

UNIVERSIDADE VILA VELHA - ES
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS FARMACÊUTICAS

**COMPARAÇÃO ENTRE DOIS PROGRAMAS PARA ANÁLISE DE
IMAGEM, APRENDIZADO DE MÁQUINA E SUBSEQUENTE
CLASSIFICAÇÃO**

GABRIELLY PEREIRA RIBEIRO

VILA VELHA
JUNHO / 2019

UNIVERSIDADE VILA VELHA - ES
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS FARMACÊUTICAS

**COMPARAÇÃO ENTRE DOIS PROGRAMAS PARA ANÁLISE DE
IMAGEM, APRENDIZADO DE MÁQUINA E SUBSEQUENTE
CLASSIFICAÇÃO**

Dissertação apresentada à Universidade Vila Velha, como pré-requisito do Programa de Pós-graduação em Ciências Farmacêuticas, para a obtenção do grau de Mestre em Ciências Farmacêuticas.

GABRIELLY PEREIRA RIBEIRO

VILA VELHA
JUNHO / 2019

Catálogo na publicação elaborada pela Biblioteca Central / UVV-ES

R484c Ribeiro, Gabrielly Pereira.
Comparação entre dois programas para análise de imagens,
aprendizado de máquina e subsequente classificação / Gabrielly
Pereira Ribeiro. – 2019.

27 f. : il.

Orientador: Dominik Lenz.
Dissertação (mestrado em Ciências Farmacêuticas)
Universidade Vila Velha, 2019.
Inclui bibliografias.

1. Farmacologia e terapêutica. 2. Diagnóstico por imagem
I. Lenz, Dominik. II. Universidade Vila Velha. III. Título.

CDD 615

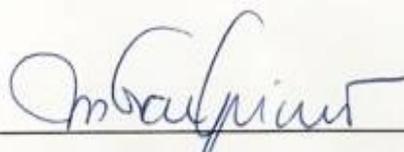
GABRIELLY PEREIRA RIBEIRO

COMPARAÇÃO ENTRE DOIS PROGRAMAS PARA ANÁLISE DE
IMAGEM, APRENDIZADO DE MÁQUINA E SUBSEQUENTE
CLASSIFICAÇÃO

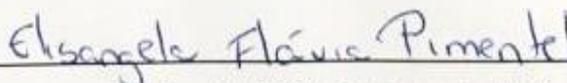
Dissertação apresentada à Universidade
Vila Velha, como pré-requisito do
Programa de Pós-graduação em Ciências
Farmacêuticas, para a obtenção do grau
de Mestre em Ciências Farmacêuticas.

Aprovada em 10 de junho de 2019,

Banca Examinadora:



Marco Cesar Cunegundes Guimarães - UFES



Elisângela Flávia Pimentel - UVV



Dominik Lenz - UVV

Orientador

Dedico esta dissertação aos meus pais Nair Maria Pereira e Osvaldo José da Costa, que sempre me apoiaram nos estudos e na realização dos meus sonhos.

AGRADECIMENTOS

Primeiramente a Deus que me concedeu forças e sabedoria.

Ao meu orientador, Professor Dr. Dominik Lenz, por seu apoio e amizade que durante 5 anos me motivou e ensinou com muita paciência. Por sua dedicação e competência para que juntos pudéssemos realizar o melhor trabalho. Por sua disponibilidade em sempre se manter presente para transmitir mais conhecimento. Pelos sucos de laranja de todas as tardes que me deixavam cada vez mais apaixonada pela ciência. Por me tornar uma pesquisadora. Obrigada Campeão.

A toda minha família que sempre me apoiou a continuar no caminho dos estudos. Em especial minha mãe Nair Maria Pereira que não mediu esforços para que eu realizasse mais essa etapa. Obrigada mãe, sem você nada disso seria possível.

A minha querida Ada Maria Kerr por ser incentivadora principiante deste título, por todo apoio, confiança e investimento em mim. Pelo carinho que me ofereceu em cada ligação preocupada com meu futuro, muito obrigada.

Ao meu irmão Tiago Pereira da Costa que sempre acreditou na minha capacidade de ir além da minha zona de conforto. Obrigada pelo carinho e várias palavras de ânimo, que me motivaram a não desistir dos meus sonhos.

Ao meu noivo Rodrigo Oliveira das Neves que compreendeu quando eu precisei trocar os preparativos do casamento pelos da defesa do mestrado. Por me motivar a crescer cada vez mais, e ao mesmo tempo ser humilde reconhecendo todos ao meu lado que cooperaram para este crescimento. Obrigada amor.

A todos os professores do mestrado que de alguma forma contribuíram para minha formação.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela bolsa de estudo concedida.

A todos os meus amigos, tanto antigos como os novos que o mestrado me concedeu, com vocês todo esse tempo se tornou mais leve para a realização do trabalho.

À Universidade Vila Velha e ao Programa de Pós-graduação em Ciências Farmacêuticas que abriu as portas para meu futuro.

Muito obrigada a todos.

“Armed with will and determination
And grace, too.”

Gordon Downie

SUMÁRIO

LISTA DE ABREVIATURAS
RESUMO
ABSTRACT
1. INTRODUÇÃO	12
2. MATERIAL E MÉTODOS	13
2.1 DELINEAMENTO DO ESTUDO.....	14
2.1.1 DECLARAÇÃO ÉTICA.....	14
2.2 MATERIAIS.....	14
2.2.1 AQUISIÇÃO DE IMAGENS.....	14
2.2.2 PREPARAÇÃO DO SLIDE - IMUNO-HISTOQUÍMICA.....	14
2.3 ANÁLISE DE IMAGEM.....	15
2.4 APRENDIZAGEM DA MÁQUINA.....	16
2.4.1 CELPROFILLER.....	16
2.4.2 QUPATH.....	17
3. RESULTADOS	18
4. DISCUSSÃO	20
5. CONCLUSÕES	23
6. REFERÊNCIAS	24
7. ANEXOS	27

LISTA DE ABREVIATURAS

CI - Citometria de Imagem

CP - Cellprofiler

CPA - Cellprofiler Analyst

EPM - Erro padrão da média

CVL - Citometria por varredura a laser

QP - Qupath

RESUMO

RIBEIRO, Gabrielly Pereira, Universidade Vila Velha – ES, junho de 2019.
Comparação entre dois programas para análise de imagens, aprendizado de máquina e subsequente classificação.

Orientador: Prof. Dr. Dominik Lenz.

No início dos anos 50, a citometria de fluxo foi desenvolvida como o primeiro método para análise celular quantitativa automatizada. No início dos anos 90, o primeiro equipamento para citometria de imagem (citometria por varredura a laser, CVL) tornou-se comercialmente disponível. Como a citometria de fluxo foi considerada o padrão-ouro, vários estudos descobriram que os resultados da citometria de fluxo e da CVL geraram resultados comparáveis. Um dos primeiros programas de análise de imagens que incluiu parâmetros morfológicos foi o ImageJ, publicado em 1997. Um dos mais recentes programas de análise de imagens que não se limitam a imagens de fluorescência é o software livre CellProfiler. Em 2008, o mesmo grupo publicou um novo software, o CellProfilerAnalyst. Uma parte do CellProfiler Analyst é um classificador supervisionado baseado em aprendizado de máquina que permite aos usuários realizar diagnósticos baseados em imagem, por exemplo, diagnóstico celular baseado em morfologia. Outro software livre relativamente novo para análise de imagens é o QuPath. O objetivo do presente estudo foi comparar dois programas gratuitos para a realização de análise de imagens, CellProfiler e QuPath, e a subsequente classificação baseada no aprendizado de máquina. Para este estudo, imagens de tecido renal foram analisadas e os objetos identificados foram classificados. As mesmas imagens foram carregadas em ambos os programas de software. Análise estatística avançada foi utilizada para comparar os dois métodos. O ensaio de Bland-Altman mostrou que todas as diferenças estavam dentro da média $\pm 1,96$ * desvio padrão, ou seja, as diferenças são normalmente distribuídas, e os programas de software são comparáveis. Para as amostras analisadas (tecido renal corado com HIF e TUNEL), o uso do QuPath foi mais fácil, pois oferece análise de imagem sem processamento prévio das imagens (por exemplo, conversão em escala de cinza, intensidades invertidas) e um processo de aprendizado automatizado não supervisionado.

Palavras-chave: Citometria de Imagem; Aprendizado de máquina; Diagnóstico celular.

ABSTRACT

RIBEIRO, Gabrielly Pereira, University of Vila Velha – ES, June 2019. **Comparison between two programs for image analysis, machine learning and subsequent classification.**

Advisor: Dominik Lenz.

In the early 1950s, flow cytometry was developed as the first method for automated quantitative cellular analysis. In the early 1990s, the first equipment for image cytometry (laser scanning cytometry, LSC) became commercially available. As flow cytometry was considered the gold standard, various studies found that the results of flow cytometry and LSC generated comparable results. One of the first programs for image analysis that included morphological parameters was ImageJ, published in 1997. One of the newer programs for image analysis that is not limited to fluorescence images is the free software CellProfiler. In 2008, the same group published a new software, CellProfilerAnalyst. One part of CellProfilerAnalyst is a supervised machine-learning-based classifier that allows users to conduct imaging-based diagnoses, e.g., cellular diagnosis based on morphology. Another relatively new, free software for image analysis is QuPath. The aim of the present study was to compare two free programs for conducting image analysis, CellProfiler and QuPath, and the subsequent classification based on machine learning. For this study, images of renal tissue were analyzed, and the identified objects were classified. The same images were loaded in both software programs. Advanced statistical analysis was used to compare the two methods. The Bland-Altman assay showed that all of the differences were within the mean ± 1.96 * standard deviation, i.e., the differences are normally distributed, and the software programs are comparable. For the analyzed samples (renal tissue stained with HIF and TUNEL), the use of QuPath was easier because it offers image analysis without a previous processing of the images (e.g., conversion to grayscale, inverted intensities) and an unsupervised machine learning process.

Keywords: Image Cytometry; Machine learning; Cellular diagnosis.

1. INTRODUÇÃO

No início dos anos 50, a citometria de fluxo foi desenvolvida como o primeiro método para análise celular quantitativa automatizada (Givan, 2011). A citometria de fluxo permaneceu como padrão-ouro até o início dos anos 90, quando o primeiro equipamento para citometria de imagem (Citometria de Varredura a Laser, CVL) se tornou comercialmente disponível (Kamentsky e Kamentsky, 1991).

Com base na possibilidade de verificar morfologicamente a análise, esta nova técnica foi considerada como uma ponte para a lacuna entre a citometria multiparamétrica de alto rendimento, análise e documentação morfológica detalhada (Gerstner et al., 2006). Durante o processo de varredura, imagens de cada canal de fluorescência, bem como a dispersão direta foram geradas, mas não foram salvas no disco rígido. Este primeiro tipo de citometria de imagem foi projetado para analisar os mesmos parâmetros analisados por citometria de fluxo, ou seja, contagem de células e intensidade de fluorescência. Como a citometria de fluxo foi considerada o padrão ouro, vários estudos descobriram que a citometria de fluxo e a CVL geraram resultados comparáveis (Kamentsky e Kamentsky, 1991; Gerstner et al., 2006). Como uma vantagem da CVL, muitos estudos mencionaram a possibilidade de análise de tecidos e o fato de que um número substancialmente menor de amostras e menor volume de reagentes foram necessários (Gerstner et al., 2002; Lenz et al., 2005; Lenz et al., 2005b; Mittag et al., 2005; Laffers et al., 2006; Mosch et al., 2006).

Um dos primeiros programas de análise de imagens que incluiu parâmetros morfológicos foi o ImageJ, publicado em 1997 (Schneider et al., 2012).

Como um dos mais novos programas de análise de imagens, não se limitando a imagens de fluorescência, o software livre CellProfiler (CP) deve ser mencionado (Carpenter et al., 2006). Mais de 70 publicações revisadas por pares (em abril de 2018) representam sua importância científica. Em 2008, o mesmo grupo publicou um novo software, CellProfilerAnalyst (CPA) (Jones et al., 2008). Este software destina-se a trabalhar com os dados gerados com CP. Uma parte do CPA é um classificador supervisionado baseado em aprendizado de máquina que permite aos usuários realizar diagnósticos baseados em imagens, por exemplo, diagnóstico celular baseado em morfologia.

Outro novo software livre para análise de imagens é o QuPath (QP) (Paulik et al., 2017), lançado em 2017, portanto ainda não existem muitos estudos publicados que utilizaram este software. Além disso, os desenvolvedores não publicaram um manual para o QP. Os desenvolvedores afirmam que o software realiza diagnósticos celulares baseados em aprendizado de máquina sem supervisão de imagens inteiras de slides (Paulik et al., 2017).

O objetivo do presente estudo foi, portanto, comparar o software bem conhecido e um software relativamente novo em termos de análise de imagem, aprendizado de máquina e classificação celular. Para o presente estudo, imagens de tecido renal foram analisadas e os objetos identificados foram classificados. Os dois programas foram comparados usando estatísticas avançadas.

2. MATERIAL E MÉTODOS

2.1 DELINEAMENTO DO ESTUDO

O estudo foi realizado no Laboratório de Citometria de Imagem da Universidade de Vila Velha (UVV), localizado na cidade de Vila Velha, Espírito Santo, Brasil.

2.1.1 DECLARAÇÃO ÉTICA

As lâminas histológicas foram fornecidas pela Universidade Federal do Espírito Santo (UFES) e seu uso foi aprovado pelo Comitê de Ética CEUA / UFES (Protocolo nº 050/2013) (ANEXO).

2.2 MATERIAIS

Imagens de lâminas histológicas de tecido renal de ratos Wistar machos foram analisadas em duas amostras com diferentes marcadores, sendo 13 imagens marcadas com HIF-1 α (fator 1 hipóxia indutível α) e 17 imagens marcadas com TUNEL (terminal desoxinucleotidil transferase dUTP nick end rotulagem). Cada conjunto de dados continha mais de 10.000 objetos identificados, analisados e classificados.

2.2.1 AQUISIÇÃO DE IMAGEM

As lâminas foram fotografadas usando um microscópio Axio Vert.A1 Zeiss com lente objetiva 40x acoplada a uma câmera AxioCam MRC Zeiss e ao software Zen 2012. As imagens de cada slide foram salvas no formato *.tif e armazenadas individualmente em pastas separadas e corretamente identificadas pelo usuário.

2.2.2 PREPARAÇÃO DO SLIDE - IMUNOHISTOQUÍMICA

TÚNEL: A fragmentação de ADN internucleossômico foi detectada in situ utilizando o ensaio de finalização com marcador de uridina trifosfato (dUTP) terminal de desoxinucleotidiltransferase (TDT) (TUNEL) (kit TUNEL POD, Roche Diagnostics, Alemanha). Secções de parafina foram desidratadas e à base de xileno montadas sob lamínulas de vidro. A placenta humana foi usada como controle positivo. Os mesmos procedimentos foram aplicados aos controles negativo e positivo, com a exceção de que o controle negativo foi incubado na ausência da enzima TdT. Para visualizar a coloração, as secções foram tratadas com tetracloridrato de 3,3-diaminobenzidina (DAB; Sigma Chemical Co., St. Louis, MO, EUA). Os espécimes foram então

levemente contrastados com a hematoxilina de Mayer, desidratados e montados com meio de montagem à base de xileno sob lamínulas de vidro. O conteúdo da área de coloração positiva foi a média dos campos escolhidos aleatoriamente ao longo do córtex e da medula (Macedo et al. 2017).

HIF-1 α : Quatro micrômetros de secções de parafina em série do rim foram coradas com HIF-1 α (ab8366) anticorpo anti-rato monoclonal de rato (Abcam, Reino Unido, 1: 500). A coloração foi visualizada sob reação de peroxidase com tetracloridrato de 3,3-diaminobenzidina (DAB; Sigma Chemical Co., EUA). Os espécimes foram então ligeiramente contrastados com a hematoxilina de Mayer, desidratada e à base de xileno montada sob lamínulas de vidro. A amostra de controle positivo foi placenta humana, enquanto o controle negativo foi o mesmo material e incubado apenas com diluente de anticorpo (Buzin et al., 2015).

2.3 ANÁLISE DE IMAGEM

As imagens foram analisadas utilizando o software livre QP (versão 0.1.2) (<https://qpath.github.io/>) e CP (versão 2.1.1) (<https://cellprofiler.org/>).

No software QP, o algoritmo foi preparado para a análise das imagens obtidas. As imagens foram carregadas, as regiões foram selecionadas manualmente através do módulo de anotação do retângulo e posteriormente classificadas pelo módulo "detecção de célula positiva". Os valores estatísticos para as células contidas na anotação foram então visualizados em uma planilha onde medições de detecção foram compiladas. Depois de analisar a porcentagem de células positivas e negativas por área, os dados foram exportados como um arquivo para uma planilha do Excel. Um novo algoritmo foi programado para cada amostra em ambos os programas (tabela 1).

Tabela 1: Pipeline usado para análise de imagem com os dois programas.

<i>HIF</i>	<i>TUNEL</i>
Load Images	Load Images

Color To Gray	Color To Gray
Morph	Image Math - Invert
Identify Primary Objects	Identify Primary Objects
Identify Secondary Objects	Measure Object Size Shape
Identify Tertiary Objects	Filter Objects
Measure Object Size Shape	Measure Object Size Shape
Measure Object Intensity	Measure Object Intensity
Measure Object Size Shape	
Tools – Rectangle Tool	Tools – Rectangle Tool
Analyze	Analyze
Cell Analysis - Positive Cell	Cell Analysis - Positive Cell
Detection	Detection
Measure – Show Detection	Measure – Show Detection
Measurements	Measurements

No software CP, um algoritmo foi programado para realizar uma análise celular automatizada e reproduzível. As imagens foram convertidas em escala de cinza e as intensidades foram invertidas. Após a análise, os dados foram exportados para um formato (SQLite) para posterior análise e classificação utilizando o CPA. Este último software foi utilizado para realizar uma classificação após um processo de aprendizado de máquina supervisionado e para calcular a precisão da classificação. Por outro lado, o QP analisou as imagens sem a necessidade de processar as imagens e classificou os objetos sem supervisão, após ter escolhido o algoritmo apropriado para aprendizado da máquina.

2.4 APRENDIZAGEM DA MÁQUINA

2.4.1 CellProfiler

As classes foram separadas em células positivas e negativas, ou seja, com e sem a presença do marcador. O aprendizado da máquina, que consiste em treinar o software para reconhecer as diferentes classes, foi realizado simultaneamente.

Os algoritmos foram ensinados a partir do conjunto de dados de treino escolhidos durante a separação de marcadores celulares positivos e negativos. Após o treinamento, o algoritmo realizou previsões interativas sobre os dados de treinamento e foi corrigido quando houve uma discrepância. A aprendizagem foi concluída quando o algoritmo atingiu um nível aceitável de desempenho.

2.4.2 QuPath

A identificação das células na QP foi realizada por algoritmos específicos na aprendizagem não supervisionada (Bankhead et al., 2018). Essa classificação de objetos de aprendizado não supervisionados não tem resultados para prever, mas precisa encontrar padrões que ocorram naturalmente dentro dos dados (Deo, 2015). Os dados foram fornecidos, e uma classificação independente foi obtida. Houve uma caracterização para cada imagem de entrada, criando um modelo de recursos chamado clustering (Jordan e Mitchell, 2015), e foi realizado através de deconvolução de cores seguida de inteligência operacional para identificar os objetos no canal da hematoxilina ou a soma das densidades ópticas (Ruifrok e Johnston, 2001).

3. RESULTADOS

CP e QP são capazes de analisar imagens microscópicas de amostras de tecido coradas. Os objetos identificados neste estudo foram classificados como positivos ou negativos, o número total de objetos em cada amostra e os demais parâmetros avaliados estão representados na tabela 2.

Tabela 2: Comparação entre os dois softwares com diferentes parâmetros avaliados.

	CELLPROFILER	QUPATH
CARREGAR IMAGENS	unlimited	limited
APRENDIZADO DE MÁQUINA	Supervised	Unsupervised
TEMPO/ANÁLISE	≅ 2 hours	≅ 3 minutes
TUNEL/CONTAGEM	10050 objects	14148 objects
HIF/CONTAGEM	15484 objects	13660 objects

O CP suporta teoricamente um número ilimitado de imagens por análise, enquanto a análise de imagens com o QP é limitada a uma imagem por análise. Além disso, CP suporta a análise de culturas de células ou leucócitos, como descrito por Buzin et al. (2015) e Mittag et al. (2011). Finalmente, o software CP é mais flexível em termos de programação de um algoritmo e parâmetros a serem analisados.

Por outro lado, o QP é muito fácil de trabalhar quando se analisam amostras de tecido com células de diferentes classes visualmente detectáveis. Durante a análise de amostras com diferentes classes que não são claramente visíveis, a aplicação de QP para amostras é difícil porque o usuário não pode influenciar ativamente o processo de aprendizado de máquina não supervisionado. No entanto, em amostras histológicas com diferentes classes que são claramente visíveis, o software QP conduz uma análise de imagem rápida e subsequente classificação. Para culturas de células e amostras com classes de morfologia similar, CP é a melhor opção, pois o usuário pode interferir ativamente.

Imagens de leucócitos e esferas fluorescentes como estudado por Mittag et, al. (2011) e Buzin et al. (2015) foram analisados usando CPA, mas não por QP (dados

não mostrados). Foi obtida uma comparação por imagem de escore (CPA) e detecção de célula positiva (QP) - Por meio dessas ferramentas foi possível verificar qualitativa e quantitativamente os resultados em ambas as classificações. Os resultados da análise de imagem revelaram as capacidades de classificação celular de CPA e QP (figuras 1 e 2).

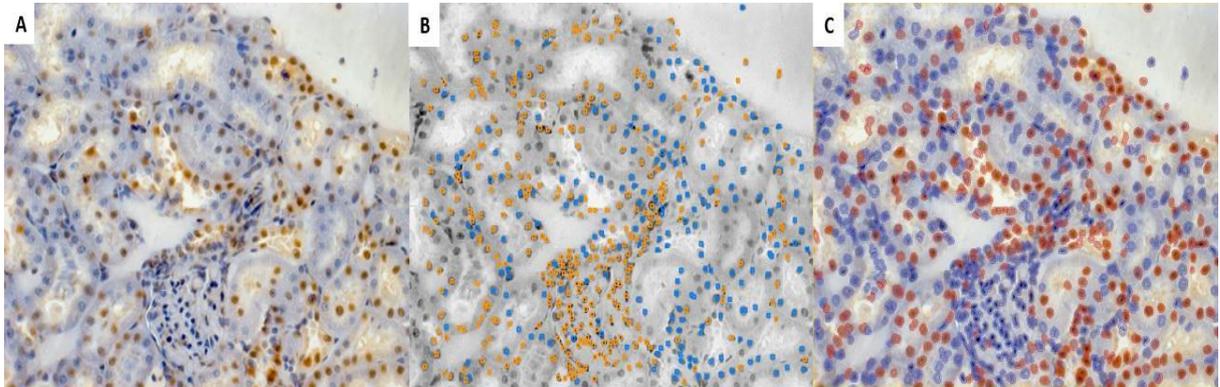


Figura 1: Tecido renal corado com TUNEL. (a) imagem original; Células laranja são positivas para TUNEL, objetos azuis são negativos; b) Células classificadas por CPA - os pontos alaranjados são positivos, os pontos azuis são negativos; (c) Células classificadas por QP, pontos azuis são positivos, pontos vermelhos são negativos.

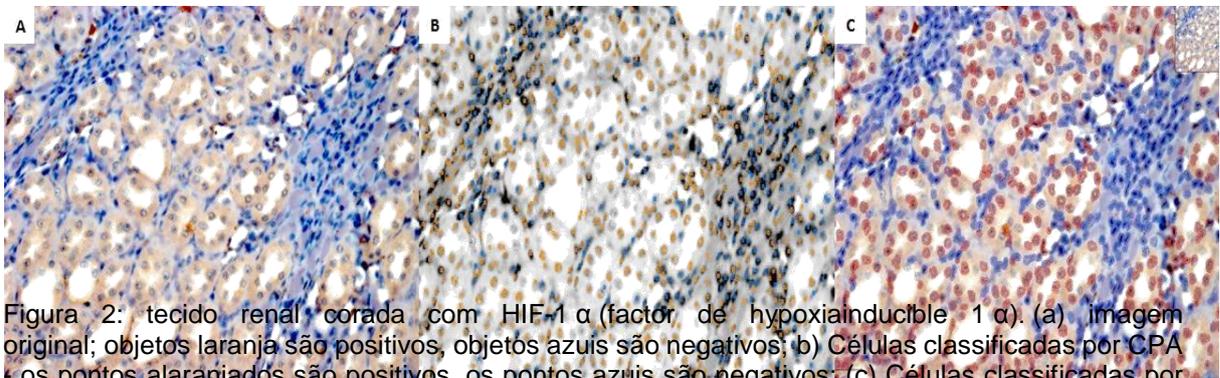


Figura 2: tecido renal corado com HIF-1 α (factor de hypoxia-inducible 1 α). (a) imagem original; objetos laranja são positivos, objetos azuis são negativos; b) Células classificadas por CPA - os pontos alaranjados são positivos, os pontos azuis são negativos; (c) Células classificadas por QP, pontos azuis são positivos, pontos vermelhos são negativos.

A fim de comparar as contagens de HIF e TUNEL com CP e QP, foi realizada uma comparação de Bland-Altman (Bland e Altman , 2007) (Figura 3 e 4).



Figura 3: Gráfico de dispersão de Bland-Altman para o HIF-1 α , a comparação das contagens de células em tecido renal corado com HIF. O eixo x representa a média das duas análises; o eixo y representa a diferença entre as duas análises. A linha média do eixo y representa a média das diferenças (140,3). As linhas superior e inferior são a média $\pm 1,96 * \text{desvio padrão}$ das diferenças.

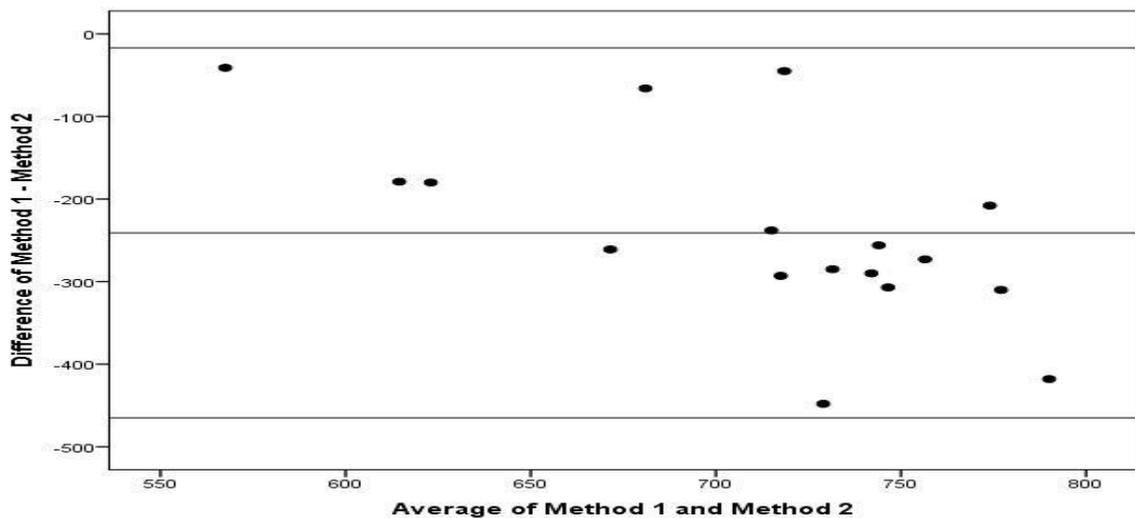


Figura 4: Gráfico de dispersão Bland-Altman para TUNEL: O eixo x representa a média das duas análises; o eixo y representa a diferença entre as duas análises. A linha média do eixo y representa a média das diferenças (-241,06). As linhas superior e inferior são a média $\pm 1,96 * \text{desvio padrão}$ das diferenças.

Como é visível nas figuras 3 e 4, todas as diferenças estão dentro da média $\pm 1,96 * \text{desvio padrão}$, ou seja, as diferenças são normalmente distribuídas, ou seja, os resultados dos dois programas são comparáveis.

Os coeficientes de correlação de concordância para o HIF -1 α foi de 0,717, ou seja, a correlação é significativa no nível 0,01 (2 extremidades-0,006). E para TUNEL, foi 0,063, ou seja, a correlação é significativa no nível 0,05 (2 extremidades-0,010).

4. DISCUSSÃO

O objetivo do presente estudo foi comparar dois programas livres, CP e QP, para realizar análise de imagem e subsequente classificação baseada na aprendizagem da máquina (Carpenter et al., 2006; Bankhead et al., 2017).

A comparação não foi fácil de conduzir porque os desenvolvedores dos dois programas tinham objetivos diferentes. Os desenvolvedores do CP afirmam que o software foi desenvolvido para imagens biológicas (Lamprecht, Sabatini e Carpenter, 2007; Vokes e Carpenter, 2008). Os desenvolvedores do QP afirmaram que seu software foi desenvolvido para a análise de imagens de slides inteiros (Paulik et al., 2017).

Por outro lado, uma comparação entre esses dois programas pode ser importante, já que, do melhor conhecimento dos autores, esses são os únicos dois programas livres nos quais os pesquisadores podem conduzir tanto a análise de imagens quanto a aprendizagem de máquina, bem como a classificação subsequente.

Mais uma razão para comparar os dois programas é a falta de literatura - e, portanto, o conhecimento ausente - sobre o software QP. Na biblioteca PubMed, existem quatro estudos revisados por pares sobre QP e mais de 70 em CP (a partir de maio de 2018). Para o QP, não há manual como existe para o CP. Para ambos os programas, no entanto, existe um fórum online.

Para o presente estudo, foram analisadas amostras de diferentes origens, nomeadamente, células inteiras, secções de tecidos e imagens de lâminas inteiras. Tanto quanto é do conhecimento dos autores, este é o primeiro estudo a combinar amostras de diferentes origens para a comparação de CP e QP. Outro estudo comparando CP e QP usou apenas um tipo de tecido, câncer de mama (Paulik et al., 2017). Os autores do último estudo também apresentaram resultados comparáveis.

Se houver uma seção de tecido com uma diferença visual óbvia entre células de classes diferentes, o QP deve ser a escolha para análise. A

identificação de células sem a necessidade de processar as imagens e o processo de aprendizado de máquina não supervisionado e os processos de classificação subsequentes permitem um fluxo de trabalho rápido. Por outro lado, se a origem da amostra for um tipo de célula única ou um tecido sem diferenças visuais óbvias entre diferentes classes de células, recomenda-se o CP.

CP e contagem manual são comparáveis (Macedo et al., 2017). Assim como QP e contagem manual também são comparáveis, pode-se concluir que as contagens de células de QP e CP também são comparáveis.

As limitações do presente estudo podem resultar da falta de uma contagem manual e da falta de literatura revisada por pares sobre o software QP. Mais estudos são necessários para verificar os resultados do presente estudo e aumentar o conhecimento e os limites do QP.

5. CONCLUSÕES

A análise quantitativa de imagens requer conhecimento dos programas de software disponíveis e sua aplicabilidade a tipos específicos de imagens, porque a escolha do software depende inteiramente dos tipos de amostras que serão analisadas. Para as amostras analisadas (tecido renal corado com HIF e TUNEL), o uso de QP foi mais fácil, pois oferece análise de imagem sem processamento prévio das imagens (por exemplo, conversão em escala de cinza, intensidades invertidas) e um processo de aprendizado não supervisionado. Mais estudos são necessários para obter mais conhecimento sobre o software QP, especialmente sobre sua capacidade de analisar imagens de leucócitos e culturas de células.

6. REFERÊNCIAS

- Bankhead P, Loughrey MB, Fernández JA, Dombrowski Y, McArt DG, Dunne PD, McQuaid S, Gray RT, Murray LJ, Coleman HG, James JA, Salto-Tellez M, Hamilton PW. QuPath: Open source software for digital pathology image analysis. *Sci Rep*. 2017; 7(1): 16878.
- Bankhead P, Fernández JA, McArt DG, Boyle DP, Li G, Loughrey MB, Irwin GW, Harkin DP, James JA, McQuaid S, Salto-Tellez M, Hamilton PW. Integrated tumor identification and automated scoring minimizes pathologist involvement and provides new insights to key biomarkers in breast cancer. 2018; 98(1):15-26.
- Bland JM, Altman DG. Agreement between methods of measurement with multiple observations per individual. *J Biopharm Stat* 2007;17:571– 82.
- Buzin AR, Pinto FE, Nieschke K, Mittag A, de Andrade TU, Endringer DC, Tarnok A, Lenz D. Replacement of specific markers for apoptosis and necrosis by nuclear morphology for affordable cytometry. *J Immunol Methods*. 2015; 420: 24-30.
- Carpenter AE, Jones TR, Lamprecht MR, Clarke C, Kang IH, Friman O, Guertin DA, Chang JH, Lindquist RA, Moffat J, Golland P, Sabatini DM. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol*. 2006; 7 (10): R100.
- Deo RC. Machinelearning in medicine. *Circulation*. 2015; 132(20):1920–1930.
- Gerstner AO, Lenz D, Laffers W, Hoffman RA, Steinbrecher M, Bootz F, Tárnok A. Near-infrared dyes for six-color immunophenotyping by laser scanning cytometry. *Cytometry*. 2002; 1; 48(3):115-23.
- Gerstner AO, Mittag A, Laffers W, Dähnert I, Lenz D, Bootz F, Bocsi J, Tárnok A. Comparison of immunophenotyping by slide-based cytometry and by flow cytometry. *J Immunol Methods*. 2006; 311(1-2): 130-8.
- Givan AL. Flow cytometry: an introduction. *Methods Mol Biol*. 2011; 699: 1-29.
- Jones TR, Kang IH, Wheeler DB, Lindquist RA, Papallo A, Sabatini DM, Golland P, Carpenter AE. CellProfiler Analyst: data exploration and analysis software for complex image-based screens. *BMC Bioinformatics*. 2008; 9: 482.

Jordan MI, Mitchell TM. Machinelearning: Trends, perspectives, and prospects. *Science*. 2015; 349:255–260.

Kamentsky LA, Kamentsky LD. Microscope-based multiparameter laser scanning cytometer yielding data comparable to flow cytometry data. *Cytometry*. 1991; 12 (5): 381-7

Laffers W, Mittag A, Lenz D, Tárnok A, Gerstner AO. Iterative restaining as a pivotal tool for n-color immunophenotyping by slide-based cytometry. *Cytometry A*. 2006 Mar; 69(3):127-30.

Lamprecht MR, Sabatini DM, Carpenter AE. CellProfiler: free, versatile software for automated biological image analysis. *Biotechniques*. 2007; 42(1): 71-5.

Lenz D, Barten MJ, Hiller S, Tárnok A, Sack U. Regenerative and predictive medicine of cardiovascular disease: the 9th Leipziger Workshop and the 2nd International Workshop on slide based cytometry. *Cytometry A*. 2005 Apr; 64(2):110-4.

Lenz D, Lenk K, Mittag A, Adams V, Kränkel N, Boldt A, Gerstner AO, Raida M, Weiss T, Hambrecht R, Tarnok A. Detection and quantification of endothelial progenitor cells by flow and laser scanning cytometry. *J BiolRegulHomeostAgents*. 2005b; 19 (3-4): 180-7.

Macedo ND, Buzin AR, de Araujo IB, Nogueira BV, de Andrade TU, Endringer DC, Lenz D. Objective detection of apoptosis in rat renal tissue sections using light microscopy and free image analysis software with subsequent machine learning: Detection of apoptosis in renal tissue. 2017; 49(1):22-27.

Mittag A., Lenz D., Gerstner A. O. H., Sack U., Steinbrecher M., Kokschi M., Raffael A., Bócsi J., Tárnok A. Polychromatic (eight-color) slide-based cytometry for the phenotyping of leukocyte, NK, and NKT subsets. *Cytometry A*. 2005; 65A (2), 103–115.

Mittag A, Pinto FE, Endringer DC, Tarnok A, Lenz D. Cellular analysis by open-source software for affordable cytometry. *Scanning*. 2011; 33 (1): 33-40.

Mosch B, Mittag A, Lenz D, Arendt T, Tárnok A. Laser scanning cytometry in human brain slices. *Cytometry A*. 2006; 69(3):135-8.

Paulik R, Micsik T, Kiszler G, Kaszál P, Székely J, Paulik N, Várhalmi E, Prémusz V, Krenács T, Molnár B. An optimized image analysis algorithm for detecting nuclear signals in digital whole slides for histopathology. *Cytometry A*. 2017 Jun; 91(6):595-608.

Ruifrok AC, Johnston DA. Quantification of histochemical staining by color deconvolution. *Quant Anal Cytol Histol*. 2001; 23: 291 – 299.

Schneider C, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* 2012; 9(7). 671-5

Vokes MS, Carpenter AE. Using CellProfiler for automatic identification and measurement of biological objects in images. *CurrProtoc Mol Biol*. 2008; Chapter 14: Unit 14.17.

*Normas da revista: Tissue and Cell.

7. ANEXOS

I) Declaração Ética: Protocolo nº 050/2013



II) Artigo Publicado, 25/03/2019. Ref: TISSUEANDCELL_2019_19_R1
Title: COMPARISON BETWEEN TWO PROGRAMS FOR IMAGE ANALYSIS, MACHINE LEARNING AND SUBSEQUENT CLASSIFICATION. Journal: Tissue and Cell. DOI: <https://doi.org/10.1016/j.tice.2019.03.002>.