

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

**CITOMETRIA DE IMAGEM: UM MÉTODO ACESSÍVEL PARA
REALIZAR DIAGNÓSTICO CELULAR E CONTROLE DE
QUALIDADE CONFIÁVEIS**

GABRIELLY RIBEIRO OLIVEIRA

VILA VELHA-ES
JULHO/2023

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

**CITOMETRIA DE IMAGEM: UM MÉTODO ACESSÍVEL PARA
REALIZAR DIAGNÓSTICO CELULAR E CONTROLE DE
QUALIDADE CONFIÁVEIS**

Tese de doutorado apresentada à Universidade Vila Velha como pré-requisito do Programa de Pós-graduação em Ciências Farmacêuticas, para obtenção do grau de Doutora em Ciências Farmacêuticas.

GABRIELLY RIBEIRO OLIVEIRA

VILA VELHA-ES
JULHO/2023

Catálogo na publicação elaborada pela Biblioteca Central / UVV-ES

O48c

Oliveira, Gabrielly Ribeiro.

Citometria de imagem : um método acessível para realizar diagnóstico celular e controle de qualidade confiáveis / Gabrielly Ribeiro Oliveira. – 2023.

42f.: il.

Orientador: Bianca Prandi Campagnaro

Tese (Doutorado em Ciências Farmacêuticas) -
Universidade Vila Velha, 2023.

Inclui bibliografias.

1. Farmacologia e terapêutica. 2. Estresse oxidativo.
3. Ferimentos e lesões - Tratamento. I. Campagnaro, Bianca Prandi. II. Universidade Vila Velha. III. Título.

CDD 615

GABRIELLY RIBEIRO OLIVEIRA

**CITOMETRIA DE IMAGEM: UM MÉTODO ACESSÍVEL PARA
REALIZAR DIAGNÓSTICO CELULAR E CONTROLE DE
QUALIDADE CONFIÁVEIS**

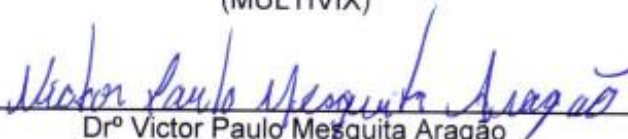
Tese apresentada à Universidade Vila Velha
como pré-requisito do Programa de Pós-
graduação em Ciências Farmacêuticas, para
obtenção do grau de Doutora em Ciências
Farmacêuticas.

Aprovada em 28 de julho de 2023,

Banca Examinadora:



Dr^a Karla Oliveira dos Santos Cassaro
(MULTIVIX)



Dr^o Victor Paulo Mesquita Aragão
(UVV - Universidade Vila Velha)



Dr^a Elisângela Flávia Pimentel Schmitt
(UVV - Universidade Vila Velha)



Dr^a Bianca Prandi Campagnaro
(UVV - Universidade Vila Velha)
Orientadora

Dedico este trabalho ao meu grande amigo e orientador Dominik Lenz.

AGRADECIMENTOS

Agradeço a Deus por todas as bênçãos derramadas sobre mim, Ele é meu refúgio.

Ao meu orientador e amigo Dr. Dominik Lenz. Com você aprendi ser uma cientista, uma profissional. Obrigada por me tornar uma campeã. Deus me apresentar você foi um grande marco para minha carreira profissional. Thanks for all. We are the champions, my friend, and we'll keep on fighting till the end.

À minha mãe, uma eterna gratidão, Nair Maria Pereira, ela que nunca mediu esforços para me ver estudando, que sempre sonhou com uma filha doutora e que lutou como uma verdadeira guerreira para tornar este sonho, o nosso sonho! Obrigada também ao meu pai Osvaldo, por segurar a minha mão e ter compreensão por tudo que minha mãe renunciou ao lar deles para me sustentar até aqui. Obrigada Tiago, meu irmão querido, você é meu anjo nesta terra, grata sou por todo carinho e amor que tem por mim, obrigado por tudo. Eu amo muito vocês, minha base.

Ao meu querido esposo Rodrigo Oliveira, obrigada por ser tão especial na minha vida, obrigada por estar comigo em todos os momentos difíceis, momentos estes que várias vezes me deixavam abalada ao ponto de não ser a melhor esposa que eu poderia ser naqueles dias, grata sou a você por sempre me motivar a fazer mais, te amo. À nossa filha Sarah que me deu um novo sentido para continuar mais forte que antes.

À Renata Mazuco, obrigada amiga, com você minha fé foi fortalecida. Que Deus te ilumine por onde for. À Ada Kerr, Walber Kerr, Dejanira Monteiro, Carlos Musso, Laíz Soares, Daniela Souza, Andréia Araújo, Jessica Oliveira, muito obrigada por todo apoio e palavras de ânimo. Que Deus os abençoe.

Aos queridos professores Marcio Fronza e Denise Endringer, quero agradecer com todo meu coração pelas oportunidades, pelo apoio com as dificuldades e por serem uma inspiração durante estes 10 anos que estive aqui nesta instituição.

À querida professora Bianca por aceitar me ajudar nesta reta final, obrigada por toda ajuda, orientação e carinho.

À CAPES:PROSUP, pela concessão da bolsa de doutorado, e a Universidade Vila Velha pela oportunidade de crescimento profissional.

RESUMO

Gabrielly Pereira Ribeiro; Dr; Universidade Vila Velha - ES; Julho 2023; **Citometria De Imagem: um método acessível para realizar diagnóstico celular e controle de qualidade confiáveis.** Orientadora: Bianca Prandi Campagnaro.

A prática de análise de células está inovando constantemente, migrando para o uso de mídias com visualização das lâminas em formato digital. A análise quantitativa automatizada de objetos de imagem é atualmente considerada uma ótima opção porque a comunidade científica comprova que a interação homem-máquina tem o potencial de reduzir o esforço humano e aumentar a confiança na tomada de decisão algorítmica. Aprendizagem de máquina é definido como um método que permite que computadores aprendam com padrões de dados para analisar e classificar fazendo previsões assertivas e melhorar continuamente os resultados. A análise manual de imagens tem algumas limitações quando se trata de análise em massa, enquanto os slides de imagens digitais são interativos, fáceis de compartilhar, envolvem menos tempo de preparação e são reproduzíveis. Diante do crescente aumento de números de amostras biológicas, se faz necessário o uso de uma metodologia automatizada para acelerar o processo e conduzir mais análises em menos tempo. O objetivo do presente estudo foi aplicar esta metodologia automatizada em diferentes áreas da pesquisa científica, demonstrando a versatilidade e desempenho com resultados confiáveis em um formato acessível para ser realizado em qualquer laboratório, uma vez que para realizar a análise é preciso somente microscopia de luz, câmera digital e um computador.

Palavras-chaves: *CellProfiler; Citometria de Imagem; Análise Celular Automatizada.*

ABSTRACT

Gabrielly Pereira Ribeiro; PhD; Vila Velha University - ES; July 2023; **Imaging Cytometry: an inexpensive method to perform reliable cell diagnosis and quality control.** Advisor: Bianca Prandi Campagnaro.

The practice of cell analysis is constantly innovating, migrating to the use of media with slide visualization in digital format. Automated quantitative analysis of image objects is currently considered a successful option because the scientific community proves that human-machine interaction has the potential to reduce human effort and increase confidence in algorithmic decision making. Machine learning is defined as a method that allows computers to learn from data patterns to analyze and classify, making assertive predictions and continuously improving results. Manual image analysis has some limitations when it comes to mass analysis, while digital image slides are interactive, easy to share, involve less preparation time and are reproducible. In view of the increasing number of biological samples, it is necessary to use an automated methodology to speed up the process and conduct more analyzes in less time. The objective of the present study was to apply this automated methodology in different areas of scientific research, demonstrating its versatility and performance with reliable results in an accessible format to be carried out in any laboratory, since to perform the analysis only light microscopy is needed, digital camera and a computer.

Keywords: CellProfiler; Image Cytometry; Automated Cell Analysis.

SUMÁRIO

| | |
|--|-----------|
| LISTA DE FIGURAS..... | 09 |
| 1. INTRODUÇÃO..... | 10 |
| 2. REVISÃO BIBLIOGRÁFICA..... | 11 |
| 2.1 Breve Histórico: Análise de Imagens..... | 11 |
| 2.2 Citometria de Imagem..... | 12 |
| 2.3 Inteligência Artificial: Aprendizado de Máquina..... | 12 |
| 2.4 Hipótese..... | 13 |
| 2.5 Referências Bibliográficas..... | 13 |
| 3. OBJETIVOS..... | 17 |
| 3.1 Objetivo Geral..... | 17 |
| 3.2 Objetivos Específicos..... | 17 |
| 4. CAPÍTULO 1 - Artigo 1. An inexpensive, automated and reproducible method to conduct quality control in nanoparticles..... | 18 |
| Abstract..... | 19 |
| Introduction..... | 20 |
| Sample Preparation..... | 21 |
| Materials..... | 21 |
| Results and discussion..... | 23 |
| Conclusions..... | 25 |
| References..... | 26 |
| 5. CAPÍTULO 2 - Artigo 2. Automated cytopathology of fine needle aspiration for the detection of malignancy in thyroid cells..... | 28 |
| Abstract..... | 29 |
| Introduction..... | 30 |
| Materials and methods..... | 30 |
| Results..... | 33 |
| Discussion..... | 37 |
| References..... | 38 |
| 6. CONSIDERAÇÕES FINAIS..... | 40 |

LISTA DE FIGURAS

I. CAPÍTULO 1

Figure 1: Original image and image after to count.

Figure 2: Original image, referring to module 1 of table 1 (a); initial identification of the objects of interest, referring to module 3 of table 1 (b); Remaining objects after filtering the objects, referring to module 5 of table 1 (c).

Figure 3: The boxplot represents the variability of the manual and automatic data in the study. The central line represents the median and the top and bottom lines are the data range (maximum - minimum). The spaces between the different parts of the box indicate the degree of dispersion.

II. CAPÍTULO 2

Figure 1: Pipeline: First all the images were loaded, identified by the unique format known in all .tif, the intensities were inverted for the objects of interest asking for light on a dark background, later the primary object was identified as nuclei 1, after filtering the objects (eliminating objects unexpected) a new count of the parameters was performed and then data nuclei 2 was found, only these were exported to carry out the classification.

Figure 2: Analysis performed in Cellprofiler of a malignant sample. A: Original Image, B: Color inversion to grayscale, C: Nuclei identification D: Filter applied to exclude the artifacts.

Figure 3: Analysis performed in Cellprofiler of a benignant sample. A: Original Image, B: Color inversion to grayscale, C: Nuclei identification D: Filter applied to exclude the artifacts.

Figure 4: ROC curve between automated and manual results.

Figure 5: Classification obtained in the TANAGRA software (Scatter plot: red-benign / green-malignant).

Figure 6: Linear discriminant analysis between benignant and malignant objects.

Figure 7: Training with two classes, malignant (lower right) and benign (lower left) cells.

Figure 8: 100 benignant correct objects requested after training.

Figure 9: 100 malignant correct objects requested after training.

Figure 10: Score Image: This tool allows the analyst to see the full extent of the image and the objects there identified according to the classification performed under each class after training. A: Malignant objects identified B: Benignant objects identified. - 97.08% of accuracy.

1. INTRODUÇÃO

A citometria de imagem é uma técnica de análise quantitativa de imagens que permite extrair dados numéricos de objetos previamente identificados como objetos de interesse (por exemplo, núcleos ou células). Para realizar essa análise, é necessário um *software* capaz de gerar parâmetros quantitativos. Quanto maior o número de parâmetros disponíveis, melhores serão os resultados, pois isso permite uma análise multiparamétrica.

Em 1997, foi lançado o primeiro *software* livre chamado *Image J*, conhecido por realizar análises de imagens que incluíam parâmetros morfológicos [1]. Em 2006, o *software CellProfiler* também foi lançado, tornando-se outra plataforma fundamental para a análise de imagens. Assim como o *Image J*, o *CellProfiler* está disponível gratuitamente. O desenvolvimento do *CellProfiler* ocorreu devido à necessidade de realizar análises de imagens com múltiplos parâmetros para cada objeto de interesse [2].

A popularidade do *Image J* e do *CellProfiler* é evidente na literatura científica recente. Até julho de 2023, mais de 20.000 artigos revisados por pares foram publicados usando o *ImageJ*. O principal recurso do *ImageJ* é o processamento e a investigação de imagens individuais, enquanto o *CellProfiler* foi projetado para construir pipelines de análises em grande escala. Esses *softwares* são aplicáveis em diversas áreas da ciência como, biologia, engenharia e medicina [3-5].

Devido à necessidade dos laboratórios de conduzir análises em grande escala, os métodos de análise manual têm se tornado obsoletos. Isso ocorre porque eles estão sujeitos a erros humanos e são demorados, exigindo muito trabalho e tempo. Felizmente essas medições podem ser realizadas facilmente por *softwares* específicos, sem a necessidade de conhecimento sobre algoritmos e aprendizado de máquina, permitindo que qualquer pessoa interessada em análise de imagens possa utilizá-los [6-9].

Em síntese, a análise automatizada de imagens gera conjuntos de dados quantitativos reproduzíveis e possui vantagens em relação à análise manual, reduzindo significativamente o viés humano, e tornando-se o método preferencial para análise de imagens [10-12].

2. REVISÃO BIBLIOGRÁFICA

2.1 Breve histórico: Análise de Imagens

Na patologia a análise de imagens é amplamente utilizada para obter dados precisos por meio da identificação e contagem de objetos, a fim de obter um diagnóstico. A mensuração de objetos microscópicos remonta aos primórdios do próprio microscópio, inventado por Hans e Zacharias Janssen por volta de 1590 [13]. No século XVII, Antonie van Leeuwenhoek desenvolveu um método para medir objetos microscópicos usando grãos de areia e fios de cabelos de diferentes tamanhos como referência [14,15].

A descoberta das células sanguíneas logo após a invenção do microscópio despertou o interesse dos pesquisadores em determinar a quantidade dessas células [16]. Em 1873 o pesquisador Louis Charles desenvolveu um método quantitativo, o hemocitômetro, também chamado de câmara de Neubauer [17]. Essa descoberta impulsionou uma série de estudos e avanços, que, ao serem combinados e aprimoradas ao longo dos anos, resultaram no desenvolvimento do que conhecemos hoje como Citometria de Fluxo. Esse método avalia células suspensas em meio líquido em fluxo no espectrofotômetro, permitindo a análise simultânea de várias propriedades espectrofotométricas de células com taxas superiores a 500 células por segundo [18].

Novas metodologias para análises quantitativas [19-21] continuaram a surgir, porém ainda apresentavam algumas limitações, como falta de objetividade e reprodutibilidade, impossibilidade de coletar amostras de recém-nascidos ou animais devido à necessidade de maior volume sanguíneo e do uso de reagentes, entre outras [22]. Para superar essas limitações foi desenvolvido o Citômetro de Varredura a Laser, que utiliza a emissão de fluorescência excitada por um laser em células individuais para medir as células em várias faixas de comprimento de onda (até 5000 células/minutos, ou seja, 83 células/segundos) [23-25]. Além disso, essa técnica requer uma menor quantidade de amostras e de reagentes [23-25]. Essas são as vantagens dessa técnica em relação à citometria de fluxo descrita anteriormente, porém, ainda está limitada a imagens de fluorescência e à quantidade de células analisadas por segundo é menor em comparação com a Citometria de Fluxo, sendo que esta última ainda lidera as análises em laboratórios.

A análise de imagens é de grande importância no campo científico, pois fornece informações detalhadas para o diagnóstico e prognóstico de doenças, especialmente por meio da análise morfológica de células e seus componentes [26].

A morfometria nuclear, por exemplo, estuda parâmetros do núcleo como área, forma, intensidade entre outros. A visualização dessas características em células malignas é fundamental para um diagnóstico reprodutível, uma vez que as informações são diferentes daquelas observadas nas não-malignas. Com base nessas diferenças, alguns autores propuseram que a morfometria nuclear para aprimorar os parâmetros diagnósticos e fornecer fatores prognósticos confiáveis [27,28].

2.2 Citometria de Imagem

A Citometria de Imagem é uma técnica de análise quantitativa automatizada que surgiu no final dos anos 90. Diferentemente das técnicas anteriores, essa abordagem permite a análise de parâmetros morfológicos que não se limitam apenas à fluorescência [29]. Nas últimas décadas, houve um desenvolvimento significativo de algoritmos personalizados para identificação e classificação de objetos e esses foram inseridos dentro de *softwares* para realizar o aprendizado de máquina. Esta metodologia é considerada reprodutível [30-32] e de baixo custo [24,33,34]. Os softwares gratuitos mais amplamente utilizados nessa área são o Image J, pioneiro na citometria de imagem, o *Cellprofiler* e o *QuPath* [1,2,35].

Na Tabela 1 abaixo estão as principais diferenças que ocorreram após a descoberta da Citometria de Imagem.

| ANÁLISE MANUAL | CITOMETRIA |
|-------------------------|-------------------------|
| MICROSCOPIA - MANUAL | SOFTWARE – AUTOMATIZADA |
| QUALITATIVA | QUANTITATIVA |
| BAIXA REPRODUTIBILIDADE | ALTA REPRODUTIBILIDADE |
| PARÂMETROS LIMITADOS | MULTIPARAMÉTRICA |

Tabela 1: Diferenças entre a Análise manual e Citometria de Imagem.

Conforme apontado por Buskermolen e colaboradores [26], a análise de imagens patológicas por meio da quantificação manual resulta em uma detecção insuficiente e não é capaz de descrever mudanças sutis que são relevantes para o diagnóstico.

2.3 Inteligência Artificial: Aprendizado de máquina

A Inteligência Artificial é uma ferramenta da tecnologia moderna que permite o desenvolvimento de *softwares* capazes de realizar diagnósticos clínicos e tomar decisões terapêuticas, ou seja, é a transmissão da inteligência humana para a máquina. Na área da medicina, diversas metodologias de aprendizado de máquina têm sido aplicadas com sucesso [36-41].

A análise de imagens é um dos campos em que a Inteligência Artificial tem apresentado avanços significativos. O aprendizado de máquina, que é parte integrante da Inteligência Artificial, utiliza dados como recurso de entrada, permitindo que a máquina aprenda com esses dados [42]. Por meio do uso de funções matemáticas pré-definidas no *software*, é possível obter resultados (classificação ou regressão) que frequentemente são difíceis de serem alcançados rapidamente pelos seres humanos.

2.4 Hipótese

Com base nos conhecimentos apresentados, a hipótese do presente estudo é que a realização de uma análise automatizada utilizando um software gratuito resulta em menor custo, alta precisão e pode ser reproduzida em qualquer laboratório.

2.5 Referências Bibliográficas

- [1] Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*. 2012;Jul;9(7):671-5.
- [2] 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.
- [3] Collette, R., Douglas, J., Patterson, L., Bahun, G., King, J., Keiser Jr, D., & Schulthess, J. Benefits of utilizing CellProfiler as a characterization tool for U-10Mo nuclear fuel. *Materials Characterization*. 2015; 105, 71-81.
- [4] Miller, K., Eggenberger, A. L., Lee, K., Liu, F., Kang, M., Drent, M., ... & Jiang, S. An improved biolistic delivery and analysis method for evaluation of DNA and CRISPR-Cas delivery efficacy in plant tissue. *Scientific Reports*. 2021; 11(1), 7695.
- [5] Sangi S, Santos MLC, Alexandrino CR, Da Cunha M, Coelho FS, Ribeiro GP, Lenz D, Ballesteros H, Hemerly AS, Venâncio TM, Oliveira AEA, Grativol C. Cell wall dynamics and gene expression on soybean embryonic axes during germination. *Planta*. 2019 Oct;250(4):1325-1337.
- [6] Ribeiro GP, Endringer DC, De Andrade TU, Lenz D. Comparison between two programs for image analysis, machine learning and subsequent classification. *Tissue Cell*. 2019 Jun;58:12-16.
- [7] Buzin AR, Macedo ND, De Araujo IB, Nogueira BV, de Andrade TU, Endringer DC,

Lenz D Automatic detection of hypoxia in renal tissue stained with HIF-1alpha. *J Immunol Methods*. 2017; 444:47-50.

[8] Macedo ND, Buzin AR, Araujo IB, Nogueira B, Andrade TU, Endringer DC, Lenz D. Automated and Reproducible Detection of Vascular Endothelial Growth Factor (VEGF) in Renal Tissue Sections. *J Immunol Res*. 2019; 2019:7232781.

[9] Chen PC, Gadepalli K, MacDonald R, Liu Y, Kadowaki S, Nagpal K, Kohlberger T, Dean J, Corrado GS, Hipp JD, Mermel CH, Stumpe MC. An augmented reality microscope with real-time artificial intelligence integration for cancer diagnosis. *Nat Med*. 2019; Sep;25(9):1453-1457.

[10] Deng C, Ji X, Rainey C, Zhang J, Lu W. Integrating Machine Learning with Human Knowledge. *iScience*. 2020;23(11):101656.

[11] Zou Z, Shi Z, Guo Y, Ye J. Object detection in 20 years: a survey. *ArXiv:1905.05055*. 2019;[online] Available: <https://arxiv.org/abs/1905.05055v2>.

[12] Norgeot B, Glicksberg BS, Butte AJ. A call for deep-learning healthcare. *Nat Med*. 2019;25(1):14-15.

[13] Davidson, M. Pioneers in Optics: Zacharias Janssen and Johannes Kepler. *Microscopy Today*, 2009; 17(6), 44-47.

[14] Van Leeuwenhoek A. *Arcana Natura Detecta*, Delphis, Batav, 1695.

[15] Meijer GA, Beliën JA, van Diest PJ, Baak JP. Origins of ... image analysis in clinical pathology. *J Clin Pathol*. 1997; May;50(5):365-70.

[16] Hajdu, S. I. The discovery of blood cells. *Annals of Clinical & Laboratory Science*. 2003;33(2), 237-238.

[17] Malassez L. *De la Numération des Globules Rouges du Sang*. Delahaye; Paris, France: 1873.

[18] Kamensky LA, Melamed MR, Derman H. Spectrophotometer: new instrument for ultrarapid cell analysis. *Science*. 1965; Oct 29;150(3696):630-1.

[19] Moldavan, A. Photo-Electric Technique For The Counting Of Microscopical Cells. *Science* 1934; 188-189.

[20] Van Dilla MA, Trujillo TT, Mullaney PF, Coulter JR. Cell microfluorometry: A method for rapid fluorescence measurement. *Science*. 1969; 163:1213–1214.

[21] Kohler G and Milstein C. Continuous cultures of fused cells secreting antibody of

predefined specificity. *Nature*. 1975; 256, 495.

[22] Pozarowski P, Holden E, Darzynkiewicz Z. Laser scanning cytometry: principles and applications-an update. *Methods Mol Biol*. 2013; 931:187-212.

[23] Darzynkiewicz Z, Bedner E, Li X, Gorczyca W, Melamed MR. Laser-scanning cytometry: A new instrumentation with many applications. *Exp Cell Res*. 1999; May 25;249(1):1-12.

[24] 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*. 2005a; Apr;64(2):110-4.

[25] 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, 180-87.

[26] Buskermolen ABC, Kurniawan NA, Bouten CVC. An automated quantitative analysis of cell, nucleus and focal adhesion morphology. *PLoS One*. 2018;13(3):e0195201.

[27] Millot, C.; Dufer, J. Clinical applications of image cytometry to human tumor analysis. *Histol Histopathol*, 2000; 15:1185–2000.

[28] Buzin AR, Pinto FE, Nieschke K, et al. Replacement of specific markers for apoptosis and necrosis by nuclear morphology for affordable cytometry. *J Immunol Methods*. 2015; 420:24-30.

[29] Myers G. Why bioimage informatics matters. *Nat Methods*. 2012; Jun 28;9(7):659-60.

[30] Mittag A., Lenz D., Gerstner A. O. H., Sack U., Steinbrecher M., Kokschi M., Raffael A., Bocsi 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.

[31] Rexhepaj E, Agnarsdóttir M, Bergman J, Edqvist PH, Bergqvist M, Uhlén M, Gallagher WM, Lundberg E, Ponten F. A texture based pattern recognition approach to distinguish melanoma from non-melanoma cells in histopathological tissue microarray sections. *PLoS One*. 2013 May 17;8(5):e62070.

[32] 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.

[33] Gerstner AO, Tárnok A. Analysis of fine-needle aspirate biopsies from solid tumors

by laser scanning cytometry (LSC). *Curr Protoc Cytom.* 2002;Chapter 7:Unit 7.

[34] Laffers W, Mittag A, Lenz D, et al. Iterative restaining as a pivotal tool for n-color immunophenotyping by slide-based cytometry. *Cytometry A.* 2006;69(3):127-130.

[35] 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; Dec 4;7(1):16878.

[36] Gulshan, V. et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA.* 2016; 316(22), 2402–2410.

[37] Karri, S. P., Chakraborty, D. & Chatterjee, J. Transfer learning based classification of optical coherence tomography images with diabetic macular edema and dry age-related macular degeneration. *Biomed. Opt. Express.* 2017; 8(2), 579–592.

[38] Hood, D. C. & De Moraes, C. G. Efficacy of a deep learning system for detecting glaucomatous optic neuropathy based on color fundus photographs. *Ophthalmology.* 2018; 125(8), 1207–1208.

[39] Kermany, D. S. et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell.* 2018; 172(5), 1122–1131.

[40] Mazo, C., Bernal, J., Trujillo, M. & Alegre, E. Transfer learning for classification of cardiovascular tissues in histological images. *Comput. Methods Prog. Biomed.* 2018; 165, 69–76.

[41] Tschandl, P. et al. Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: an open, web-based, international, diagnostic study. *Lancet Oncol.* 2019; 20(7), 938–947.

[42] Stafford I., Kellermann M., Mossotto E., Beattie R., MacArthur B., Ennis S. A systematic review of the applications of artificial intelligence and machine learning in autoimmune diseases. *NPJ Digit. Med.* 2020; 3:1–11.

3. OBJETIVOS

3.1 Objetivo Geral

- Aplicar uma metodologia automatizada em duas diferentes áreas da pesquisa científica, demonstrando sua versatilidade e desempenho na obtenção de resultados confiáveis de forma acessível.

3.2 Objetivos Específicos

- Analisar as imagens para a obtenção de resultados multiparamétrico.
- Analisar imagens feitas em microscopia de luz de amostras citopatológicas de células tireoidianas utilizando software livre e inteligência artificial.
- Analisar a forma e o tamanho de objetos de interesse (nanopartículas) com um software gratuito para análise de imagens e posteriormente comparar os resultados com os de uma análise manual.
- Demonstrar uma comparação estatística de alta precisão.

Capítulo 1

Artigo 1. An inexpensive, automated and reproducible method to conduct quality control in nanoparticles

**An inexpensive, automated and reproducible method to conduct quality control
in nanoparticles**

**Gabrielly Pereira Ribeiro¹, Rafaela Spessemille Valotto², Jairo Pinto de Oliveira³,
Marco Cesar Cunegundes Guimarães² and Dominik Lenz¹**

¹ *Post-Graduated Program of Pharmaceutical Sciences University, Vila Velha, ES,
Brazil*

² *Federal University of Espírito Santo, Vitória, ES, Brazil*

³ *Federal University of Espírito Santo, Alegre, ES, Brazil*

Present address: University, Vila Velha, ES, Brazil

Corresponding author, e-mail: dominik.lenz@gmail.com

Abstract

Multiple methods have been developed to estimate size and shape of nanoparticles from transmission electron microscopy images. However, some of them have limiting characteristics such as hard to use, expensive or have steps that make the results unreal. We present here a comparison of a manual (IMAGE J) and an automated analysis, both free methods. The gold nanoparticles were synthesized by the traditional Turkevich method using sodium citrate. We used 23 images of gold nanoparticles. The results of our study show that the conglomerated particles can be excluded - given that a sufficient amount of particles is included in the analysis. We show that CellProfiler can be highly efficient to assumptions about size and shape of nanoparticles. These findings provide a new approach of free software that can be applied as machine learning and therefore automate the processing of nanoparticles data.

Keywords: Nanoparticles; CellProfiler; transmission electron microscopy; automated analysis.

Introduction

Nanoparticles are defined as material between 1 and 100 nanometers in size with a mass of 10^5 atoms (Issa et al., 2013). Nanoparticles are known to have a wide range of diagnostic and therapeutic applications. As examples, vascular and tumor imaging, drug delivery and iron supplement may be mentioned (Arami et al., 2015). The strong interest in nanoparticles results from the unique physical properties such as the area of their surface and the surface tension (Paranjpe et al., 2014). The therapeutic effect of nanoparticles, however, is strongly dependent on their size (de Jong et al., 2008; Dwivedi et al., 2014). The importance of nanoparticles including quality control was mentioned in the recent literature (Kite et al., 2019; Pantil & Bhosale 2018; Pantil et al., 2019; Pantil et al., 2020). This frequent control of the synthesized nanoparticles is therefore a condition to maximize the therapeutic effect (Nischwitz et al., 2012; Vippola et al., 2016). The gold standard for quality control is the analysis of images created by transmission electron microscopy, scanning electron microscopy, scanning transmission electron microscopy, small-angle X-ray scattering, wide-angle X-ray scattering, atomic force microscopy, or dynamic light scattering (Vippola et al., 2016). The nature of the image analysis may be manual or automated. A manual analysis is tedious, time-consuming and, most important, not reproducible. An automated analysis on the other hand is reproducible, fast, and manages bigger sample sizes (Buzin et al., 2015; Buzin et al., 2017; Macedo et al., 2017). Last, not least, the automated analysis does not require a highly educated specialist, resulting in a more inexpensive analysis. A recent study stated the inability of common image analysis software to identify agglomerated nanoparticles. To this end, the authors developed a software capable to identify single nanoparticles also when agglomerated. The authors compared their results with Image J and a manual analysis (Vippola et al., 2016). The results of the three methods seemed comparable. Gold nanoparticles (AuNPs) have been researched for applications in medical images, according to recent research; they are considered image contrast agents that benefit clinical practice (Han et al., 2019).

The authors of the present study hypothesized, that agglomerated events might be excluded from an automated analysis, given, that a sufficient quantity of nanoparticles per sample is analyzed.

The aim of the present study is to analyze the shape and size of nanoparticles with a free software for image analysis and subsequently compare the results with those of a manual analysis.

Sample preparation

In Synthesis and characterization of gold nanoparticles

Gold nanoparticles (AuNPs) were synthesized based on (Turkevich et al., 1951) method with modifications. Briefly, 15 mL of Chloroauric acid (HAuCl₄, Merck Co. Ltd., Brazil) at 2.5×10^{-4} M was maintained under agitation (200 rpm). The gold precursor was reacted with 1 mL of trisodium citrate dihydrate at 1% m/v (Na₃C₆H₅O₇·2H₂O, Merck Co. Ltd., Brazil) for 15 minutes at 100°C. The solution was cooled in ice bath for 30 minutes. AuNPs were collected and their optical properties were evaluated by UV-vis spectroscopy (FEMTO 800 XI). Size and morphology were examined by transmission electron microscope (JEOL, JEM-1400) operated at 120 kV with Lab₆ filament. AuNPs were characterized using Image J and Cellprofiler.

Materials

Image J

The monodispersity of the AuNPs was determined based on the aspect ratio manually (major axis length and minor axis length) using the ImageJ software as shown in figure 1.

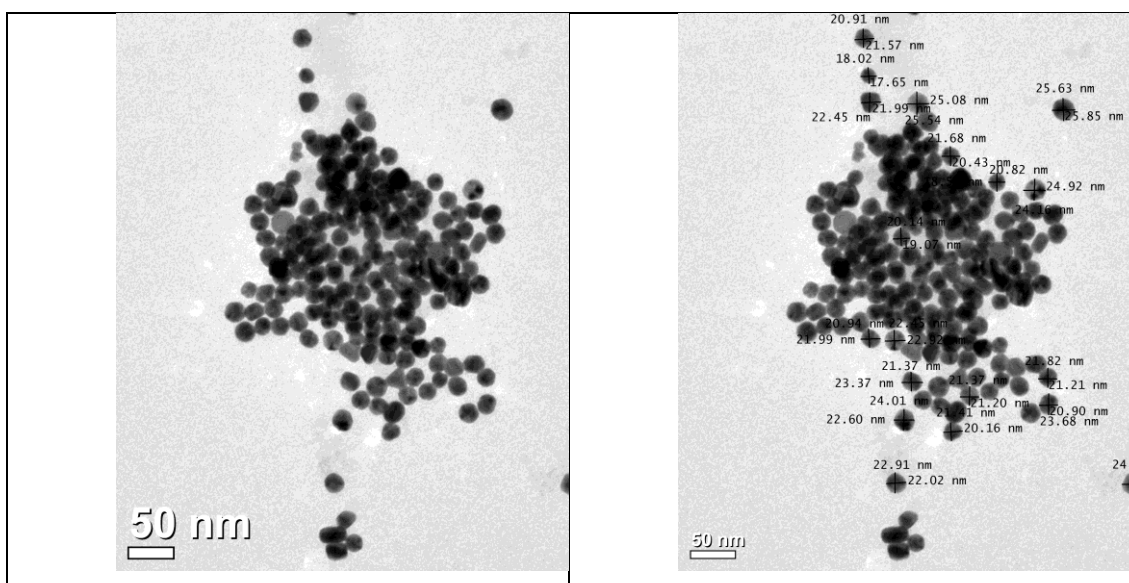


Figure 1. Original image and image after to count.

CellProfiler

The automated analysis was conducted using the free software CellProfiler (version 2.1.1), where 23 images ($N=423$ nanoparticles), each image was processed individually. The identified objects that did not reach a certain minimum in terms of a user-defined, image dependent form factor were not included in the subsequent analysis, as described in table 1 (5- Filter Objects):

Table 1. Algorithm used for the image analysis with CellProfiler.

| <i>Module</i> | <i>Operation</i> |
|-------------------------------------|--|
| 1- <i>Load Images</i> | Identify and load image by image Name of this loaded image: i1 |
| 2- <i>Image Math</i> | Operation: Invert Select the first image: i1 Name of the output image: i2 |
| 3- <i>Identify Primary Objects</i> | Identify an object of interest Name of the primary objects: spheres Threshold strategy: Adaptive |
| 4- <i>Measure Object Size Shape</i> | Measurement objects: spheres |
| 5- <i>Filter Objects</i> | Select the objects to filter: spheres Remaining objects: spheres2 |
| 6- <i>Measure Object Size Shape</i> | Measurement objects: spheres2 |

- 1) Upload images; 2) ImageMath: reverting the intensity of the images; 3) Identify of the objects of interest in the images. "Objects" is used as a generic term in CellProfiler to refer to a specific pattern identified in the image, in this case, the spheres. 4) Measure the features of the identified object; 5) Filter the objects to exclude all objects that do not fulfill certain limits – in this case a Form Factor < user-defined value; 6) Measure the features of the identified object: Major Axis Length and Minor Axis Length.

The original image, the initial identification and the result of the filtering are depicted in figure 2.

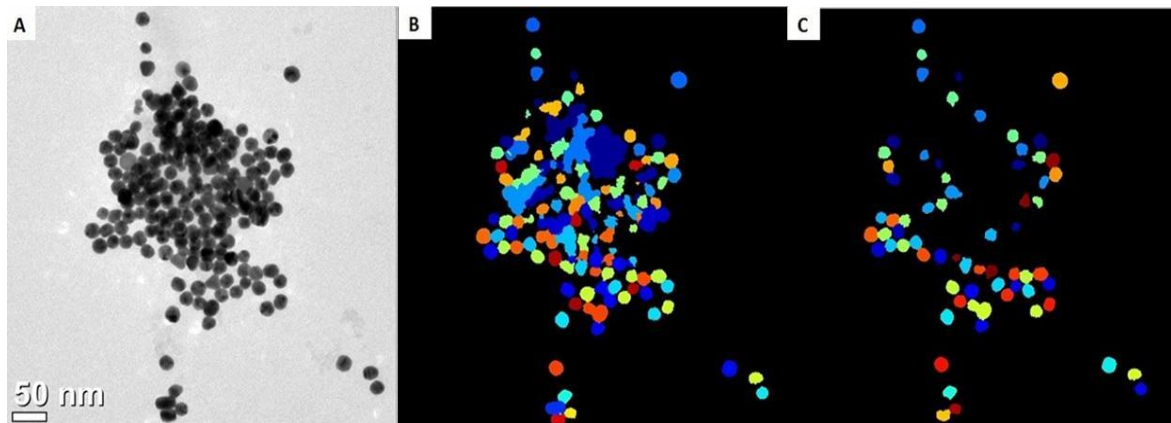


Figure 2. Original image, referring to module 1 of table 1 (a); initial identification of the objects of interest, referring to module 3 of table 1 (b); Remaining objects after filtering the objects, referring to module 5 of table 1 (c).

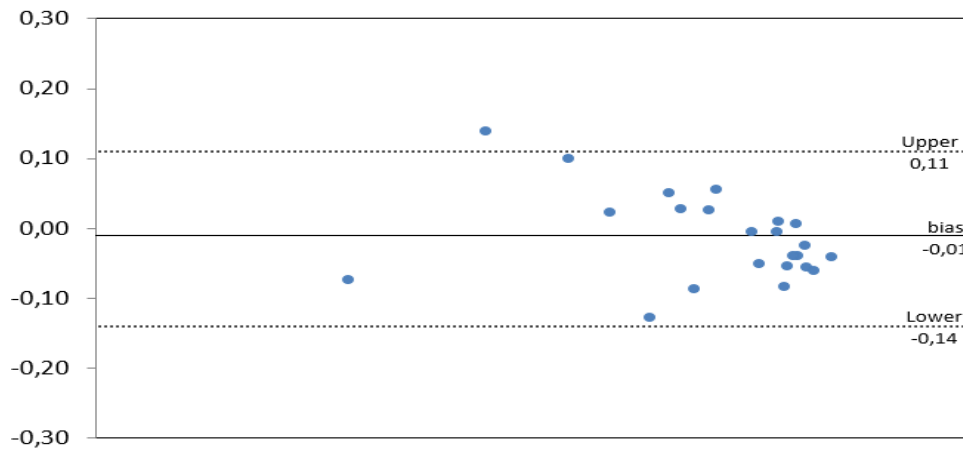
Results and discussion

A quality control, i.e., a morphological analysis of nanoparticles was conducted manually and automated. A mean of the values was described in table 2. Subsequently the results were compared. The results of the comparison are depicted (Table 3).

Table 2: Mean and standard deviation of manual and automated count.

| <i>IMAGES</i> | <i>IMAGE J</i> | <i>CELLPROFILER</i> |
|----------------------|-----------------------|----------------------------|
| <i>Mean</i> | 0,93 | 0,95 |
| <i>SD</i> | 0,03 | 0,04 |

Table 3. A Bland-Altman graph for the comparison of manual and automated analysis. The compared parameters were minor axis length divided by the major axis length. The x-axis represents the media of the two analyses; the y-axis represents the difference between the two analyses. The middle line of the y-axis represents the mean of the differences (-0.01). The upper and lower lines are mean + 1.96 * standard deviation of the differences.



The Bland-Altman assay show the mean of the differences is close to zero. There is no systematic bias, i.e., the differences are mean, above and below the mean. The higher the mean the smaller the differences between the two methods. Almost all the differences are within mean + 1.96 * standard deviation, i.e., the differences are normally distributed.

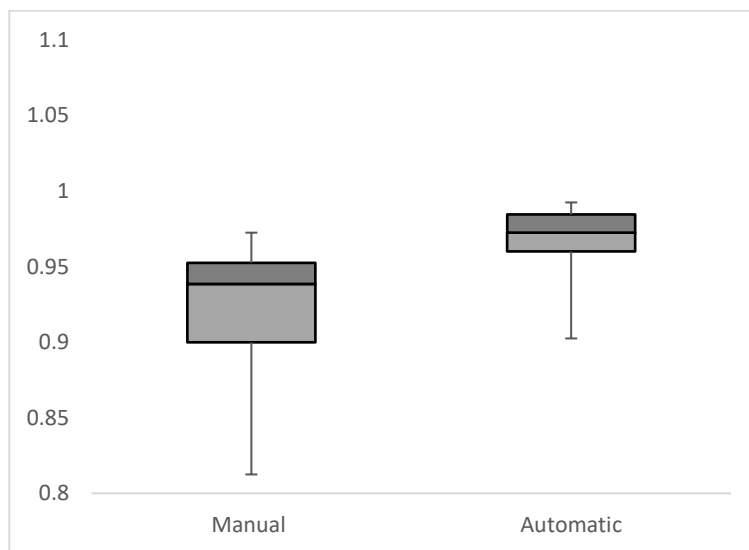


Figure 3: The boxplot represents the variability of the manual and automatic data in the study. The central line represents the median and the top and bottom lines are the data range (maximum - minimum). The spaces between the different parts of the box indicate the degree of dispersion.

The median of the manual method was 0.94 and in the automatic method, it was 0.97, as shown in the graph by the central line. The automatic method showed a more symmetrical distribution than the manual method. There is also a low variability and low standard deviation compared to the manual method.

The effect of nanoparticles is strongly dependent on their size (De Jong et al., 2008; Dwivedi et al., 2014). A frequent quality control of the synthesized nanoparticles is therefore a very important measure to maximize the effect of the objects (Nischwitz et al., 2012; Vippola et al., 2016).

For the present study a quality control, i.e. a morphological analysis of nanoparticles was conducted manually and automated. Subsequently the results were compared. The data of the present study indicate that the results of the manual and automated analysis are comparable. In a recent study, a software was developed, capable to conduct a quality control of nanoparticles including the identification and analysis of single objects even when conglomerated (Vippola et al., 2016). The authors had compared the results obtained with their new software with those of the software *image J* and a manual analysis. The results of the three different methods seemed comparable; however, no statistical analysis of the comparability of the results was conducted. As the authors of the present study hypothesized that conglomerated objects can be excluded, the results of the present study indicate that the results of a manual analysis, including conglomerated objects, is comparable to those of an automated analysis that excluded conglomerated objects.

Conclusions

The results of the present study indicate, that an automated analysis excluding conglomerated objects may deliver the same results as automated analysis including, given that the first method analyses a representative count of objects. Another conclusion of the present study is that an automated analysis is more accurate and faster as compared to a manual analysis. It may be concluded that there is no necessity to spend labor, money, and time to develop a software.

Further comparisons may be necessary to verify the truth of the author's hypothesis of the present study.

Acknowledgements. The financial support of CAPES (1740334) is greatly acknowledged.

References

- Arami H, Khandhar A, Liggitt D, Krishnan KM. *ChemSoc Rev.* (2015); 44: (23): 8576. DOI: 10.1039/c5cs00541h
- Buzin AR, Macedo ND, De Araujo IB, Nogueira BV, de Andrade TU, Endringer DC, Lenz D. *J ImmunolMethods.* (2017); 444: 47. DOI: 10.1016/j.jim.2017.02.005
- Buzin AR, Pinto FE, Nieschke K, Mittag A, de Andrade TU, Endringer DC, Tarnok A, Lenz D. *J ImmunolMethods.* (2015); 420: 24. DOI: 10.1016/j.jim.2015.03.011
- De Jong WH, Hagens WI, Krystek P, Burger MC, Sips AJ, Geertsma RE. *Biomaterials.* (2008) ;29 (12): 1912. DOI: 10.1016/j.biomaterials.2007.12.037
- Dwivedi MV, Harishchandra RK, Koshkina O, Maskos M, Galla HJ. *Biophys J.* (2014);106 (1): 289. DOI: 10.1016/j.bpj.2013.10.036
- Han X, Xu K, Taratula O, Farsad K. *Nanoscale* (2019); 11 (3): 799-819. DOI: 10.1039/C8NR07769J
- Issa B, Obaidat IM, Albiss BA, Haik Y. *Int J Mol Sci.* (2013); 14 (11): 21266. DOI:org/10.3390/ijms141121266
- Kite SV, Sathe DJ, Patil SS, Bhosale PN, Garadkar KM. *Mater. Res. Express* (2019) 6 026411. DOI: 10.1088/2053-1591/aaed81
- Macedo ND, Buzin AR, de Araujo IB, Nogueira BV, de Andrade TU, Endringer DC, Lenz D. *TissueCell.* (2017); 49 (1): 22. DOI: hdx.doi.org/10.1016/j.tice.2016.12.006
- Nischwitz V, Goenaga-Infante H. *J. Anal. At. Spectrom.,* (2012); 27, 1084. DOI: 10.1039/C2JA10387G
- Paranjpe M, Müller-Goymann CC. *Int J Mol Sci.* (2014); 15 (4): 5852. DOI: 10.3390/ijms15045852
- Patil SS & Bhosale PN, *Mater. Res. Express* (2019) 6, 126404. DOI: 10.1088/2053-1591/ab5500
- Patil SS, Mali SS, Hong CK, Bhosale PN. *Mat Sci in Semi Process* (2020); 105: 104727. DOI: 10.1016/j.mssp.2019.104727
- Patil, SS, Bagade, CS, Joshi, MP, Kharade, SD, Khot KV, Mali SS, Hong CK, Bhosale PN, *J Mater Sci: Mater Electron* (2018) 29, 19322–19335. DOI: 10.1007/s10854-018-0059-0
- Turkevich J, Stevenson PC, Hillier J. *Faraday Soc.,* (1951); 11, 55. DOI: 10.1039/DF9511100055
- Vippola M, Valkonen M, Sarlin E, Honkanen M, Huttunen H. *Nanoscale Res Lett* (2016); 11 (1): 169. DOI: 10.1186/s11671-016-1391-z

Accepted Manuscript

Chemical Papers
<https://doi.org/10.1007/s11696-020-01123-3>

ORIGINAL PAPER



An inexpensive, automated and reproducible method to conduct quality control in nanoparticles

Gabrielly Pereira Ribeiro¹ · Rafaela Spessemille Valotto² · Jairo Pinto de Oliveira³ · Marco Cesar Cunegundes Guimarães² · Dominik Lenz¹

Received: 13 November 2019 / Accepted: 26 February 2020
© Institute of Chemistry, Slovak Academy of Sciences 2020

Abstract

Multiple methods have been developed to estimate size and shape of nanoparticles from transmission electron microscopy images. However, some of them have limiting characteristics such as hard to use, expensive or have steps that make the results unreal. We present here a comparison of a manual (IMAGEJ) and an automated analysis, both free methods. The gold nanoparticles were synthesized by the traditional Turkevich method using sodium citrate. We used 23 images of gold nanoparticles. The results of our study show that the conglomerated particles can be excluded—given that a sufficient amount of particles is included in the analysis. We show that CellProfiler can be highly efficient to assumptions about size and shape of nanoparticles. These findings provide a new approach of free software that can be applied as machine learning and therefore automate the processing of nanoparticles data.

Keywords Nanoparticles · CellProfiler · Transmission electron microscopy · Automated analysis

Capítulo 2

Artigo 2. Automated cytopathology of fine needle aspiration for the detection of malignancy in thyroid cells

AUTOMATED CYTOPATHOLOGY OF FINE NEEDLE ASPIRATION FOR THE DETECTION OF MALIGNANCY IN THYROID CELLS

Gabrielly Pereira Ribeiro¹, Carlos Musso² and Dominik Lenz¹

¹ Post-Graduated Program of Pharmaceutical Sciences University, Vila Velha, ES, Brazil

² Federal University of Espírito Santo, Vitória, ES, Brazil

Address: Avenida Comissário José Dantas de Melo, 21 - Boa Vista II, Vila Velha - ES, 29102-920.

55 (27) 3421-2001/ 55 (27) 998733994 – corresponding author:
gabriellyribeiro.uvv@gmail.com

This study was financed by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) - Brazil.

ABSTRACT

Cytopathology of thyroid cells is an established method to detect malignancies in the thyroid. It is of advantage because an anesthesia and a diagnostic laparotomy is not necessary. There are however not yet many studies about automated cytopathology in thyroid cells. To this end, the aim of the present study was to establish an automated diagnosis of malignancy using image analysis and subsequent machine learning and Artificial intelligence. Light microscopy images of 52 patients were analyzed and the results were compared to those of pathology. The results of the automated analysis yielded a sensitivity of 0.94 and a specificity of 0.91 when compared to those of the pathologic diagnoses. The process of machine learning yielded an under curve area of 0.91 as calculated by a ROC-curve. The software used for image analysis, machine learning and classification (diagnosis) are open-source software, respectively. It may be concluded that the method presented in the present study may be used for a low-cost verification of pathological diagnosis. The equipment used for this study (light microscopy, digital camera and a Personal Computer) is available in almost every laboratory.

Keywords: Machine learning; CellProfiler; Image Cytometry; Cytopathology; Papillary Thyroid Carcinoma.

INTRODUCTION

Histopathology (HP), a technique that aims to diagnose diseases in tissue sections using microscopy, established in 1893, remains the gold standard until nowadays (Tiford et al., 2006; den Baaker, 2017). In 1928, Papanicolau described a new technique of cell-based diagnosis called Cytopathology (CP) (Van den Tweel and Taylor, 2010).

Results, advantages and disadvantages of HP and CP have always been subject of comparison (Crino et al., 2021; Gerstner, Tarnok, 2002; de Ruiter et al., 1994). Recent literature reveals that the results of HP and CP in terms of accuracy became comparable, independent if CP is conducted manually (de Ruiter et al., 1994) or automatically (Gerstner et al., 2002; Gerstner et al., 2009).

The studies of Gerstner et al. used Laser Scanning Cytometry (LSC) aiming to define the DNA-ploidy as a marker of malignancy. LSC is automated and thus reproducible, but at least the equipment is cost-intensive. A manual count is low cost but tedious, time-consuming and its results are not reproducible. To this end the authors of the present study aimed to combine an automated analysis that is less expensive. The authors hypothesized that the morphological features yield information about the malignancy/benignancy of the respective sample.

The aim of the present study was to analyze light microscopic images of cytopathologic samples of thyroid cells using free software. The classification (diagnosis) was aimed to be conducted by artificial intelligence after a process of machine learning.

Samples of thyroid were chosen because the incidence of diagnosed thyroid cancers is increasing at least in Brazil and the USA. At least in those two countries, thyroid cancer is the fifth most frequent cancer in women. As stated in the recent scientific literature, an inexpensive but reliable methodology will always be an important issue to lower the costs for reliable medical diagnosis (Gerstner et al., 2002; Mittag et al., 2006; Gerstner et al., 2009; Buzin et al., 2015; Laffers et al., 2006; Melzer et al., 2016).

MATERIALS AND METHODS

Study Design

Manual diagnostic data from cytopathological aspirates were provided by Cassiano Antonio Moraes University Hospital. The automated analysis was conducted at the Image Cytometry Laboratory 10 (B) at the Biopractices Campus of Vila Velha University. The type of study was cross-sectional where the observation of data will occur within a

time interval, without the need for direct interaction with the patients.

Obtaining of Samples and Images

In the present study 52 samples submitted to fine needle aspiration stained by pap smear were collected with a diagnosis made by the pathologist Dr. Carlos Musso; selected from March/2021 to March/2022. The photos taken by the pathologist himself were collected and sent to the researcher. The study was previously approved by the ethics committee of the institution Vila Velha University - (No. 4.094.607) and of the institution Cassiano Antonio Moraes University Hospital - (No. 5.073.110).

Image Preparation and Analysis

The images have been saved in .tif format. Samples of each patient were stored in individual folders without personal identification. The received images were analyzed and classified using free software. The software chosen for nuclei analysis was CellProfiler (CP) (version 4.2.4), after the pipeline (Figure 1) was created in this software, the images of each patient were loaded, there was a background inversion where the colors were inverted to grayscale scales for processing the pixels, presenting you light objects on a dark background, identification of all objects (nuclei) and finally filtered (Figure 2 and 3), in order to exclude non-cellular artifacts.

PIPELINE USED WITH CELLPROFILER FOR ALL SAMPLES

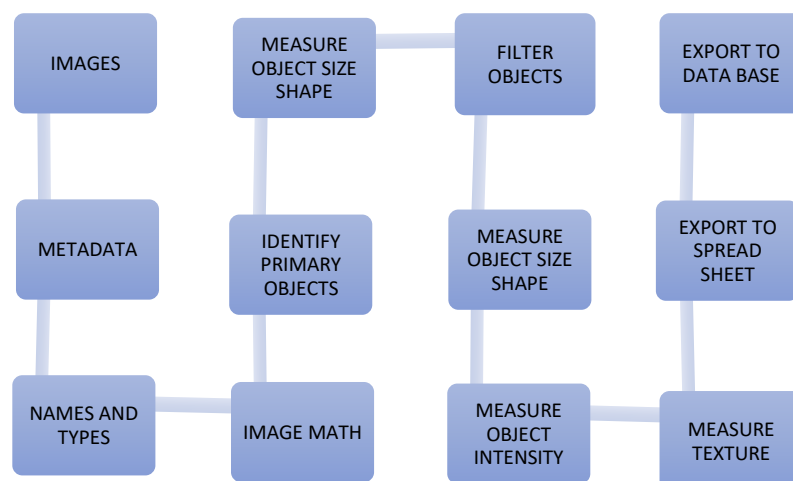


Figure 1: Pipeline: First all the images were loaded, identified by the unique format known in all .tif, the intensities were inverted for the objects of interest asking for light on a dark background, later the primary object was identified as nuclei 1, after filtering the objects (eliminating objects unexpected) a new count of the parameters was performed and then data nuclei 2 was found, only these were exported to carry out the classification.

STEPS USING CELLPROFILER – MALIGNANT SAMPLE

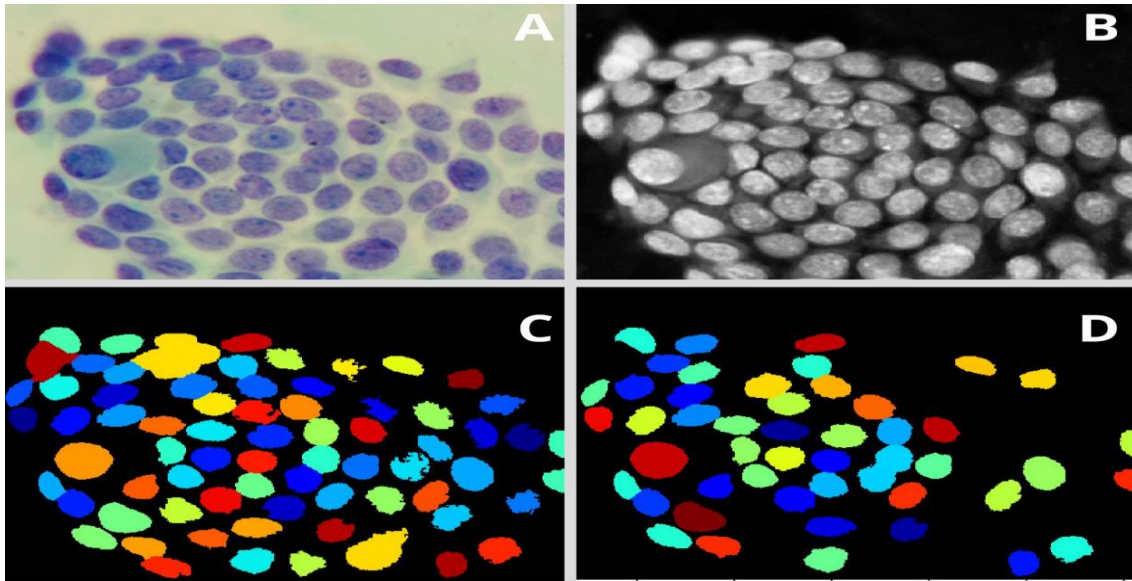


Figure 2: Analysis performed in Cellprofiler of a malignant sample. A: Original Image, B: Color inversion to grayscale, C: Nuclei identification D: Filter applied to exclude the artifacts.

STEPS USING CELLPROFILER – BENIGNANT SAMPLE

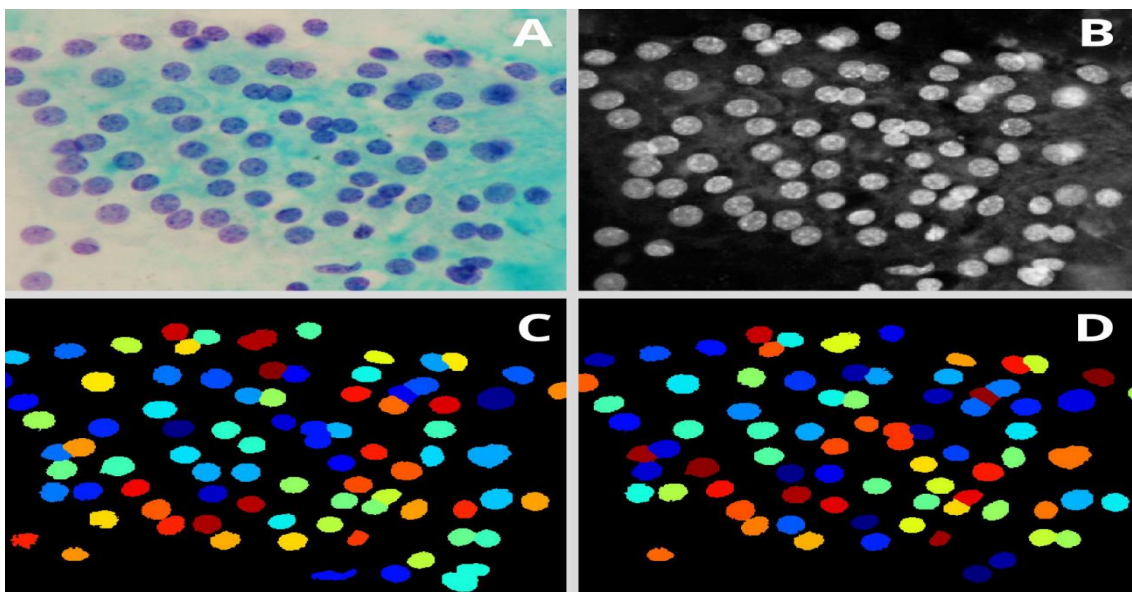


Figure 3: Analysis performed in Cellprofiler of a benignant sample. A: Original Image, B: Color inversion to grayscale, C: Nuclei identification D: Filter applied to exclude the artifacts.

The dates of the analyzed nuclei were exported to an excel file and later to another two software that performed the classification by supervised learning (Cellprofiler Analyst (CPA) and Tanagra (TG) software).

Statistical Analysis

After the complete analysis of all samples a comparison of the results of the effectiveness of the automated method and the manual diagnosis was made with MedCalc software (<https://www.medcalc.org/download/>). Also, calculations for specificity and sensitivity with CPA and TG software.

RESULTS

The ROC curve (Figure 4) shows the progress of machine learning (AUC = 0.96), and the capability to differentiate the positive and negative.

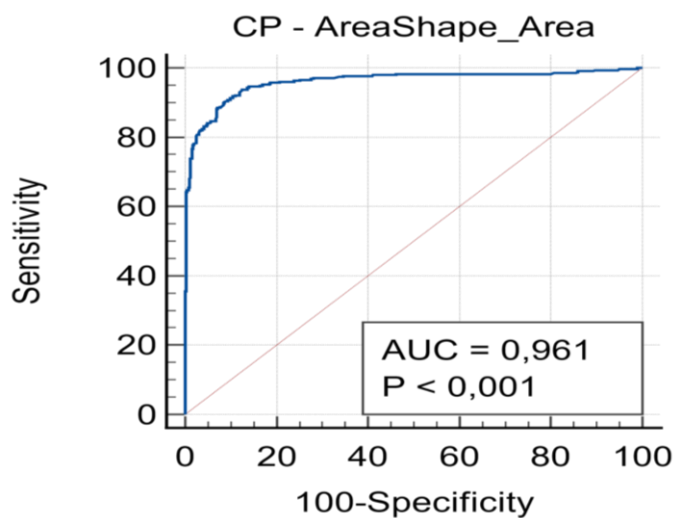


Figure 4: ROC curve between automated and manual results.

TG is a macro to be integrated in excel for classification purposes, amongst others. If, like in the present study there are multiparametric data, TG offers the possibility to choose the best parameters for a successful classification (data mining), that is, parameters that do not substantially enhance the results will be excluded in this step.

The scatterplot shown in figure 5 concluded that the software was able to correctly differentiate the classes.

SCATTERPLOT WITH SOFTWARE TANAGRA – SUPERVISED LEARNING

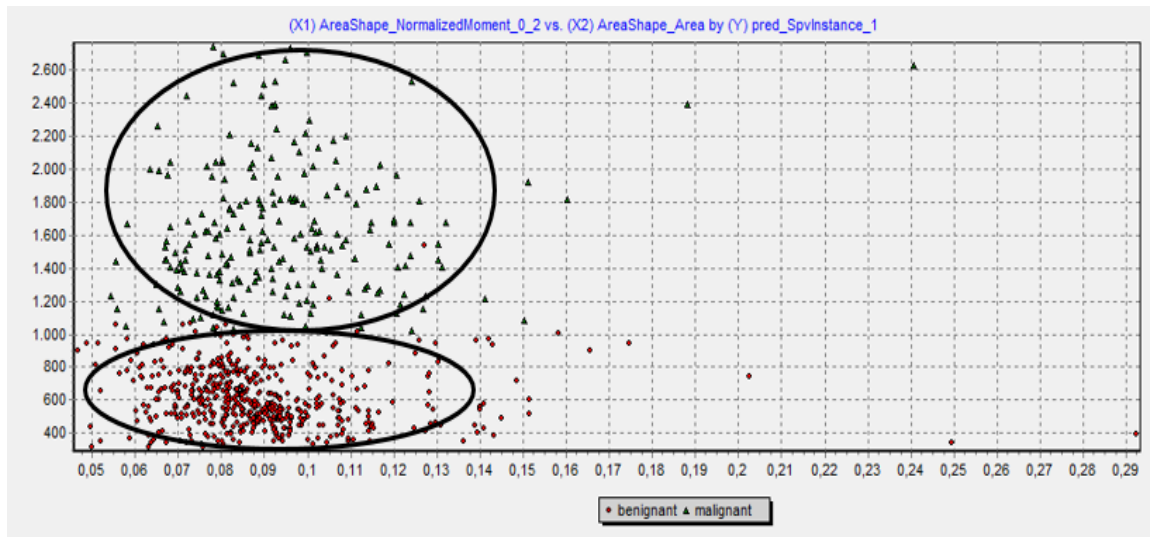


Figure 5: Classification obtained in the TANAGRA software (Scatter plot: red-benign / green-malignant).

The algorithm chosen by the authors was the Linear discriminant analysis (LDA). LDA is strongly related to analysis of variance (ANOVA) and regression analysis, which also attempt to express a dependent variable as a linear combination of other characteristics (Venkataraman et al., 2007). TG software presented data with high sensitivity (0.94) and specificity (0.91) (Figure 6).

| Values prediction | | |
|-------------------|--------|-------------|
| Value | Recall | 1-Precision |
| benignant | 0,9451 | 0,0779 |
| malignant | 0,9104 | 0,0634 |

↓ specificity

↓ sensitivity

Figure 6: Linear discriminant analysis between benignant and malignant objects.

The other software used to also classify the data was CPA, which also received the same data calculated by TG software. Figure 7 represents the training manual classification made. CPA also supervised machine learning and the algorithm used was logistic regression. In this software the results were comparable to the TG software and also showed high performance in learning – 97.08% of accuracy.

CELLPROFILER ANALYST TRAINING

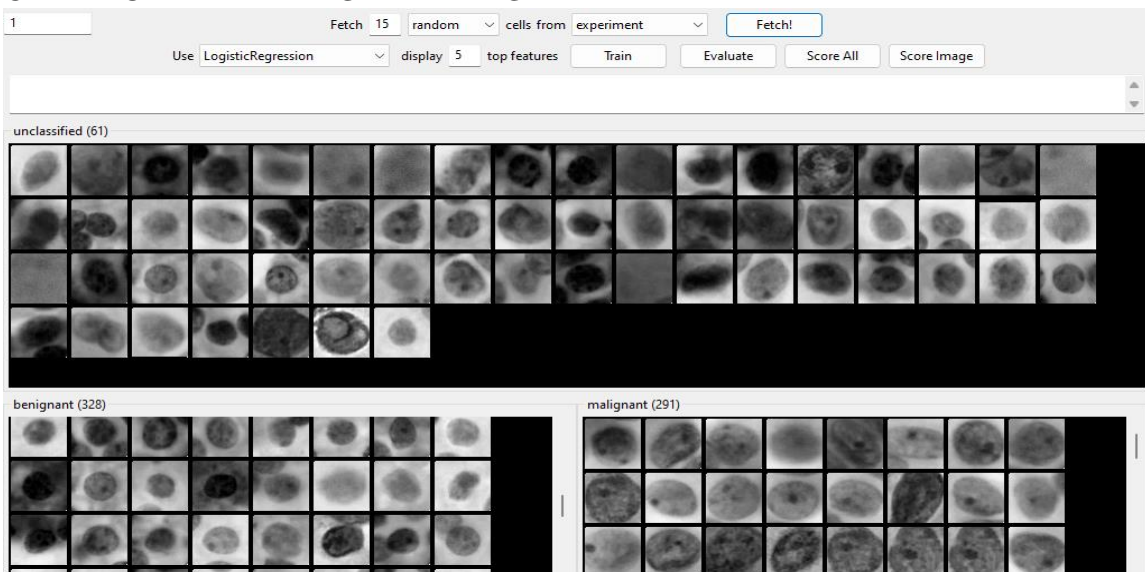


Figure 7: Training with two classes, malignant (lower right) and benign (lower left) cells.

After training benignant and malignant objects were requested to verify software performance to calculate sensitivity and specificity. (Figure 8 and 9).

CELLPROFILER ANALYST PERFORMANCE - BENIGNANT

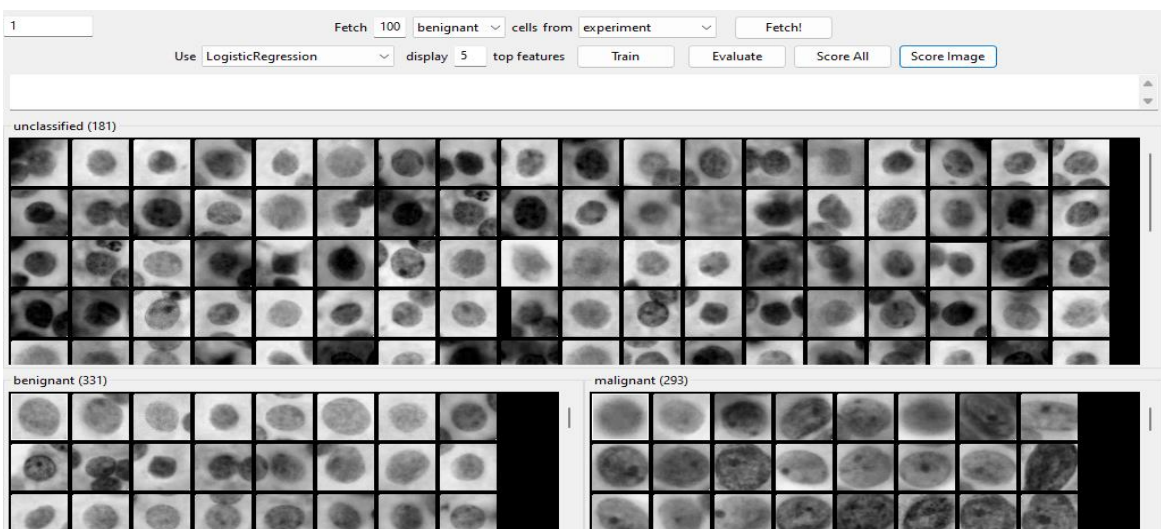


Figure 8: 100 benignant correct objects requested after training.

CELLPROFILER ANALYST PERFORMANCE - MALIGNANT

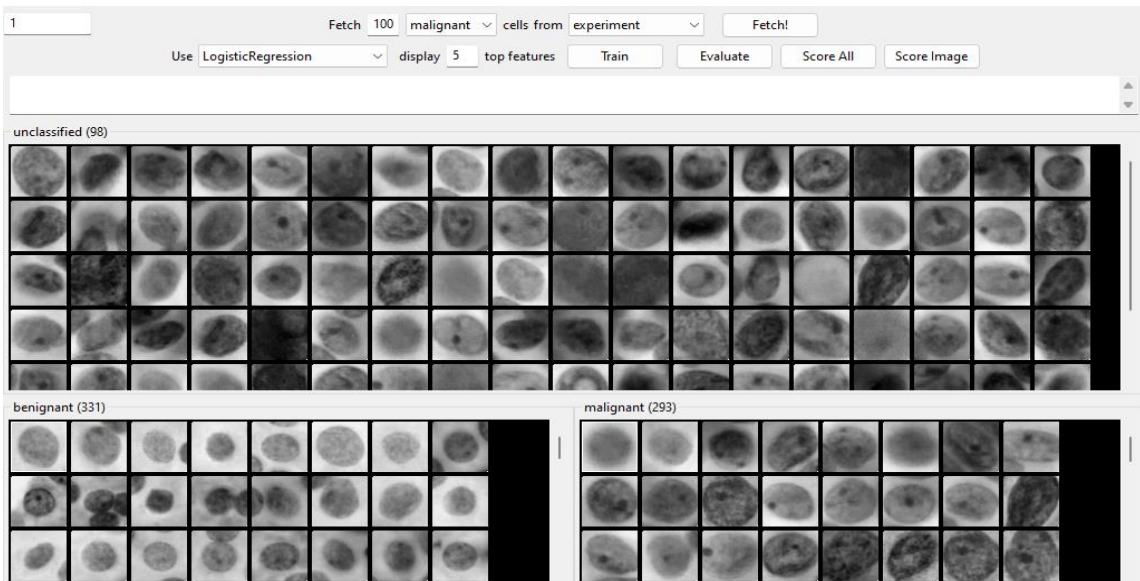


Figure 9: 100 malignant correct objects requested after training.

To check the machine learning progress, a score image tool is displayed after training to show the correctness of the machine (Figure 10).

SCORE IMAGE OF MALIGNANT AND BENIGNANT NUCLEUS – CELLPROFILER ANALYST

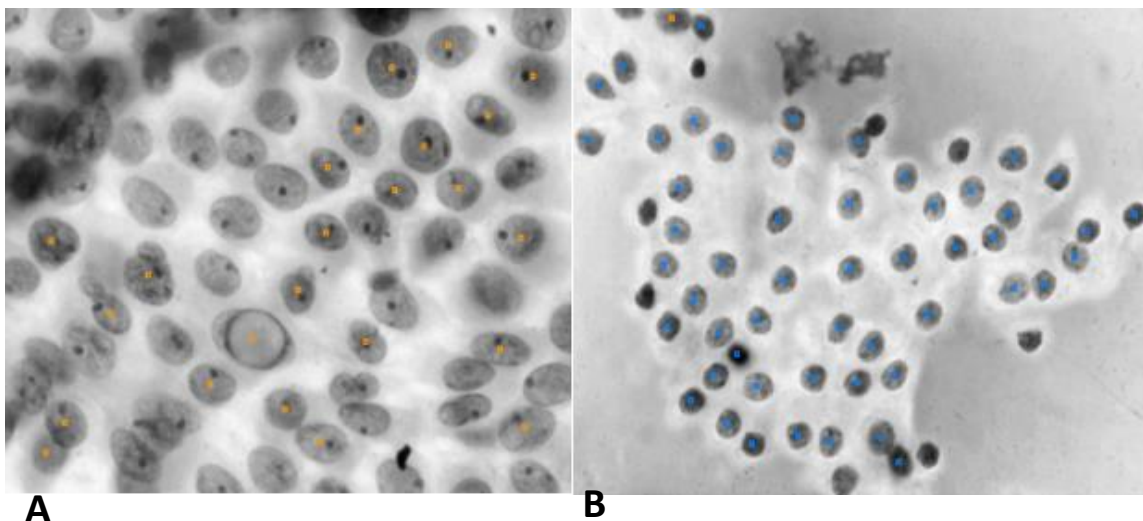


Figure 10: Score Image: This tool allows the analyst to see the full extent of the image and the objects there identified according to the classification performed under each class after training. A: Malignant objects identified B: Benignant objects identified. - 97.08% of accuracy.

The results calculated by CPA and TG obtained accuracy above 90%. Both software were able to distinguish the difference between classes. Regardless of the number of samples, the algorithm is able to learn and perform the correct classification of objects, as shown in the results of the present study.

DISCUSSION

Fine-needle aspiration cytology is a routinely used, reliable, quick, highly reproducible, minimally invasive, and cost-effective method for diagnosing nodular lesions of the thyroid (Rosai, 1996). It has been suggested as a first step test in the diagnosis of thyroid nodules (Segev et al., 2003; Aguilar et al., 1997). The most important contribution of this method to the diagnostic process is its power to distinguish patients who need surgical intervention from those who do not (Koss, 1992). Manual diagnosis, however, is time-consuming and not reproducible.

Currently there are free software where it is possible to view all objects and classify them with the best parameters. After training all objects, the software can even independently predict a correct diagnosis in new objects. Other studies also had satisfying results with data found in automated analysis of cellular objects (Deka et al., 2017).

An interesting study combining machine learning and papillary thyroid carcinoma published in 2021 showed several machine learning models were built to predict disease recurrence and all models had a confirmed accuracy of 90% or more (Park & Lee, 2021).

Automated quantitative analysis of cellular objects is currently considered the viable option in developed countries, because the scientific community proves that human-machine interaction has the potential to reduce human effort and increase confidence in algorithmic decision to solve clinical diagnoses (Jayatilake & Ganegoda, 2021; Sarker, 2021).

In the literature, there are few studies with images of the cell classification process (Wang et al., 2010; Wang et al., 2019), unlike this study, which the objective was to show step by step.

In terms of DNA ploidy there may be a low percentage of false positive artifacts. However, those may not have influenced the results of the present study as proven by the high accuracy of the results (Gerstner et al., 2006). The ROC curve was created only using the area of the nuclei, concluding the area of the nuclei is referring to the amount of DNA, which makes an analysis of the ploidy relatively easy, even when using light microscopy as in the present study.

The term digital pathology is considered as the future of histopathology. The authors agree that the integration of images and digital results is not yet standard, but it will significantly contribute to receiving digital pathology as part of the clinical evaluation (Athelougou & Huss, 2018; Ribeiro et al., 2019). Other studies concluded that after several highly accurate and scientifically proven studies, histopathology is ready for automation (Ribeiro et al., 2020, Lucas et al., 2021).

The main purpose of this study was to compare the diagnosis manual analysis (gold standard) and automated analysis. The results are comparable to the gold standard with high specificity and high sensitivity in two different classifiers.

CONFLICTS OF INTEREST

Nothing to declare.

ACKNOWLEDGMENTS

The authors would like to thank the National Council for Scientific Technological Research (No. process: 88887.370523/2019-00) for their financial support. This study was financed by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES). Thanks very much.

REFERENCES

- Tiford M. A Short History of Histopathology Technique. *J Histotechnol.* 2006;29:99-110.
- den Bakker MA. Histopathologisch onderzoek als gouden standaard? [Is histopathology still the gold standard?]. *Ned Tijdschr Geneeskd.* 2017;160:D981.
- Van den Tweel JG, Taylor CR. Introduction to the History of Pathology series. *Virchows Arch.* 2010;457(1):1-2.
- Crinò SF, Ammendola S, Meneghetti A, et al. Comparison between EUS-guided fine-needle aspiration cytology and EUS-guided fine-needle biopsy histology for the evaluation of pancreatic neuroendocrine tumors. *Pancreatology.* 2021;21(2):443-450.
- Gerstner AO, Tárnok A. Analysis of fine-needle aspirate biopsies from solid tumors by laser scanning cytometry (LSC). *Curr Protoc Cytom.* 2002;Chapter 7:Unit 7.
- de Ruiter A, Carter P, Katz DR, et al. A comparison between cytology and histology to detect anal intraepithelial neoplasia. *Genitourin Med.* 1994;70(1):22-25.
- Gerstner AO, Laffers W, Tárnok A. Clinical applications of slide-based cytometry--an update. *J Biophotonics.* 2009;2(8-9):463-9.
- Mittag A, Lenz D, Bocsi J, et al. Sequential photobleaching of fluorochromes for polychromatic slide-based cytometry. *Cytometry A.* 2006;69(3):139-41.

Buzin AR, Pinto FE, Nieschke K, et al. Replacement of specific markers for apoptosis and necrosis by nuclear morphology for affordable cytometry. *J Immunol Methods*. 2015;420:24-30.

Laffers W, Mittag A, Lenz D, et al. Iterative restaining as a pivotal tool for n-color immunophenotyping by slide-based cytometry. *Cytometry A*. 2006;69(3):127-130.

Melzer S, Nunes CS, Endringer DC, et al. Trypan blue as an affordable marker for automated live-dead cell analysis in image cytometry. *Scanning*. 2016;38(6):857-863.

Rosai J, editor. Ackerman's surgical pathology. Volume 2. 8th ed. St Louis: Mosby. 1996;493–567.

Segev DL, Clark DP, Zeiger MA, et al. Beyond the suspicious thyroid fine needle aspirate: a review. *Acta Cytol* 2003; 47:709-722.

Aguilar-Diosdado M, Contreras A, Gavilán I, et al. Thyroid nodules. Role of fine needle aspiration and intraoperative frozen section examination. *Acta Cytol*. 1997;41(3):677-682.

Koss LG. Diagnostic cytology and its histopathologic basis. 4th ed. Philadelphia: JB Lippincott. 1992;2665–2708.

Deka L, Gupta S, Gupta R, et al. Nuclear morphometry and texture analysis on cytological smears of thyroid neoplasms: a study of 50 cases. *Malays J Pathol*. 2017;39(1):33-37.

Park YM, Lee BJ. Machine learning-based prediction model using clinico-pathologic factors for papillary thyroid carcinoma recurrence. *Sci Rep*. 2021;11(1):4948.

Jayatilake SMDAC, Ganegoda GU. Involvement of Machine Learning Tools in Healthcare Decision Making. *J Healthc Eng*. 2021;2021:6679512.

Sarker IH. Machine Learning: Algorithms, Real-World Applications and Research Directions. SN

Comput Sci. 2021;2(3):160.

Wang W, Ozolek JA, Rohde GK. Detection and classification of thyroid follicular lesions based on nuclear structure from histopathology images. *Cytometry A*. 2010;77(5):485-494.

Wang Y, Guan Q, Lao I, et al. Using deep convolutional neural networks for multi-classification of thyroid tumor by histopathology: a large-scale pilot study. *Ann Transl Med*. 2019;7(18):468.

Gerstner AO, Thiele A, Tárnok A, et al. Prediction of upper aerodigestive tract cancer by slide-based cytometry. *Cytometry A*. 2006;69(7):582-587.

Athelogou M, Huss R. Tissue phenomics in diagnostic pathology. In Tissue Phenomics – Profiling Cancer Patients for Treatment Decisions, (1st ed.). *Jenny Stanford Publishing* 2018; 175–184.

Ribeiro GP, Endringer DC, De Andrade TU, et al. Comparison between two programs for image analysis, machine learning and subsequent classification. *Tissue Cell*. 2019;58:12-16.

Ribeiro GP, Valotto RS, Oliveira J, et al. An inexpensive, automated and reproducible method to conduct quality control in nanoparticles. *Chem. Pap*. 2020;74:2821–2824.

Lucas AM, Ryder PV, Li B, et al. Open-source deep-learning software for bioimage segmentation. *Mol Biol Cell*. 2021 Apr 19;32(9):823-829.



GLOBAL JOURNAL OF MEDICAL RESEARCH: C
MICROBIOLOGY AND PATHOLOGY
Volume 23 Issue 2 Version 1.0 Year 2023
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals
Online ISSN: 2249-4618 & Print ISSN: 0975-5888

Automated Cytopathology of Fine Needle Aspiration for the Detection of Malignancy in Thyroid Cells

By Gabrielly Pereira Ribeiro, Carlos Musso & Dominik Lenz

Pharmaceutical Sciences University

Abstract- Cytopathology of thyroid cells is an established method to detect malignancies in the thyroid. It is of advantage because an anesthesia and a diagnostic laparotomy is not necessary. There are however not yet many studies about automated cytopathology in thyroid cells. To this end, the aim of the present study was to establish an automated diagnosis of malignancy using image analysis and subsequent machine learning and Artificial intelligence. Light microscopy images of 52 patients were analyzed and the results were compared to those of pathology. The results of the automated analysis yielded a sensitivity of 0.94 and a specificity of 0.91 when compared to those of the pathologic diagnoses. The process of machine learning yielded an under curve area of 0.91 as calculated by a ROC-curve. The software used for image analysis, machine learning and classification (diagnosis) are open-source software, respectively.

Keywords: *machine learning; cellprofiler; image cytometry; cytopathology; papillary thyroid carcinoma.*

6. CONSIDERAÇÕES FINAIS

A análise de imagens e os algoritmos de aprendizado de máquina têm a capacidade de identificar automaticamente os dados de interesse e segmentar células individuais ou outros recursos específicos. É importante ressaltar que os sistemas de automação não substituem os seres humanos, mas auxiliam a otimizar o uso do tempo e obter resultados de maior qualidade, uma vez que estes resultados são quantitativos, não se limitando apenas a aspectos qualitativos.

A automação e a alta precisão dos resultados são importantes para o desenvolvimento de métodos de citometria e para torná-los mais conhecidos na literatura científica. Isso permite que descobertas científicas transcendam diferentes conjuntos de dados experimentais, possibilitando que as descobertas relatadas neste estudo e em estudos futuros sejam validadas em laboratórios ao redor do mundo.

APÊNDICES

➔ DURANTE O ÚLTIMO ANO DO DOUTORADO RECEBI UM CONVITE PARA SER PALESTRANTE NO EVENTO FESBE: FEDERAÇÃO DA SOCIEDADE DE BIOLOGIA EXPERIMENTAL PARA APRESENTAR RESULTADOS COM O USO DE CITOMETRIA DE IMAGEM EM SOFTWARES GRATUITOS.



Certificado

Certificamos que

GABRIELLY RIBEIRO

ministrou a Palestra: "Análise de Imagem, Aprendizado de Máquina e Inteligência Artificial para um diagnóstico automático e reproduzível - Seeing is believing." na XXXVII Reunião Anual da FeSBE realizada no período 27 a 30 de agosto de 2023, no Hotel Atlântico Búzios, em Búzios, RJ.


Prof. Dr. Eduardo Colombari
Presidente da FeSBE

Patrocinador

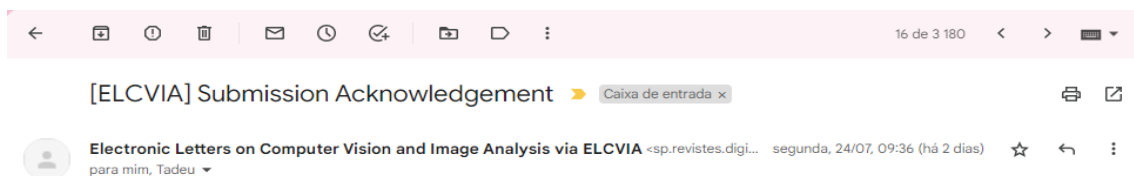


Apoio



➡ ALÉM DOS RESULTADOS APRESENTADOS, OUTROS 2 ARTIGOS DESENVOLVIDOS DURANTE O DOUTORADO ESTÃO SUBMETIDOS, AGUARDANDO RETORNO DE REVISORES.

ARTIGO I:



Images analysis, machine learning and automated classification: A review for those who want to facilitate cellular diagnostics.

Gabrielly Pereira Ribeiro¹; Dominik Lenz¹

¹ Post-Graduated Program of Pharmaceutical Sciences University, Vila Velha, ES, Brazil

Corresponding author: gabriellyribeiro.uvv@gmail.com / dominik.lenz@gmail.com

ARTIGO II:

Laser Scanning Cytometry - The future of cytometry became past.

Gabrielly Ribeiro¹ Dominik Lenz^{1*}

¹ Post-Graduated Program of Pharmaceutical Sciences University, Vila Velha, Espírito Santo, Brazil

Address: Avenida Comissário José Dantas de Melo, 21 - Boa Vista II, Vila Velha - ES, 29102-920.