada método no grupo 1 do Experimento II

	Divergência D_{φ_1}	Divergência D_{φ_2}	Divergência D_{φ_3}	K-means
Imagem 1	0,9580	0,9620	0,950	0,950
Imagem 2	0,9620	0,960	0,9610	0,9620
Imagem 3	0,9720	0,970	0,9760	0,9770
Imagem 4	0,960	0,9520	0,9570	0,9650
Imagem 5	0,910	0,8960	0,9590	0,960

Tabela 14 – Tabela dos resultados da precisão para cada método no grupo 1 do Experimento II

	Divergência D_{φ_1}	Divergência D_{φ_2}	Divergência D_{φ_3}	K-means
Imagem 1	0,7150	0,7375	0,6785	0,6790
Imagem 2	0,930	0,9330	0,8735	0,8805
Imagem 3	0,9690	0,9670	0,9480	0,9480
Imagem 4	0,9695	0,9715	0,9455	0,9560
Imagem 5	0,9320	0,9270	0,960	0,9605

Tabela 15 – Tabela dos resultados da sensibilidade para cada método no grupo 1 do Experimento II

	Divergência D_{φ_1}	Divergência D_{φ_2}	Divergência D_{φ_3}	K-means
Imagem 1	0,8795	0,8860	0,8725	0,8755
Imagem 2	0,8970	0,8860	0,9195	0,9210
Imagem 3	0,7240	0,7070	0,7910	0,7950
Imagem 4	0,9465	0,9340	0,9475	0,9590
Imagem 5	0,8925	0,8765	0,9640	0,9660

Tabela 16 – Tabela dos resultados do CSD para cada método no grupo 1 do Experimento II

	Divergência D_{φ_1}	Divergência D_{φ_2}	Divergência D_{φ_3}	K-means
Imagem 1	0,8627	0,8752	0,8423	0,8447
Imagem 2	0,9261	0,9220	0,9168	0,920
Imagem 3	0,8288	0,8168	0,8624	0,8648
Imagem 4	0,9602	0,9519	0,9582	0,9651
Imagem 5	0,9116	0,8978	0,9580	0,9601

Com base nos resultados da acurácia na Tabela 13 e do CSD na Tabela 16 observa-se que não há uma diferença considerável entre os testes com a clusterização baseada na divergência generalizada e o *K-means*, com exceção da imagem 5 em que o *K-means* apresentou um resultado, para ambas as métricas, acima de 5% a mais que o do método proposto para as exponenciais φ_1 e φ_2 .

Quanto a precisão e sensibilidade, observando as Tabelas 14 e 15 percebe-se um comportamento similar ao do experimento 1 para a maioria das imagens, tendo o método proposto resultados levemente superiores para precisão e levemente inferiores para sensibilidade em relação ao *K-means*. Assim como na acurácia e CSD, o *K-means* resultou numa precisão e sensibilidade maior para a Imagem 5 que o método proposto para as exponenciais deformadas φ_1 e φ_2 .

Na Tabela 17 são apresentadas as médias das notas obtidas a partir da avaliação qualitativa. Considerando esses resultados e o do CSD é possível observar que o método proposto obteve uma boa segmentação para todas as imagens, notas acima de 7 e CSD acima de 0,7. Em comparação com a técnica *K-means*, para ambos os métodos de avaliação, o método proposto mostrou resultados tão bons quanto a técnica *K-means*, sendo a Imagem 5 a que o método apresentou o menores valores em relação ao *K-means*. Observa-se na Figura 13 que no melhor resultado do método proposto em comparação com o *K-means* a maioria dos *pixels* possuem intensidade acima de 200 para os três níveis RGB, já no pior resultado os níveis G e B estão abaixo de 100 para a maioria dos pixels.

Tabela 17 – Tabela da média das notas para cada método no grupo 1 do Experimento II

	Divergência D_{φ_1}	Divergência D_{φ_2}	Divergência D_{φ_3}	K-means
Imagem 1	8,2	8,4	8,4	8,4
Imagem 2	8,3	8,2	7,8	7,8
Imagem 3	8,0	8,0	8,2	8,1
Imagem 4	7,9	7,8	8,4	8,6
Imagem 5	7,4	7,3	8,6	8,8

4.2.2 Grupo 2

Os testes realizados no segundo grupo de imagens coloridas utilizam os pixels da imagem no espaço HSV como conjunto de entrada dos algoritmos de clusterização. Os resultados das métricas acurácia, precisão, sensibilidade e CSD são apresentados nas Tabelas 18, 19, 20 e 21, respectivamente.

Dados os resultados da acurácia, precisão e sensibilidade é possível observar que o método proposto utilizando as exponenciais deformadas φ_3 e a técnica *K-means* apresentaram valores próximos em cada métrica para a maioria das amostras. Observa-se também que o método proposto, em pelo menos uma das exponenciais, apresentou melhores resultados para todas as imagens na acurácia e em quase todas as imagens na precisão e na sensibilidade.

Figura 13 – Histograma dos níveis RGB das imagens

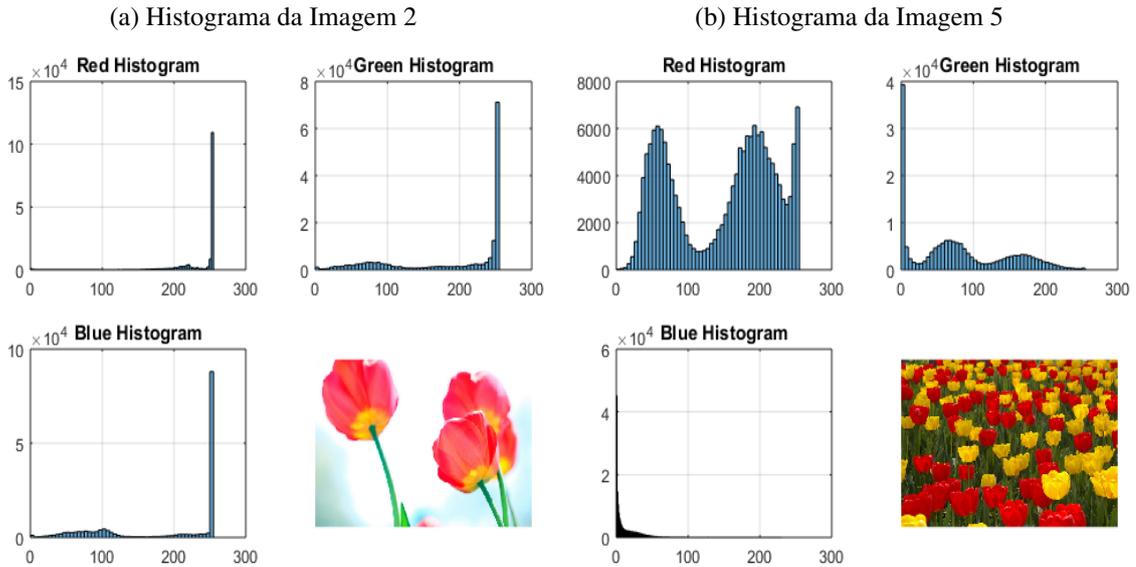


Tabela 18 – Tabela dos resultados da acurácia para cada método no grupo 2 do Experimento II

	Divergência D_{φ_1}	Divergência D_{φ_2}	Divergência D_{φ_3}	K-means
Imagem 1	0,9550	0,9760	0,9310	0,9210
Imagem 2	0,9630	0,9610	0,7850	0,770
Imagem 3	0,9780	0,9780	0,9620	0,9680
Imagem 4	0,9760	0,9750	0,9750	0,9680
Imagem 5	0,9530	0,9540	0,9190	0,9150

Tabela 19 – Tabela dos resultados da precisão para cada método no grupo 2 do Experimento II

	Divergência D_{φ_1}	Divergência D_{φ_2}	Divergência D_{φ_3}	K-means
Imagem 1	0,8220	0,9010	0,7670	0,7470
Imagem 2	0,9970	0,9980	0,7540	0,7290
Imagem 3	0,9470	0,940	0,9620	0,9620
Imagem 4	0,9780	0,9770	0,9760	0,9450
Imagem 5	0,8240	0,8260	0,7290	0,720

Tabela 20 – Tabela dos resultados da sensibilidade para cada método no grupo 2 do Experimento II

	Divergência D_{φ_1}	Divergência D_{φ_2}	Divergência D_{φ_3}	K-means
Imagem 1	0,9880	0,9850	0,9340	0,9080
Imagem 2	0,8970	0,8890	0,5680	0,5430
Imagem 3	0,9840	0,9880	0,910	0,930
Imagem 4	0,9620	0,9620	0,9630	0,9780
Imagem 5	0,9980	0,9980	0,9950	0,9950

Considerando os resultados do CSD na Tabela 21 é possível considerar que o método proposto e a técnica *K-means* obtiveram bons resultados de segmentação para a maioria das imagens. Com exceção da Imagem 2, em que a exponencial deformada φ_3 e o *K-means* apresentaram $CSD < 0,65$, ambas as metodologias apresentaram resultados acima de 0,8. Nota-

Tabela 21 – Tabela dos resultados do CSD para cada método no grupo 2 do Experimento II

	Divergência D_{φ_1}	Divergência D_{φ_2}	Divergência D_{φ_3}	K-means
Imagem 1	0,8974	0,9411	0,8423	0,8197
Imagem 2	0,9444	0,9404	0,6479	0,6224
Imagem 3	0,9652	0,9634	0,9353	0,9650
Imagem 4	0,9699	0,9694	0,9694	0,9612
Imagem 5	0,9027	0,9039	0,8415	0,8355

se também que a clusterização utilizando a exponencial φ_2 apresentou CSD acima de 0,9 para todas as imagens.

Os resultados obtidos pela avaliação qualitativa são apresentados na Tabela 22. Diante desses resultados e do CSD é possível concluir que o método proposto utilizando as exponenciais deformadas φ_1 e φ_2 apresentaram melhores segmentações que o K-means para todas as imagens testadas no espaço HSV. Na Figura 14 é apresentado os histogramas do espaço HSV das imagens 2 e 3, em que há a maior e a menor diferença do método proposto em relação ao K-means, respectivamente. Na Imagem 3 os *pixels* possuem níveis mais distribuídos entre a escala, enquanto na Imagem 2 os níveis estão mais concentrados em valores próximos.

Tabela 22 – Tabela da média das notas para cada método no grupo 2 do Experimento II

	Divergência D_{φ_1}	Divergência D_{φ_2}	Divergência D_{φ_3}	K-means
Imagem 1	7,6	7,9	6,8	6,7
Imagem 2	8,7	8,7	6,3	6,3
Imagem 3	8,1	8,1	7,8	8,0
Imagem 4	8,7	8,7	8,3	7,8
Imagem 5	7,6	7,5	6,4	6,3

4.2.3 Grupo 3

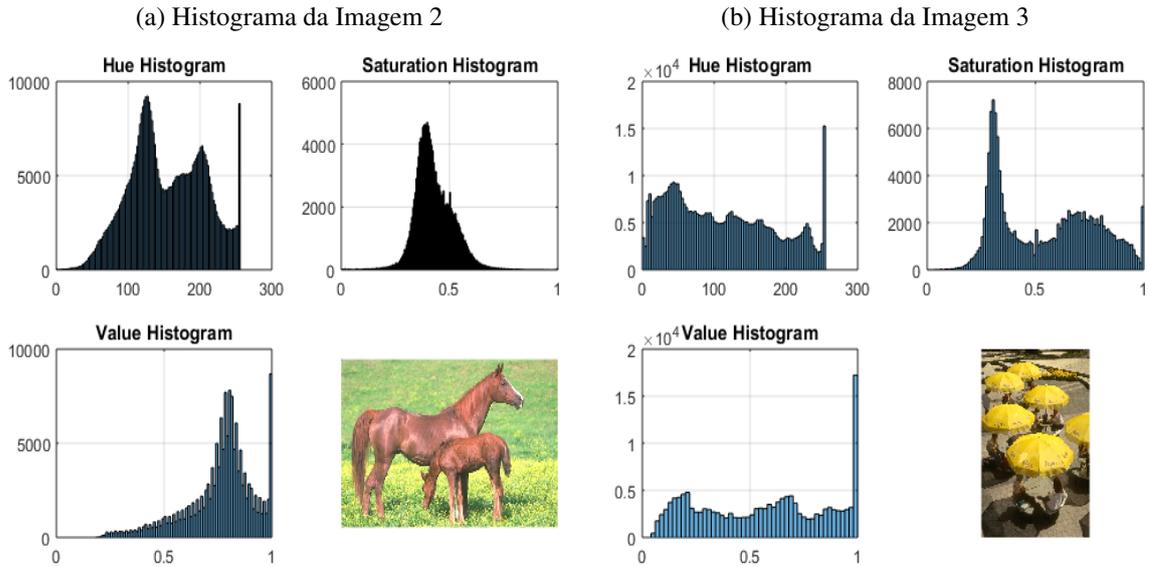
No terceiro grupo de imagens coloridas os testes são realizados após um pré-processamento de normalização dos níveis RGB, sendo os novos valores obtidos pelas seguintes equações:

$$r = \frac{R}{(R + G + B)}, \quad (4.1)$$

$$g = \frac{G}{(R + G + B)}, \quad (4.2)$$

$$b = \frac{B}{(R + G + B)}, \quad (4.3)$$

Figura 14 – Histograma dos níveis HSV das imagens



para $R + G + B > 0$, sendo R, G, B os valores dos níveis RGB na escala $[0,1]$. Os resultados obtidos para acurácia, precisão, sensibilidade e CSD das segmentações de cada imagem são apresentados nas Tabelas 23, 24, 25 e 26, respectivamente.

Tabela 23 – Tabela dos resultados da acurácia para cada método no grupo 3 do Experimento II

	Divergência D_{φ_1}	Divergência D_{φ_2}	Divergência D_{φ_3}	K-means
Imagem 1	0,9850	0,9850	0,9860	0,9860
Imagem 2	0,9680	0,9680	0,9680	0,9680
Imagem 3	0,9740	0,9750	0,9730	0,7120
Imagem 4	0,9320	0,9280	0,9420	0,9360
Imagem 5	0,9910	0,990	0,9930	0,9930

Tabela 24 – Tabela dos resultados da precisão para cada método no grupo 3 do Experimento II

	Divergência D_{φ_1}	Divergência D_{φ_2}	Divergência D_{φ_3}	K-means
Imagem 1	0,9520	0,9530	0,950	0,9490
Imagem 2	0,9915	0,9915	0,9915	0,9915
Imagem 3	0,5910	0,6040	0,5810	0,1040
Imagem 4	0,9850	0,9850	0,9890	0,9920
Imagem 5	0,9880	0,9890	0,9860	0,9870

Diante dos valores expostos por cada métrica é possível observar que, exceto para a Imagem 3, o método proposto e o *K-means* mostraram resultados bem próximos para o conjunto de imagens com pré-processamento. Comparando com as médias das notas, apresentadas na Tabela 27, é possível notar o mesmo comportamento. Na Figura 15 são observados os histogramas das Imagens 3 e 5.

Tabela 25 – Tabela dos resultados da sensibilidade para cada método no grupo 3 do Experimento II

	Divergência D_{φ_1}	Divergência D_{φ_2}	Divergência D_{φ_3}	K-means
Imagem 1	0,9380	0,9360	0,9450	0,9470
Imagem 2	0,9240	0,9240	0,9255	0,9255
Imagem 3	0,9260	0,9230	0,9250	0,9190
Imagem 4	0,7310	0,7170	0,7710	0,7420
Imagem 5	0,8640	0,8550	0,8930	0,8940

Tabela 26 – Tabela dos resultados do CSD para cada método no grupo 3 do Experimento II

	Divergência D_{φ_1}	Divergência D_{φ_2}	Divergência D_{φ_3}	K-means
Imagem 1	0,9450	0,9444	0,9475	0,9480
Imagem 2	0,9564	0,9564	0,9572	0,9572
Imagem 3	0,7215	0,7302	0,7137	0,1868
Imagem 4	0,8392	0,8299	0,8665	0,8490
Imagem 5	0,9218	0,9171	0,9372	0,9382

Tabela 27 – Tabela da média das notas para cada método no grupo 3 do Experimento II

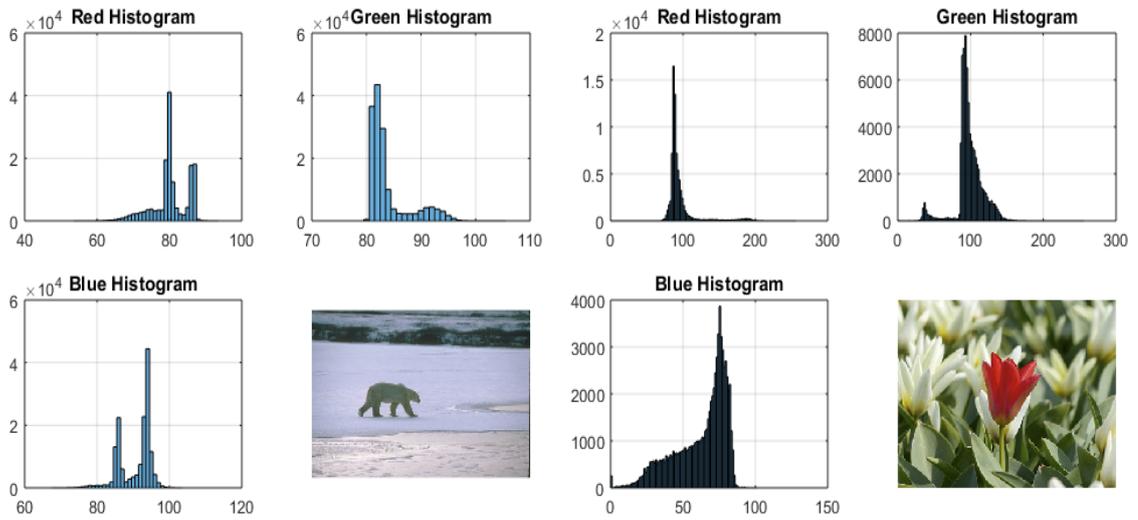
	Divergência D_{φ_1}	Divergência D_{φ_2}	Divergência D_{φ_3}	K-means
Imagem 1	8,0	8,1	8,1	8,0
Imagem 2	8,6	8,6	8,6	8,6
Imagem 3	7,9	7,8	7,9	5,7
Imagem 4	7,9	7,9	8,0	8,0
Imagem 5	8,1	8,4	8,4	8,5

4.3 Conclusão dos Resultados dos Experimentos

Com base nos resultados das métricas discutidos para cada experimento é possível observar que a clusterização baseada na divergência generalizada obteve um bom desempenho para todos os grupos de testes apresentando acurácia acima de 0,9 para todas as imagens testadas utilizando a exponencial φ_1 , para 24 das 25 imagens com a exponencial φ_2 e em 23 com a exponencial φ_3 . O método proposto também apresentou CSD > 0,7 para 23 das 25 imagens testadas utilizando as exponenciais φ_1 e φ_2 e para 22 com a exponencial φ_3 . Enquanto o K-means alcançou acurácia acima de 0,9 para 21 das 25 imagens e CSD > 0,7 para 19 das 25. Quanto à avaliação qualitativa, o método proposto obteve nota média igual ou acima de 7 para 24 imagens para as exponenciais φ_1 e φ_2 e para 21 imagens com a exponencial φ_3 , já o K-means para 17 imagens.

Diante da performance dos testes realizados com as imagens em escala de cinza e coloridas é possível concluir que a clusterização baseada na divergência generalizada se apresentou eficiente e promissora para realizar segmentação de imagens obtendo resultados tão

Figura 15 – Histograma dos níveis RGB das imagens depois do pré-processamento
 (a) Histograma da Imagem 3 (b) Histograma da Imagem 5



bons quanto ao *K-means* e limiar de Otsu ou até mais eficaz para algumas imagens.

5 CONCLUSÕES E TRABALHOS FUTUROS

As considerações finais e contribuições sobre o trabalho desenvolvido são abordadas neste capítulo. São apresentadas também algumas sugestões de trabalhos futuros.

5.1 Considerações Finais

Neste trabalho é realizado um estudo sobre a entropia relativa generalizada e sobre o uso dessa divergência generalizada como medida de dissimilaridade no algoritmo de clusterização aplicado a segmentação de imagens. O objetivo desse trabalho foi desenvolver um caminho alternativo para segmentações de imagens utilizando a divergência obtida a partir da generalização das entropias relativas de Shannon e Tsallis e o algoritmo de clusterização *hard* de Bregman como método de agrupamento de dados. Para tanto, primeiramente foram realizadas pesquisas sobre o uso de divergências como medida de dissimilaridade em algoritmos de clusterização e aplicados a segmentação de imagens. Em seguida foram selecionadas imagens para realizar os testes, tanto para imagens em escala de cinza quanto coloridas. Como para a aplicação da divergência generalizada necessita-se especificar a função exponencial deformada φ , o método proposto foi implementado comparando três funções exponenciais deformadas diferentes. Para avaliar a metodologia foram utilizadas métricas pertinentes ao contexto de segmentação de imagens, como acurácia, precisão, sensibilidade e o Coeficiente de Dice (CSD), sendo comparadas com outras técnicas de segmentação já estabelecidas. Também avaliações visuais foram realizadas e notas foram atribuídas para as segmentações. Foram observadas, em pelo menos duas das três exponenciais deformadas, CSD acima de 0,7 para 23 das 25 imagens e nota média acima 7 para 24 das 25 imagens. Portanto, diante dos resultados obtidos é possível concluir que o método proposto se apresentou eficaz para o cumprimento dos objetivos estabelecidos.

A execução do método proposto teve um custo computacional mais alto que a técnica *K-means* e o Limiar de Otsu. Pois a clusterização baseada na divergência generalizada utiliza mais operações do que o *K-means* clássico e muitas bibliotecas já possuem função pronta para técnicas de clusterização conhecidas. Contudo, o método proposto pode apresentar vantagens que compensem o aumento do custo computacional. Pode-se destacar como principal contribuição deste trabalho a aplicação eficaz de uma distância não métrica como medida de dissimilaridade que permite ao algoritmo de clusterização uma maior flexibilidade de implementação. Pois, dependendo da função exponencial deformada para a divergência generalizada, é possível haver

o acréscimo de parâmetros e com a alteração desses parâmetros obter resultados diferentes para um mesmo número de cluster.

5.2 Trabalhos Futuros

Em vista do desempenho obtido, o método proposto se apresentou promissor em aplicações envolvendo segmentação de imagens. Portanto, propõe-se implementar a metodologia para imagens em aplicações específicas. Planeja-se também realizar testes com outras funções exponenciais deformadas, assim como utilizar a divergência generalizada em outros algoritmos de clusterização. Aplicar a divergência generalizada em outros problemas envolvendo análise de imagem, como na recuperação de conteúdo.

REFERÊNCIAS

- AFROZ, A. S. An alternative approach of evaluating dice similarity index for pre-segmented blood vessels. The 8th International Conference on Software, Knowledge, Information Management and Applications (SKIMA 2014), p. 6, 2014.
- ARBELAEZ, P.; MAIRE, M.; FOWLKES, C.; MALIK, J. Contour detection and hierarchical image segmentation. **IEEE Trans. Pattern Anal. Mach. Intell.**, IEEE Computer Society, Washington, DC, USA, v. 33, n. 5, p. 898–916, 2011. ISSN 0162-8828. Disponível em: <<http://dx.doi.org/10.1109/TPAMI.2010.161>>.
- BANERJEE, A.; MERUGU, S.; DHILLON, I. S.; GHOSH, J. Clustering with bregman divergences. **J. Mach. Learn. Res.**, JMLR.org, v. 6, p. 1705–1749, dez. 2005. ISSN 1532-4435.
- BARRETO, J. B.; OLIVEIRA, D. A.; VIGELIS, R. F.; PAULA JÚNIOR, I. C. Identificação de lesões periapicais em exames de radiografia utilizando clusterização baseada na φ -divergência e algoritmos de classificação. XXXVIII Simpósio Brasileiro de Telecomunicações e Processamento de Sinais - SBrT2020, Florianópolis, p. 5, 2020.
- BARRETO, J. B.; VIGELIS, R. F. Clusterização baseada na φ -divergência aplicada à segmentação de imagens. XXXVII Simpósio Brasileiro de Telecomunicações e Processamento de Sinais - SBrT2019, Petrópolis, p. 5, 2019.
- BORLAND, L.; PLASTINO, A. R.; TSALLIS, C. Information gain within nonextensive thermostatics. *Journal of Mathematical Physics*, v. 39, n. 12, p. 6490–6501, 1998.
- BREGMAN, L. The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. **USSR Computational Mathematics and Mathematical Physics**, v. 7, n. 3, p. 200 – 217, 1967. ISSN 0041-5553.
- BURGER, M. Bregman distances in inverse problems and partial differential equations. In: _____. [S.l.: s.n.], 2016. p. 3–33. ISBN 978-3-319-30784-8.
- BURNEY, S. M. A.; TARIQ, H. K-means cluster analysis for image segmentation. *International Journal of Computer Applications*, v. 96, n. 4, p. 1–8, 2014.
- BUZO, A.; JR, A.; GRAY, R.; MARKEL, J. Speech coding based upon vector quantization. **Acoustics, Speech and Signal Processing, IEEE Transactions on**, v. 28, p. 562 – 574, 11 1980.
- CAYTON, L. Fast nearest neighbor retrieval for bregman divergences. In: **Proceedings of the 25th International Conference on Machine Learning**. New York, NY, USA: Association for Computing Machinery, 2008. (ICML '08), p. 112–119. ISBN 9781605582054. Disponível em: <<https://doi.org/10.1145/1390156.1390171>>.
- CHABRIER, S.; EMILE, B.; ROSENBERGER, C.; LAURENT, H. Unsupervised performance evaluation of image segmentation. *EURASIP Journal on Applied Signal Processing*, p. 1–12, 2006.
- CHABRIER, S.; LAURENT, H.; EMILE, B.; ROSENBURGER, C.; MARCHE, P. A comparative study of supervised evaluation criteria for image segmentation. *EUSIPCO*, p. 1143–1146, 2004.

CHEN, Y.; MING, D.; ZHAO, L.; LV, B.; ZHOU, K.; QING, Y. Review on high spatial resolution remote sensing image segmentation evaluation. **Photogrammetric Engineering & Remote Sensing**, v. 84, p. 629 – 646, 2018.

COVER, T. M.; THOMAS, J. A. **Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)**. USA: Wiley-Interscience, 2006. ISBN 0471241954.

CSISZÁR, I.; SHIELDS, P. C. Information theory and statistics: A tutorial. **Commun. Inf. Theory**, Now Publishers Inc., Hanover, MA, USA, v. 1, n. 4, p. 417–528, dez. 2004. Disponível em: <<https://doi.org/10.1561/01000000004>>.

DE, S.; BHATTACHARYYA, S.; CHAKRABORTY, S.; DUTTA, P. **Hybrid Soft Computing for Multilevel Image and Data Segmentation**. [S. l.]: Springer International Publishing, 2016. (Computational Intelligence Methods and Applications). ISBN 978-3-319-47524-0.

DHILLON, I. S.; MALLELA, S.; KUMAR, R. A divisive information theoretic feature clustering algorithm for text classification. **J. Mach. Learn. Res.**, JMLR.org, v. 3, n. null, p. 1265–1287, mar. 2003. ISSN 1532-4435.

DICE, L. R. Measures of the amount of ecologic association between species. In: **Ecology**. [S.l.]: Ecological Society of America, 1945. v. 26, n. 3, p. 297–302.

DUNN, C. E.; WANG, P. S. P. Character segmentation techniques for handwritten text—a survey. In: **Proceedings., 11th IAPR International Conference on Pattern Recognition. Vol.II. Conference B: Pattern Recognition Methodology and Systems**. [S.l.: s.n.], 1992. p. 577–580.

ENNEKHLI, M.; OUAZZANI, R. E.; HAZITI, M. E. A new content-based image retrieval system using parzen relevance feedback and kullbackleibler divergence. **International Journal of Computer Information Systems and Industrial Management Applications**, v. 10, p. 18–27, 2018. ISSN 2150-7988.

ESQUEF, I. A.; ALBUQUERQUE, M. P.; ALBUQUERQUE, M. P. Processamento digital de imagens. Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, p. 12, 2003.

FERREIRA, D. P. L. **Divergências de Bregman e Total Bregman aplicadas na análise de imagens**. Dissertação (Mestrado) — Universidade Federal de Uberlândia, Uberlândia, 2017.

FU, K. S.; MUI, J. K. A survey on image segmentation. *Pattern Recognition*, New York: Springer Verlag, v. 13, p. 3–16, 1981.

GARCIA-LAMONT, F.; CERVANTES, J.; LÓPEZ, A.; RODRIGUEZ, L. Segmentation of images by color features: A survey. *Neurocomputing*, v. 292, p. 1–27, 2018.

GAURI, A. A.; SHARMA, I. Segmenting coloured images using bregman divergence. In: **2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)**. [S.l.: s.n.], 2018. p. 613–617.

GOLDBERGER, J.; ROWEIS, S.; HINTON, G.; SALAKHUTDINOV, R. Neighbourhood components analysis. In: **Proceedings of the 17th International Conference on Neural Information Processing Systems**. Cambridge, MA, USA: MIT Press, 2004. (NIPS'04), p. 513–520.

GOMALAVALLI, R.; MUTTAN, S.; VENKATASAI, P. M. Determination of dice coefficient of diseased renal images. *IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)*, v. 11, n. 1, p. 37–44, 2016.

GONZALEZ, R. C.; WOODS, R. E. **Processamento Digital de Imagens**. 3. ed. [S. l.]: Prentice Hall, 2011. 624 p.

GUHA, S.; RASTOGI, R.; SHIM, K. Rock: A robust clustering algorithm for categorical attributes. In: **Proceedings of the 15th International Conference on Data Engineering**. USA: IEEE Computer Society, 1999. (ICDE '99), p. 512. ISBN 0769500714.

GUO, X.; HONG, J.; YANG, N. Ambiguity set and learning via bregman and wasserstein. *ArXiv*, abs/1705.08056, p. 14, 05 2017.

GUPTA, M. D.; SRINIVASA, S.; MADHUKARA, J.; ANTONY, M. K1 divergence based agglomerative clustering for automated vitiligo grading. In: **The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. [S.l.: s.n.], 2015. p. 2700 – 2709.

HAMEED, M.; SHARIF, M.; RAZA, M.; HAIDER, S.; IQBAL, M. Framework for the comparison of classifiers for medical image segmentation with transform and moment based features. *Research Journal of Recent Sciences*, v. 2, p. 1–10, 01 2013.

HOSSIN, M.; SULAIMAN, M. N. A review on evaluation metrics for data classification evaluations. *International Journal of Data Mining and Knowledge Management Process (IJDKP)*, v. 5, n. 2, p. 1–11, 2015.

HRUSCHKA, E. R.; EBECKEN, N. F. F. A genetic algorithm for cluster analysis. *Intell. Data Anal.*, IOS Press, NLD, v. 7, n. 1, p. 15–25, jan. 2003. ISSN 1088-467X.

JAIN, A. K. Data clustering: 50 years beyond k-means. *Pattern Recogn. Lett.*, Elsevier Science Inc., USA, v. 31, n. 8, p. 651–666, jun. 2010. ISSN 0167-8655. Disponível em: <<https://doi.org/10.1016/j.patrec.2009.09.011>>.

JAIN, A. K.; DUBES, R. C. **Algorithms for Clustering Data**. USA: Prentice-Hall, Inc., 1988. ISBN 013022278X.

KAMENCAY, P.; RADILOVA, M.; HUDEC, R.; JARINA, R.; BENCO, M.; HLUBIK, J. A novel approach to face recognition using image segmentation based on spca-knn method. *Radioengineering*, v. 22, p. 92–99, 04 2013.

KANIADAKIS, G. H-theorem and generalized entropies within the framework of nonlinear kinetics. *Physics Letters A*, Elsevier BV, v. 288, n. 5-6, p. 283–291, Oct 2001. ISSN 0375-9601.

KATKAR, J. A.; BARASKAR, T. Medical image segmentation using pca and k-mean clustering algorithm. In: . [S.l.]: Fourth Post Graduate Conference, 2015. p. 1–6.

KHAN, W. Image segmentation techniques: A survey. *Journal of Image and Graphics*, v. 1, p. 166–170, 01 2014.

KIWIEL, K. C. Proximal minimization methods with generalized bregman functions. *SIAM J. Control Optim.*, Society for Industrial and Applied Mathematics, USA, v. 35, n. 4, p. 1142–1168, 1997. ISSN 0363-0129. Disponível em: <<https://doi.org/10.1137/S0363012995281742>>.

KONDRAK, G.; MARCU, D.; KNIGHT, K. Cognates can improve statistical translation models. In: **Companion Volume of the Proceedings of HLT-NAACL 2003 - Short Papers**. [s.n.], 2003. p. 46–48. Disponível em: <<https://www.aclweb.org/anthology/N03-2016>>.

KULLBACK, S.; LEIBLER, R. A. On information and sufficiency. **Ann. Math. Statist.**, v. 22, n. 1, p. 79–86, 1951.

LINDE, Y.; BUZO, A.; GRAY, R. M. An algorithm for vector quantizer design. **IEEE Transactions on Communications**, v. 28, p. 84–95, 1980.

LIU, C.; BELKIN, M. Clustering with bregman divergences: An asymptotic analysis. In: **Proceedings of the 30th International Conference on Neural Information Processing Systems**. Red Hook, NY, USA: Curran Associates Inc., 2016. (NIPS'16), p. 2351–2359. ISBN 9781510838819.

LIU, M. **Total Bregman Divergence, a Robust Divergence Measure, and Its Applications**. Tese (Doutorado) — University of Florida, USA, 2011.

LIU, M.; VEMURI, B. C.; AMARI, S.-i.; NIELSEN, F. Shape retrieval using hierarchical total bregman soft clustering. **IEEE Trans. Pattern Anal. Mach. Intell.**, v. 34, n. 12, p. 2407–2419, 2012. Disponível em: <<http://dblp.uni-trier.de/db/journals/pami/pami34.html#LiuVAN12>>.

MACQUEEN, J. Some methods for classification and analysis of multivariate observations. In: **Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics**. Berkeley, Calif.: University of California Press, 1967. p. 281–297. Disponível em: <<https://projecteuclid.org/euclid.bsm/1200512992>>.

MAMAEV, A. **Flowers Recognition**. 2018. Conjunto de dados disponibilizado por Kaggle, <<https://www.kaggle.com/alxmamaev/flowers-recognition>>.

MARTINEZ, F. Z.; BOQUERA, S. E.; BLEDA, J. M. C. Behaviour-based clustering of neural networks applied to document enhancement. *Computational and Ambient Intelligence*, p. 144–151, 2007.

OTSU, N. A threshold selection method from gray level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, v. 9, n. 1, p. 62–66, 1979.

PAVIOTTI, J. R. **Considerações sobre o conceito de entropia na teoria da informação**. Dissertação (Mestrado) — UniCamp, Limeira, 2019.

ROCHA, B. M. **O uso da Divergência de Kullback-Leibler e da Divergência Generalizada como medida de similaridade em sistemas CBIR**. Dissertação (Mestrado) — Universidade Federal de Uberlândia, Uberlândia, 2015.

ROMERO, A.; LACASSAGNE, L.; GOUIFFÈS, M. Total bregman divergence for multiple object tracking. In: . [S.l.: s.n.], 2013. p. 5.

SAKJI-NSIBI, S.; BENAZZA-BENYAHIA, A. Fast scalable retrieval of multispectral images with kullback-leibler divergence. In: **Proceedings of the International Conference on Image Processing, ICIP 2010, September 26-29, Hong Kong, China**. IEEE, 2010. p. 2333–2336. Disponível em: <<https://doi.org/10.1109/ICIP.2010.5653932>>.

SALVADOR, S.; CHAN, P. Determining the number of clusters/segments in hierarchical clustering/segmentation algorithms. In: **Proceedings of the 16th IEEE International Conference on Tools with Artificial Intelligence**. USA: IEEE Computer Society, 2004. (ICTAI '04), p. 576–584. ISBN 076952236X. Disponível em: <<https://doi.org/10.1109/ICTAI.2004.50>>.

SEMENOV, V. V. Modified extragradient method with bregman divergence for variational inequalities. **Journal of Automation and Information Sciences**, v. 50, p. 26–37, 2018.

SHANNON, C. E. A mathematical theory of communication. **The Bell System Technical Journal**, v. 27, p. 379–423, 623–656, July, October 1948. Disponível em: <<http://cm.bell-labs.com/cm/ms/what/shannonday/shannon1948.pdf>>.

SOKOLOVA, M.; LAPALME, G. A systematic analysis of performance measures for classification tasks. **ELSEVIER**, v. 45, p. 427–437, 2009.

SONG, Y.; YAN, H. Image segmentation techniques overview. In: **2017 Asia Modelling Symposium (AMS)**. [S.l.: s.n.], 2017. p. 103–107.

SOUZA, D. C.; CAVALCANTE, C. C.; VIGELIS, R. F. Modified maximum likelihood estimator. 2016 IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM), p. 1–5, 2016.

SØRENSEN, T. **A Method of Establishing Groups of Equal Amplitude in Plant Sociology Based on Similarity of Species Content and Its Application to Analyses of the Vegetation on Danish Commons**. [S. l.]: København: I KOMMISSION HOS EJNAR MUNKSGAARD, 1948.

TAKENOUCI, T.; ISHII, S. Binary classifiers ensemble based on bregman divergence for multi-class classification. **Neurocomputing**, v. 273, p. 424–434, 2018. ISSN 0925-2312.

TELGARSKY, M. J.; DASGUPTA, S. Agglomerative bregman clustering. In: **Proceedings of the 29th International Conference on Machine Learning, ICML 2012**. [S.l.: s.n.], 2012. v. 2, p. 1527–1534. ISBN 9781450312851. 29th International Conference on Machine Learning, ICML 2012 ; Conference date: 26-06-2012 Through 01-07-2012.

THURNER, S.; COROMINAS-MURTRA, B.; HANEL, R. Three faces of entropy for complex systems: Information, thermodynamics, and the maximum entropy principle. **Phys. Rev. E**, American Physical Society, v. 96, p. 032124, Sep 2017. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevE.96.032124>>.

TRIBUS, M.; MCIRVINE, E. C. Energy and information. **Scientific American**, Scientific American, a division of Nature America, Inc., v. 225, n. 3, p. 179–190, 1971. ISSN 00368733, 19467087. Disponível em: <<http://www.jstor.org/stable/24923125>>.

TSALLIS, C. Possible Generalization of Boltzmann-Gibbs Statistics. **J. Statist. Phys.**, v. 52, p. 479–487, 1988.

TSALLIS, C. What are the numbers that experiments provide. In: **Quimica Nova**. [S.l.: s.n.], 1994. v. 17, n. 6, p. 468–471.

VEMURI, B. C.; LIU, M.; AMARI, S.; NIELSEN, F. Total bregman divergence and its applications to dti analysis. **IEEE Transactions on Medical Imaging**, v. 30, n. 2, p. 475–483, 2011.

- Vigelis, R. F.; Andrade, L. H. F. D.; Cavalcante, C. C. Properties of a generalized divergence related to tsallis generalized divergence. **IEEE Transactions on Information Theory**, v. 66, n. 5, p. 2891–2897, 2020.
- VIGELIS, R. F.; CAVALCANTE, C. C. On φ -families of probability distributions. *Journal of Theoretical Probability*, v. 26, n. 3, p. 870–884, 2013.
- VIGELIS, R. F.; CAVALCANTE, C. C. Information geometry: An introduction to new models for signal processing. *Signals and Images - CRC Press*, p. 455–491, 2015.
- WANG, Z.; WANG, E.; ZHU, Y. Image segmentation evaluation: a survey of methods. *Artificial Intelligence Review*, v. 53, p. 5637–5674, 2020.
- WANJARI, M.; YEOTIKAR, V.; KALASKAR, K.; MAHENDRA, P.; DHORE, M. Document image segmentation using k-means clustering technique. In: . [S.l.: s.n.], 2015. p. 95–98.
- WEI, H.; CHEN, L.; GUO, L. Kl divergence-based fuzzy cluster ensemble for image segmentation. **Entropy**, MDPI AG, v. 20, p. 17, 2018. ISSN 1099-4300. Disponível em: <<http://dx.doi.org/10.3390/e20040273>>.
- WU, C.; ZHANG, X. Total bregman divergence-based fuzzy local information c-means clustering for robust image segmentation. *Applied Soft Computing*, v. 94, p. 31, 2020. ISSN 1568-4946.
- XING, E.; NG, A.; JORDAN, M.; RUSSELL, S. Distance metric learning with applications to clustering with side information. In: **16th Conference on Advances in Neural Information Processing Systems**. [s.n.], 2002. p. 505–512. Disponível em: <<http://sherry.ifi.unizh.ch/context/2203340/0>>.
- XU, B.; BU, J.; CHEN, C.; CAI, D. A bregman divergence optimization framework for ranking on data manifold and its new extensions. 2012. Disponível em: <<https://www.aaai.org/ocs/index.php/AAAI/AAAI12/paper/view/4828/5540>>.
- ZAKANI, F. R.; ARHID, K.; BOUKSIM, M.; ABOULFATAH, M.; GADI, T. New measure for objective evaluation of mesh segmentation algorithms. 2016 4th IEEE International Colloquium on Information Science and Technology (CiSt), p. 416–421, 2016.
- ZEZULA, P.; AMATO, G.; DOHNAL, V.; BATKO, M. **Similarity Search: The Metric Space Approach**. [S.l.]: Springer, 2006. v. 32. (Advances in Database Systems, v. 32). ISBN 0-387-29146-6.
- ZHANG, H.; FRITTS, J. E.; GOLDMAN, S. A. Image segmentation evaluation: A survey of unsupervised methods. *Computer Vision and Image Understanding*, v. 110, p. 260–280, 2008.
- ZHANG, J. Divergence function, duality, and convex analysis. **Neural Comput.**, MIT Press, Cambridge, MA, USA, v. 16, n. 1, p. 159–195, jan. 2004. ISSN 0899-7667. Disponível em: <<https://doi.org/10.1162/08997660460734047>>.
- ZHANG, Z.; OOI, B.; PARTHASARATHY, S.; TUNG, A. Similarity search on bregman divergence: Towards non-metric indexing. **PVLDB**, v. 2, p. 13–24, 2009.
- ZHAO, X.; LI, Y.; ZHAO, Q. Mahalanobis distance based on fuzzy clustering algorithm for image segmentation. **Digit. Signal Process.**, Academic Press, Inc., USA, v. 43, n. C, p. 8–16, 2015. ISSN 1051-2004. Disponível em: <<https://doi.org/10.1016/j.dsp.2015.04.009>>.

ZHAO, X.; TURK, M.; LI, W.; LIEN, K.-c.; WANG, G. A multilevel image thresholding segmentation algorithm based on two-dimensional k-l divergence and modified particle swarm optimization. *Applied Soft Computing*, v. 48, p. 151–159, 2016. ISSN 1568-4946. Disponível em: <<https://doi.org/10.1016/j.asoc.2016.07.016>>

ZIJDENBOS, A.; DAWANT, B.; MARGOLIN, R.; PALMER, A. Morphometric analysis of white matter lesions in mr images: method and validation. *IEEE Transactions on Medical Imaging*, v. 13, n. 4, p. 716–724, 1994.

ANEXO A – IMAGENS UTILIZADAS PARA OS TESTES

Figura 16 – Imagens do grupo 1 do Experimento I

(a) Imagem 1

A new offline handwritten database for the language, which contains full Spanish sentences, has recently been developed: the Spartacus database (which stands for Spanish Restricted-domain of Cursive Script). There were two main reasons for creating this corpus. First of all, most databases do not contain Spanish sentences, even though Spanish is a widespread major language. Another important reason was to create a corpus from semantic tasks. These tasks are commonly used in practical applications, the use of linguistic knowledge beyond the lexical process.

As the Spartacus database consisted mainly of short sentences and did not contain long paragraphs, the writers were given a set of sentences in fixed places: dedicated to specific forms. Next figure shows one of the forms used in the process. These forms also contain a brief set

(b) Imagem 2

There are several classic spatial filters for image processing or eliminating high frequency noise. The mean filter, the median filter, the opening filter, and the closing filter are frequently used. The mean filter is a lowpass or smoothing filter that averages the pixel values with the neighborhood. The median filter removes the image noise but blurs the image. The opening filter calculates the median of the values in the neighborhood for each pixel, thereby reducing the noise. Finally, the opening closing filter is a morphological filter that combines

(c) Imagem 3

A new offline handwritten database for the language, which contains full Spanish sentences, has recently been developed: the Spartacus database (which stands for Spanish Restricted-domain of Cursive Script). There were two main reasons for creating this corpus. First of all, most databases do not contain Spanish sentences, even though Spanish is a widespread major language. Another important reason was to create a corpus from semantic tasks. These tasks are commonly used in practical applications, the use of linguistic knowledge beyond the lexical process.

(d) Imagem 4

There exist several methods to design forms to be filled in. For instance, fields may be surrounded by bounding boxes, by light rectangles, or by guiding rulers. These methods are used where to write and, therefore, minimize the overlap with other parts of the form. These guides can be located on a separate sheet of paper that is located below the form or directly on the form. The use of a separate sheet is much better from the point of view of the quality of the scanned image, but requires giving more instructions to the user.

(e) Imagem 5

There exist several methods to design forms to be filled in. For instance, fields may be surrounded by light rectangles or by guiding rulers. These methods are used where to write and, therefore, minimize the overlap with other parts of the form. These guides can be located on a separate sheet of paper that is located below the form or directly on the form. The use of a separate sheet is much better from the point of view of the quality of the scanned image, but requires giving more instructions to the user. Guiding rulers printed on the form are more practical. Light rectangles can be removed more easily than dark lines whenever the handwritten text is present. Nevertheless, other practical issues must be considered. The best way to print these light rectangles

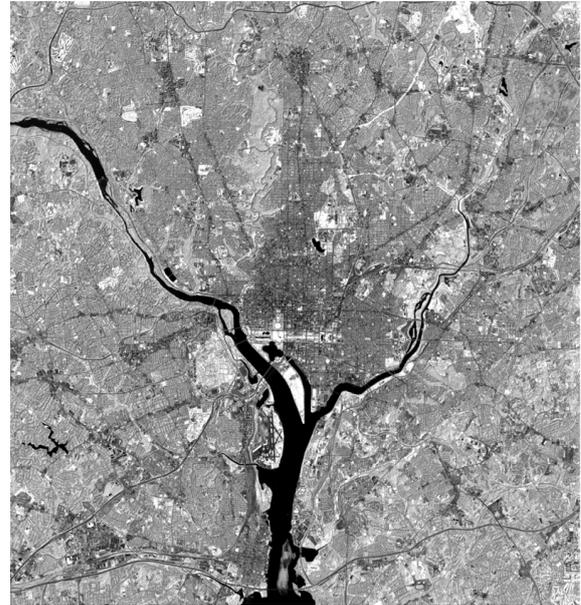
Figura 17 – Imagens do grupo 2 do Experimento I
(a) Imagem 1



(b) Imagem 2



(c) Imagem 3



(d) Imagem 4



(e) Imagem 5



Figura 18 – Imagens do grupo 1 do Experimento II

(a) Imagem 1



(b) Imagem 2



(c) Imagem 3



(d) Imagem 4



(e) Imagem 5



Figura 19 – Imagens do grupo 2 do Experimento II

(a) Imagem 1



(b) Imagem 3



(c) Imagem 2



(d) Imagem 4



(e) Imagem 5



Figura 20 – Imagens do grupo 3 do Experimento II

(a) Imagem 1



(b) Imagem 2



(d) Imagem 4



(c) Imagem 3



(e) Imagem 5

