

IMPLEMENTAÇÃO DE UMA VERSÃO DE PERCOLAÇÃO MULTIESCALA E MULTIDIMENSIONAL PARA EXTRAÇÃO DE CARACTERÍSTICAS EM IMAGENS HISTOLÓGICAS

Adão Marcelo Cuevas¹, Luiz Fernando Segato dos Santos¹

Instituto Federal do Mato Grosso do Sul – Aquidauana-MS

adao.cuevas@estudante.ifms.edu.br, luiz.santos@ifms.edu.br

Resumo

Neste trabalho foi realizado um estudo sobre importantes técnicas para extração de características de imagens no contexto da histopatologia. Primeiramente foi realizada uma pesquisa em artigos científicos sobre melhorias nas técnicas básicas e também na busca por novas técnicas. Uma nova técnica encontrada foi a da percolação multiescala e multidimensional. A partir da técnica foi realizado um estudo para entender seu funcionamento e posteriormente implementá-la e validá-la. Algumas imagens fractais foram pesquisadas e coletadas para os testes iniciais. Uma implementação da técnica em Python também foi encontrada na internet e alguns testes iniciais sobre as imagens encontradas foram realizados para uma validação inicial.

Palavras-chave: Percolação, validação, implementação.

Introdução

O *machine learning* tem contribuído na obtenção de modelos computacionais que possam auxiliar estes especialistas no reconhecimento de padrões e na classificação de imagens médicas. Diante disso a busca por técnicas capazes de representar as principais características no auxílio ao diagnóstico é uma prioridade desta área de estudo (LATEEF; SAHRAN; MAKLUMAT, 2019; DING, 2022).

A utilização de técnicas fractais e de entropia na representação de características de imagens histológicas contribuiu em diversas aplicações.

As técnicas fractais foram exploradas por meio das seguintes métricas: Dimensão Fractal (DF), Lacunaridade (LAC) e percolação (PERC). A DF foi aplicada em diversos contextos como (PANIGRAHI; et al., 2019; DA SILVA; et al., 2021). Este descritor está intimamente ligado à complexidade da textura analisada e obtém a dimensão de estruturas irregulares comuns neste tipo de imagem. A medida representa o quão irregular o objeto de interesse é. A LAC foi aplicada em (ARALICA, 2020) e é considerada uma medida complementar à DF, uma vez que proporciona uma análise em como o objeto preenche o espaço em que está contido. A medida também está mais ligada à informação espacial, enquanto que a DF é uma medida mais estatística.

A PERC por sua vez é uma medida que se baseia no conceito de percolação de fluidos. A métrica original representa um modelo matemático relacionado à conectividade de elementos em um sistema. Dada uma matriz quadrada e

pontos a e b nessa matriz, e se houver conectividade entre os pontos de tal forma que todos os pontos adjacentes a estes também estão conectados, esses pontos pertencem a um aglomerado (ESSAM, 1980). Esses pontos simbolizam a passagem do fluido entre os pontos a e b e são também chamados de poros. Existem diversos modelos de percolação para imagens, mas em geral, cada aglomerado tem uma probabilidade de se estender de uma extremidade a outra da imagem, se essa probabilidade atingir um determinado limiar então haverá a percolação. Este descritor é importante para as imagens histológicas porque suas texturas apresentam formato poroso, normalmente são compostas por células contidas em um tecido também conhecido por estroma. A percolação foi unificada à abordagens multiescala e multidimensional e aplicada com sucesso em (ROBERTO; et al., 2017).

Com relação às entropias, os modelos mais aplicados foram da entropia de Shannon e da Amostral. A entropia de Shannon foi aplicada principalmente em contextos unidimensionais e no contexto de imagens em tarefas de segmentação, mas recentemente contribuiu para tarefas de classificação de imagens colorretais H&E (DOS SANTOS; et al., 2022). A entropia Amostral também foi aplicada em diversos contextos de imagens médicas, principalmente em tarefas de classificação (DOS SANTOS; et al., 2018). Esta medida, assim como a PERC e a LAC está ligada à estrutura espacial da textura analisada.

A contribuição destas técnicas são de relevância científica para muitas bases de imagens, portanto suas implementações e validações para que possam compor sistemas CAD (*Computer Aided Diagnosis*) seria de grande ajuda aos interessados.

Com base nessas justificativas e mediante o projeto proposto foram realizadas pesquisas por trabalhos técnicos que apresentassem o código (algoritmo) ou parte deste para fins de aproveitamento. A ideia é de estudar os códigos baseados nas técnicas descritas e validá-los a fim de armazenar em repositórios específicos para que pudessem ser aproveitados em outros trabalhos do IFMS - Campus Aquidauana, como TCCs ou iniciações científicas. Para tanto os objetivos são: (a) Identificar os aprimoramentos multiescalas das principais técnicas multiescalas baseadas em modelos fractais e de entropia dedicadas à extração de características de imagens médicas histológicas; (b) Implementar as técnicas e aprimoramentos multiescalas identificados em (a) e validá-los conforme consta na literatura e; (c) Armazenar as técnicas

validadas em (b) em um repositório local para facilitar a execução outros projetos de pesquisa e de TCCs futuros que desenvolvam tecnologias para o auxílio dos profissionais da área médica.

Metodologia

Como o primeiro passo, as técnicas descritas anteriormente foram exaustivamente pesquisadas na Literatura. A partir dos artigos encontrados foi detectada uma nova abordagem de aplicação além da multiescala, a multidimensional (ROBERTO; et al., 2017). Para isso os esforços de pesquisa tiveram que ser estendidos sob essa nova abordagem.

A estratégia foi de pesquisar pelos códigos que também apresentassem a abordagem multidimensional, uma extensão das técnicas fractais e de entropia para imagens coloridas, conforme as considerações de (IVANOVICI; RICHARD; DECEAN, 2009). Durante as pesquisas foram encontrados dois códigos de dimensão fractal porém ainda não validados que serão analisados e estudados posteriormente e, um código sobre percolação, mais precisamente sobre rotulagem de aglomerados de Hoshen-Kopelman, que pode ser visto a seguir na forma de algoritmo:

Algoritmo de Hoshen-Kopelman aplicado para definir abertura de poros e aglomerados (Guilherme F. Roberto, 2017):

```

1: Label ← 0
2: for i=1 to i = height do
3: for j=1 to j = width do
4: Label ← Label +1
5: LABEL_CLUSTER(i, j)
6: end for
7: end for
8: LABEL_CLUSTER(i, j){
9: site(i, j) ← Label
10: if i< height and site(i+1, j) = -1 then
11: LABEL_CLUSTER(i+1, j)
12: else if j< width and site(i, j + 1) = -1 then
13: LABEL_CLUSTER(i, j + 1)
14: else if i>1 and site(i - 1, j) = -1 then

```

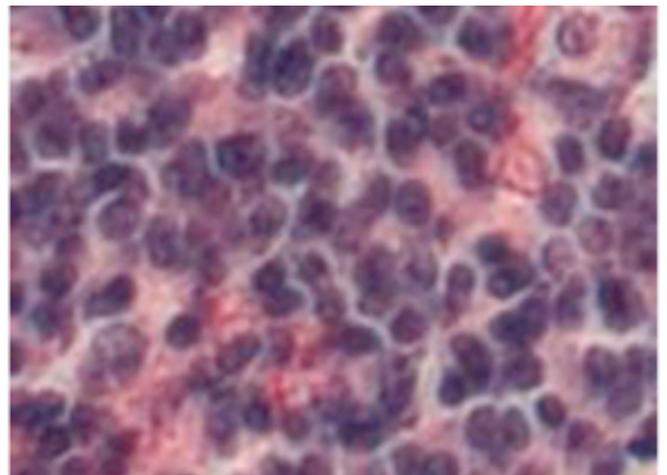
```

15: LABEL_CLUSTER(i - 1, j)
16: else if i>1 and site(i, j - 1) = -1 then
17: LABEL_CLUSTER(i, j - 1)
18: end if
19: }

```

Vários ajustes tiveram que ser realizados no código a fim de realizar os primeiros testes em linguagem Python. A partir disso foram escolhidas imagens com valores conhecidos na aplicação da técnica de percolação. Algumas destas imagens podem ser vistas a seguir:

Figura 1 - Imagem de um estrutura histológica contendo linfoma Non-Hodgkin



Para a aplicação da técnica nas imagens serão escolhidas métricas como as especificadas em (ROBERTO; et al., 2017), a saber o número médio de aglomerados ($C(L)$) por caixa de tamanho L , a quantidade de caixas percolantes ($Q(L)$) para cada escala L e a cobertura do maior aglomerado por caixa ($M(L)$) para cada caixa de lado L .

De acordo com esses mesmos autores, uma etapa para o cálculo da percolação é a rotulagem de aglomerados, normalmente implementada via algoritmo de Hoshen-Kopelman. A seguir o código encontrado será checado também para verificar se esta rotulagem e cada métrica é contemplada, caso contrário será adaptado para tal. A partir disso as métricas serão calculadas comparadas com os resultados fornecidos por (ROBERTO; et al., 2017) para validação. Apesar de testes mais detalhados não puderem ser realizados, testes iniciais com o algoritmo de percolação

encontrado foram realizados e mostraram que o mesmo pode ser aproveitado para as implementações restantes.

Resultados Esperados

Os resultados esperados para este árduo trabalho de pesquisa, implementação e validação são:

1. Um repositório (Github) de técnicas multiescalas implementadas e validadas para quantificação e extração de características de imagens histológicas da área médica, somando as novas técnicas implementadas às implementadas anteriormente;
2. O entendimento detalhado da relação entre as propriedades histológicas de imagens representativas de tecidos saudáveis e cancerígenos, e as técnicas implementadas;
3. O desenvolvimento de novos métodos, a partir dos TCCs e iniciação científica, com resultados significativos que utilizem as técnicas implementadas.

Considerações Finais

O trabalho proposto considerou um amplo levantamento de técnicas fractais e de entropia, seu entendimento para implementação e validação e a busca de códigos prontos ou semi-prontos para análises e adequações. Testes mais detalhados serão realizados posteriormente, no entanto alguns testes iniciais realizados mostraram que os códigos encontrados serão úteis ao desenvolvimento e implementação das técnicas encontradas.

Agradecimentos

Ao IFMS pela bolsa concedida para execução do projeto.

Referências

ARALICA, Gorana et al. Prognostic Significance of Lacunarity in Preoperative Biopsy of Colorectal Cancer. **Pathology & Oncology Research**, v. 26, n. 4, p. 2567-2576, 2020.

DA SILVA, Lucas Glaucio et al. Fractal dimension analysis as an easy computational approach to improve breast cancer histopathological diagnosis. **Applied Microscopy**, v. 51, n. 1, p. 1-9, 2021.

DING, Saisai et al. Fractal Graph Convolutional Network with MLP-mixer based Multi-path Feature Fusion for Classification of Histopathological Images. **Expert Systems with Applications**, p. 118793, 2022.

DOS SANTOS, Luiz Fernando Segato et al. Multidimensional and fuzzy sample entropy (SampEnMF) for quantifying H&E histological images of colorectal cancer.

Computers in biology and medicine, v. 103, p. 148-160, 2018.

DOS SANTOS, Luiz Fernando Segato et al. Multidimensional shannon entropy (H M) as an approach to classify H&E colorectal images. In: **2022 29th International Conference on Systems, Signals and Image Processing (IWSSIP)**. IEEE, 2022. p. 1-4.

ESSAM, John W. Percolation theory. **Reports on progress in physics**, v. 43, n. 7, p. 833, 1980.

IVANOVICI, M.; RICHARD, N.; DECEAN, H. Fractal dimension and lacunarity of psoriatic lesions-a colour approach. **medicine**, v. 6, n. 4, p. 7, 2009.

LATEEF, A.; SAHRAN, Shahnorbanun; MAKLUMAT, Sains. **Histopathology image classification based on color space**. Tech. rep., Universiti Kebangsaan Malaysia. <https://www.ftsm.ukm.my/cybersecurity/file/research/technicalreport/PS-FTSM-2019-002.pdf>, 2019.

PANIGRAHI, Santisudha et al. Fractal Geometry for Early Detection and Histopathological Analysis of Oral Cancer. In: **International Conference on Mining Intelligence and Knowledge Exploration**. Springer, Cham, 2019. p. 177-185.

ROBERTO, Guilherme F. et al. Features based on the percolation theory for quantification of non-hodgkin lymphomas. **Computers in biology and medicine**, v. 91, p. 135-147, 2017.

IMPLEMENTATION OF A MULTISCALE AND MULTIDIMENSIONAL PERCOLATION VERSION FOR FEATURE EXTRACTION IN HISTOLOGICAL IMAGES

Abstract: In this work, a study was conducted on important techniques for feature extract from images in the context of histopathology. First a research in scientific papers about improvements in the basic techniques and also in the search for new techniques was performed. A new technique found was the multiscale and multidimensional percolation. Based on this technique a study was done to understand how it works and later implement and validate it. Some fractal images were researched and collected for initial tests. An implementation of the technique in Python was also found on the internet and some initial tests on the images found were done for an initial validation.

Keywords: .Percolation, validation, implementation.