INTERNATIONAL DOCTORAL SCHOOL OF THE USC

Zakarya Abdullah Ayid
Al Tarawneh

PhD Thesis

Machine learning algorithms and software to quantify immunohistochemical images in biomedicine

Santiago de Compostela, 2025

Doctoral Programme in Information Technology Research

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

ESCOLA DE DOUTORAMENTO
INTERNACIONAL DA USC

TESE DE DOUTORAMENTO

Machine learning algorithms and software to quantify immunohistochemical images in biomedicine

Zakarya Abdullah Ayid Al Tarawneh

Dirección: Eva Cernadas García and Pilar Gándara Vila

Tutora: Eva Cernadas García

PROGRAMA DE DOUTORAMENTO EN INVESTIGACIÓN EN TECNOLOXÍAS DA INFORMACIÓN

SANTIAGO DE COMPOSTELA, 2025

CONFLICT OF INTERESTS

The doctoral candidate, Zakarya Abdullah Ayid Al Tarawneh, declares no conflicts of interest related to his thesis entitled: **Machine learning algorithms and software to quantify immunohistochemical images in biomedicine**

In Santiago de Compostela, Novembro de 2025

Zakarya Abdullah Ayid Al Tarawneh
Thesis author


USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

To my parents, who have been loving, praying, and believing in me unconditionally throughout my life.

To my wife, my life partner and companion in every success; her patience, love, and sacrifices have made this journey possible.

To my children, whose smiles and laughter have been a constant source of strength, joy, and inspiration.

To my sister and brother-in-law, for their continued support and encouragement.

And to my friends who have stood by me with real support and friendship, Dr. Fawaz, Dr. Omr, Dr. Mutaz.

This work is dedicated to all of you with deep love, gratitude, and respect.

Acknowledgments

I would like to express my deep and sincere gratitude to my highly valued supervisor, Professor Eva Cernadas García, for her wise guidance through the whole study process and constant support and encouragement. She has been a true source of inspiration. Due to her deep knowledge, constructive suggestions, and remarkable dedication, I reached greater maturity and precision in my works. Her contributions have played a significant role in shaping this thesis into its final form. The professionalism and humanity that characterize her have been a great honor for me to know, and I am deeply grateful to her. I would also like to extend my sincerest gratitude and appreciation to Professor Pilar Gándara Vila and Professor Manuel Fernández Delgado, who enriched this research with their contributions. Special thanks are extended to my supportive and insightful colleagues, Professor Ahmad Al-Hasanat and Dr. Ahmad Al-Tarawneh, for their continuous advice and encouragement. I am also indebted to the University of Santiago de Compostela (USC) and the CiTIUS Research Center, which provided all the infrastructures and research environment necessary for the development of this work. Lastly, I would like to thank Mutah University for making available the Artificial Intelligence Laboratory at the College of Information Technology, which gave me great opportunities to learn and experiment.

Novembro de 2025


USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Listado de publicacións

As publicacións asociadas á tese son as seguintes:

1. Z.A. Al-Tarawneh, M. Pena-Cristóbal, E. Cernadas, J.M. Suarez-Peñaranda, M. Fernán-dez-Delgado, A. Mbaidin, M. Gallas-Torreira, P. Gándara-Vila. **OralImmunoAnalyser: a software tool for immunohistochemical assessment of oral leukoplakia using image segmentation and classification models**. *Frontiers in Artificial Intelligence*, 2024, Vol. 7, pp. 1–13. ISSN 2624-8212. DOI: 10.3389/frai.2024.1324410. Índice de impacto 4.7 (2024), posicións 58 de 204 en Computer Science-Artificial Intelligence (Q2) e 56 de 258 en Computer Science-Information systems (Q1).
2. Z.A. Al-Tarawneh, A.S. Tarawneh, A. Mbaidin, M. Fernández-Delgado, P. Gándara-Vila, A. Hassanat, E. Cernadas. **Cell Detection in Biomedical Immunohistochemical Images Using Unsupervised Segmentation and Deep Learning**. *Electronics*. Special issue *Applications of Computer Vision, 3rd Edition*, 2025, Vol. 14(18), p. 3705, ISSN 2079-9292. DOI: 10.3390/electronics14183705. Índice de impacto 2.6 (2024), 172 de 366 en Engineering, Electrical & Electronic (Q2) e posición 135 de 258 en Computer Science-Information Systems (Q3).

e a afiliación das autorías é:

1. Z.A. Al-Tarawneh, A.S. Tarawneh, A. Mbaidin and A. Hassanat from the Computer Science Department, Mutah University, Karak 61711, Jordan.
2. Maite Pena-Cristóbal, Mercedes Gallas-Torreira and Pilar Gándara-Vila from Oral Medicine, Oral Surgery and Implantology Unit, MedOralRes Group of University of Santiago, Santiago de Compostela, Spain.

3. Z.A. Al-Tarawneh, Eva Cernadas and Manuel Fernández Delgado from Centro Singular de Investigación en Tecnoloxías Intelixentes da USC (CiTIUS), Universidade de Santiago de Compostela, Spain.

4. José Manuel Suarez-Peñaranda from Pathological Anatomy Service, University Hospital Complex of Santiago (CHUS), and Department of Forensic Sciences and Pathology, University of Santiago. Both in Santiago de Compostela, Spain.

Na primeira publicación o doutorando comparte co-autoría con Maite Pena Cristóbal, pertencente a Oral Medicine, Oral Surgery and Implantology Unit, MedOralRes Group, Facultade de Odontoloxía, Universidade de Santiago de Compostela, que aportou coñecemento en histopatoloxía oral para o deseño e avaliación do software. A contribución do doutorando foi o deseño e creación do software OralImmunoAnalyser, xunto coa xeración de estadísticas de evaluación do mesmo.

A segunda publicación foi realizada exclusivamente polo doutorando, polo tanto a súa contribución é do 100%, é orixinal e propia do doutorando.

Autorización das revistas

This article **OralImmunoAnalyser: a software tool for immunohistochemical assessment of oral leukoplakia using image segmentation and classification models**, published in *Frontiers in Artificial Intelligence* is distributed under the Creative Commons Attribution (CC-BY) License, which permits reuse, distribution and reproduction in any medium, provided the original work is properly cited.

https://www.frontiersin.org/journals/artificial-intelligence/about#copyright-statement

This article **Cell Detection in Biomedical Immunohistochemical Images Using Unsupervised Segmentation and Deep Learning** published in *Electronics (MDPI)* is licensed under a Creative Commons Attribution 4.0 (CC-BY 4.0) License, which permits unrestricted reuse in a thesis provided that proper attribution is given.

https://www.mdpi.com/2079-9292/14/18/3705
https://www.mdpi.com/openaccess#Permissions


USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Contents

Resumo en galego

Esta tese pertence ó ámbito da visión por computador e a aprendizaxe automática. O seu obxetivo é deseñar solucións computacionais para a análise cuantitativa de imaxes de immunohistoquímica no ámbito da odontoloxía, que axuden o estudo do cancro oral. Esta investigación está supervisada polas doutoras Eva Cernadas García e a Dra. Pilar Gándara Vila, pertencentes ó ámbito da intelixencia artificial e da odontoloxía, respectivamente.

A tese estructurase nos seguintes sete capítulos:

1. **Introduction**: onde se introduce o problema de cancro oral e as súas dificultades para a diagnose, así como unha breve revisión da investigación realizada no ámbito de patoloxía dixital neste campo.
2. **Obxetivos e hipóteses**: establecer os obxetivos e hipóteses desta tese.
3. **Materials**: que describe a selección e preparación de mostras do tecido bucal para esta investigación, a acquisición de imaxes microscópicas de immunohistoquímica e o proceso de etiquetado das mesmas.
4. **Methodology**: realiza unha breve revisión dos diferentes algoritmos de segmentación de imaxes existentes e propón e describe novas solucións para o noso problema, abordando tanto técnicas supervisadas basada en aprendizaxe profunda como técnica clásicas non supervisadas.
5. **OralImmunoAnalyser software**: é o primeiro capítulo de resultados que describe e avalía o software desenvolto, OralImmunoAnalyser, para a cuantificación de imaxes patolóxicas de tecido bucal.
6. **Results**: neste capítulo realízase unha comparación exhautiva dos algoritmos de detección de células do capítulo 4 nas imaxes patolóxicas de tecido bucal.
7. **Conclusion**: resume as aportacións máis relevantes da tese e propón as liñas futuras de avance.

O **apéndice A** contén un breve manual de uso do software OralImmunoAnalyser.

O **capítulo 1** aborda os retos ós que se enfrenta a investigación en odontoloxía, revisando as propostas existentes. Como o resto de cancros, o cancro oral representa un problema de saude pública grave, cunha incidencia de 389.846 casos novos en 2022 e unha mortalidade de 188.438 mortes, o que supón o 15th posto entre todos os tipos de tumores. Por exemplo, no caso de tumores de melanoma ocupa só un posto por debaixo en incidencia (posto 16th), pero baixa ao posto 22^{nd} en termos de mortalidade. Isto destaca claramente a importancia de buscar formas de lograr un diagnóstico precoz das lesións incipientes de cancro oral. Unha parte importante dos cancros orais desenvólvense a partir de transtornos potencialmente malignos coma a leucoplasia oral. Polo tanto, a súa identificación en etapas premalignas podería ser clave para reducir a mortalidade, morbilidade e custo do tratamento asociado co cancro oral. Un dos principais predictores desta transformación maligna da leucoplasia oral é a presencia de displasia epitelial. Normalmente, a análise immunohistoquímica (IHC polas súas siglas en inglés) é o método de referencia para axudar no diagnóstico e pronóstico do cancro. En particular, a expresión do biomarcador ki-67 é importante tanto como immunomarcador de pronóstico fiable como un factor para determinar o grao tumoral dos distintos tipos de cancro. O tecido tinguido nos cortes IHC examínase baixo o microscopio para que o persoal experto estime manualmente a proporción de núcleos celulares tinguidos na imaxe, coñecida como índice de proliferación do ki-67. A pesar do valioso desta información para o tratamento personalizado do cancro, este proceso require moito tempo e experiencia humana, ademáis de ser tedioso, cansado e existir diferencias de criterio entre quen observa. Estos factores motivaron que nos últimos anos exista un interés crecente en desenvolver algoritmos automáticos e ferramentas software que realicen automáticamente o proceso de identificación e reconto de núcleos celulares na imaxe IHC ou faciliten este proceso. Actualmente, aínda que se propuxeron algúns métodos de análise de imaxes, comerciales ou non, para identificar e cuantificar as células positivas para o biomarcador ki-67, a súa precisión é limitada. Polo tanto, é unha liña de investigación que continúa aberta. A variabilidade inherente os tecidos de distintas persoas, hai que engadir a variabilidade introducida nos procesos de tinción é adquisición das mostras de tecido bucal. Así, un dos puntos máis críticos no proceso de cuantificación das imaxes é a detección robusta dos núcleos celulares.

O **capítulo 2** aborda as hipóteses e obxetivos desta tese, que inclúe o deseño de programas que axuden a cuantificación de imaxes IHC de tecido bucal e o deseño de algoritmos automáticos para a detección de núcleos celulares en ditas imaxes.

As mostras de tecido oral utilizadas nesta investigación pertencen a pacientes do Departamento de Medicina Oral, Ciruxía Oral e Implantoloxía da Facultade de Odontoloxía da Universidade de Santiago de Compostela, que deron o seu consentemento para participar nesta investigación. O **capítulo 3** presenta os materiais utilizados na investigación. Seleccionaronse pacientes tratando de acadar unha máxima variabilidade no seu diagnóstico e, polo tanto, unha ampla variabilidade entre as imaxes de IHC. As imaxes foron seleccionadas de pacientes diagnosticados con leucoplasia oral para os que se dispuña dun informe de biopsia para confirmar o diagnóstico e o grao de displasia. As mostras cirúrxicas procesaronse no Servizo de Patoloxía do Complexo Hospitalario Universitario de Santiago de Compostela. As fotografías adquirironse cun microscopio Olympus BX51 conectado a unha cámara Olympus DP70. Especialistas en patoloxía etiquetaron 41 imaxes utilizando o software desenvolto nesta tese OralImmunoAnalyser, sinalando a posición dos núcleos celulares e o seu nivel de tinción. Considérase que as anotacións realizadas por estas persoas son o patrón verdadeiro co que comparar os resultados dos métodos automáticos de detección de núcleos celulares para a súa validación estadística. Un resultado desta tese foi a confección da base de imaxes anotadas OIADB, dispoñible no repositorio do CiTIUS[1].

Como xa se mencionou, o problema de cuantificación automática dunha imaxe IHC de tecido oral implica o recoñecemento ou detección de núcleos celulares e a súa clasificación en distintos niveis de tinción. O **capítulo 4** describe a abordaxe que se fixo para resolver este problema. Como o obxetivo só é contar obxetos (núcleos celulares) con distintos niveis de tinción, o proceso máis crítico da solución será a localización ou detección dos núcleos. Cando comezamos esta investigación non se dispoña de imaxes de IHC etiquetadas, nin existían en ningún repositorio. Polo tanto, abordouse o problema cunha filosofía clásica, un sistema de visión por computador composto das seguintes etapas:

1. Pre-procesar a imaxe IHC para atenuar o ruído inherente a mesma ou adecuala a etapa seguinte.

2. Detectar os núcleos celulares na imaxe utilizando técnicas de segmentación de imaxes non supervisadas. Ademáis de revisar as técnicas de segmentación non supervisadas clásicas, propúxose un algoritmo, basado no agrupamento de píxels da imaxe similares, para detectar núcleos celulares na imaxe de IHC. Este algoritmo denominouse RDA ou *region detection algorithm*.

3. Clasificar as células en distintos niveis de tinción, normalmente utilizando técnicas de aprendizaxe supervisadas.

4. Finalmente, realizar o reconto das células en cada nivel de tinción, deseñando un algoritmo para dividir o epitelio en tres capas e realizar o reconto para cada estracto epitelial.

Nembargantes, nas últimas décadas, a aprendizaxe profunda emerxeu como unha técnica eficiente para a segmentación ou detección supervisada de obxetos en imaxes. Entre as técnicas de aprendizaxe profunda existentes, as diferentes variantes da arquitecturas U-Net e YOLO (*You Only Look Once*) son as máis destacadas para a segmentación de imaxes e detección de obxetos respectivamente. Adaptamos o uso de U-Net e YOLO para resolver o problema de detección de núcleos celulares en imaxes IHC de cancro oral. Concretamente, o modelo U-Net emprega a regresión do mapa de calor para detectar os centros de masa dos núcleos e adaptouse YOLO

---

[1] https://gitlab.citius.gal/analyser/oiadb

con caixas delimitadoras sintéticas. Aínda que os modelos de aprendizaxe profundo se poden utilizar conxuntamente para detectar e clasificar obxetos, nesta tese so se utilizan para detectar obxetos co obxetivo de comparalos cos métodos de segmentación non supervisada.

O **capítulo 5** describe e avalía no laboratorio de biomedicina da Facultade de Odontoloxía o software OralImmunoAnalyser (OIA), unha ferramenta gráfica para cuantificar as imaxes IHC de leucoplasia oral. OIA é unha aplicación de escritorio que se executa nun ordenador de propósito xeral cos sistemas operativos Linux e Windows. Foi escrito na linguaxe de programación C++ usando a biblioteca GTK+ (GIMP Tool Kit) para desenvolver a súa interface gráfica e a biblioteca OpenCV para desenvolver os algoritmos automáticos para procesar as imaxes. OralImmuno-Analyser pódese descargar desde o repositorio do CiTIUS (https://citius.usc.es/transferencia/software/oralimmunoanalyser) para fins de investigación. OIA deseñouse coa clásica arquitectura de tres capas: a capa de interface para xestionar a nosa interación co programa, a capa de aplicación que inclúe os módulos para o procesamento de imaxe IHC e a capa de persistencia para almacenar os datos da análise e cuantificación da imaxe IHC. A capa de aplicación contén os seguintes módulos:

1. Módulo de detección de núcleos celulares: que inclúe os algoritmos de detección RDA e EDA ou *edge detection algorithm* (ambos detallados no capítulo 4). Ambos algoritmos poden aplicarse separadamente ou combinar as deteccións de ambos métodos.

2. Módulo de clasicación dos núcleos celulares segundo o seu nivel de tinción: OIA incorpora o clasificador supervisado máquina de soporte vectorial, *Support Vector Machine* (SVM) en inglés, implementado pola libraría `libsvm` coa interface en C++ para clasificar os núcleos celulares en tres niveis: tinción alta, baixa e sen tinción. Tamén permite o adestramento do clasificador en caquer momento a través da interface de OIA.

3. Módulo de conteo: que permite non só o reconto global dos núcleos celulares en cada nivel de tinción, senón que realiza o reconto para cada extracto epitelial.

Os detalles para o manexo de toda a funcionalidade de OIA pode consultarse no manual de uso do apéndice A. O funcionamento típico de OIA das persoas de biomedicina é: abrir unha imaxe IHC, debuxar manualmente a rexión de análise na imaxe, detectar automáticamente os núcleos celulares, clasificar as células, revisar a análise automática da imaxe engadindo, borrando ou modificando células e, finalmente, exportar os datos cuantitativos da análise coa porcentaxe de núcleos celulares en cada nivel de tinción para cada extracto epitelial. O funcionamento de OIA foi avaliado no traballo rutineiro do personal experto do laboratorio de biomedicina colaborador. Para iso OIA rexista en arquivos XML (*Extensible Markup Language*) a interacción que o persoal experto realiza co software. Na avaliación estatística dos módulos de procesamento automático da imaxe IHC obtiveronse os seguintes resultado:

- O módulo de detección automática de células acadou unha sensibilidade promedio do 64,4%, unha especificidade do 93% e unha precisión do 60,7%.
- A clasificación dos núcleos celulares en tres niveis de tinción (alto, baixo e sen tinción) acadou unha precisión promedio do 79,8%.

Estes resultados amosan a dificultade de acadar un sistema que funcione totalmente automático. Polo tanto, parece necesaria a supervisión experta para unha cuantificación precisa das imaxes. Nembargantes, aínda que o procesamento automático dunha imaxe IHC necesita menos dun minuto, o tempo total para analizar cada imaxe depende das necesidades de supervisión. O persoal de biomedicina estimou que o uso de OIA permitíalles ir 2.3 veces máis rápido que no reconto manual tradicional, evitando erros porque se pode comprobar en calquera momento as células contadas. Polo tanto, a pesar de que OIA non pode funcionar de forma totalmente automática, pode acelerar considerablemente a análise na práctica clínica diaria, sendo un gran avance sobre o que está dispoñible actualmente. De feito, a percepción subxetiva do persoal experto sobre OIA alcanza unha puntuación media de 80,9 no cuestionario *system usability scale* (SUS), o que significa que o sistema é de bo a excelente. Ademáis de ser sinxelo de usar e instalar, e de permitir supervisar a análise antes da cuantificación da imaxe, OIA permite dividir o estracto epitelial en tres terzos (basal/medio/superior) para comparar a positividade de ki67 entre as capas e estudar a súa relación co grao de displasia epitelial. Por outra banda, o uso de OIA no laboratorio biomédico permitiu anotar o conxunto de datos OIADB, que está dispoñible publicamente no repositorio do CiTIUS (https://gitlab.citius.gal/analyser/oiadb) para o seu uso en futuras investigacións.

A segmentación de imaxes de IHC para detectar a posición dos núcleos das células é un tema complexo na análise automática de imaxes en patoloxía. No capítulo 5 presentouse e avaliouse o software OralImmunoAnalyser (OIA), que se empregou para crear o conxunto de datos OIADB con 41 imaxes IHC de tecido bucal con anotacións. A ferramenta OIA inclúe dúas abordaxes de segmentación de imaxes para detectar os núcleos das células, que se poden aplicar ás imaxes IHC usando a interface gráfica de OIA. Non obstante, este uso non nos permite avaliar de forma independente cada abordaxe para ter unha comprensión máis profunda dos puntos fortes e débiles de cada técnica. Os dous algoritmos de segmentación non supervisada incluídos en OIA chamáronse EDA (un algoritmo de detección baseado en bordos) e RDA (un algoritmo de detección baseado en rexións). No **capítulo 6**, denomínanse OIA-EDA e OIA-RDA para salientar que están incluídos no software OIA. Este capítulo presenta unha avaliación exhautiva comparando diferentes técnicas de segmentación non supervisadas con técnicas supervisadas utilizando aprendizaxe profunda. OIA-RDA proporciona unha precisión P alta, arredor do 87,6%, e unha recuperación R e puntuación F_1 relativamente baixa (R=24,4% e F_1=35,9%), aínda que o rendemento depende da imaxe específica. Un comportamento similar ao OIA-RDA, aínda que cun rendemento global menor, obsérvase para os contornos activos (ChV) cunha media de F_1=21% (P=89,4%, R=12,6%). Por outra banda, OIA-EDA mostra un rendemento máis equilibrado (é dicir, R e P máis semellantes entre si) que OIA-RDA, acadando R=44,1% e P=46,7% con F_1=41,0%, que é o rendemento máis alto entre todos os métodos non supervisados. O método de grafos FH mostra un comportamento similar ao de OIA-EDA, pero cun rendemento xeral inferior, F_1 =34,5% (P=40,9% e R=34,6%). As abordaxes baseadas en clústeres xeralmente teñen un rendemento inferior, con implementacións de K-means tanto nos espazos de cor RGB como LAB que producen valores de R e P arredor do 12-13% e o 40%, respectivamente, e puntuacións de F_1 correspondentemente baixas (16,7-17%). O método de clústeres SLIC ten un rendemento particularmente deficiente, $R = 0,3\%$ e $P = 4,2\%$ case insignificantes. Non se proporcionan resultados cuantitativos do algoritmo MeanShift debido ao seu baixo rendemento. Os resultados demostraron que os métodos clásicos poden proporcionar un rendemento aceptable en condicións específicas. Por exemplo, o algoritmo de segmentación non supervisado OIA-RDA detecta correctamente a maioría das células altamente tinguidas. Aínda que son moi rápidos (só precisan un ou dous segundos por imaxe) e non requiren recursos de hardware específicos, xeralmente están limitados á hora de manexar a complexidade dalgunhas imaxes de IHC, especialmente en presenza de artefactos ou defectos na preparación da mostra. Globalmente, o maior rendemento promedio acadado é só de F_1=46,4% combinando o algoritmo OIA-RDA e OIA-EDA, denominado OIA-RDA-EDA. En canto ao seu rendemento na detección de células con diferentes niveis de tinción, as técnicas baseadas en rexións como OIA-RDA foron moi sensibles na detección de células altamente tinguidas (63,9% de células altamente tinguidas) e acadaron unha baixa sensibilidade co resto dos niveis de tinción. Non obstante, investigacións recentes suxiren que as células intensamente tinguidas están asociadas a fases celulares específicas, estando presentes con máis frecuencia en casos de carcinoma invasivo ou carcinoma in situ. Isto fai que a súa detección sexa particularmente importante. Por outra banda, as células non tinguidas ou lixeiramente tinguidas identifican fases celulares que poden ser consistentes coa normalidade ou coa presenza de hiperplasia. Os algoritmos baseados en bordos, como OIA-EDA, funcionan uniformemente con todos os niveis de tinción celular.

Como se esperaba, a detección de células utilizando modelos de aprendizaxe profunda conseguiu mellor rendemento que a segmentación non supervisada, porque usan o coñecemento do problema durante o proceso de adestramento do modelo e antes do seu uso para a detección das células. U-Net acada o rendemento máis alto con F_1=75,3%, R=74,4% e P=76,8%. YOLO é lixeiramente inferior, en termos de F_1, acadando F_1=74,0%, R=69% e P=81%. De feito, aplicando o test de Wilcoxon comparando a puntuación de F_1 do modelo U-Net co resto de métodos de detección de núcleos celulares concluímos que U-Net non é estatisticamente significativo só para YOLO (valor de p = 0,358122) sendo significativo para os outros métodos con $p \sim 10^{-14}$, que cumpren que $p < 0,05$. Isto significa que os métodos de aprendizaxe profunda U-Net como YOLO superan claramente os métodos non supervisados, como se esperaba. Ademais, o comportamento de ambos métodos é uniforme con todos os niveis de tinción das células. Isto destaca a eficacia da aprendizaxe profunda para a detección robusta de células na análise histopatolóxica.

Non resulta doado comparar as abordaxes de detección das células supervisada e non supervisada. Os métodos de segmentación tradicionais, como a detección de bordos ou a agrupación en clústeres, adoitan funcionar con tempos razoables (un ou dous segundos para procesar unha imaxe) para operar en tempo real en ordenadores de propósito xeral. Debido a isto, estos métodos son aceptables para institucións ou persoal de biomedicina con poucos recursos de procesamento, ou en tarefas de investigación que requiran analizar un número limitado de mostras. Por outra parte, a fase de adestramento dos modelos de aprendizaxe profunda pode ser especialmente intensiva en recursos, empregando habitualmente hardware específico como as GPU (*graphical processing unit*), xa que a miúdo require grandes conxuntos de datos e unha cantidade significativa de tempo para alcanzar o rendemento máximo. A disponibilidade de conxuntos de datos etiquetados no ámbito médico require un custe humano elevado, xa que son tarefas moi tediosas e que requiren persoal moi cualificado. Este feito tamén é unha limitación á hora de utilizar técnicas de aprendizaxe profunda. Isto contrasta fortemente coas abordaxes convencionais, que a miúdo se poden usar de inmediato sen requirir grandes conxuntos de datos nin adestramento intensivo. En xeral, aínda que os métodos de aprendizaxe profunda poden producir mellores resultados de detección, especialmente en situacións complexas, o seu maior consumo de recursos e a dependencia de hardware especializado poden facelos pouco prácticos para algunhas aplicacións, especialmente en entornos con recursos computacionais limitados.

O traballo futuro terá como obxectivo ampliar os modelos de aprendizaxe profunda YOLO e U-Net para realizar tamén clasificacións dos niveis de tinción celular dos obxectos detectados. Tamén planeamos incluír un modelo U-Net de aprendizaxe profunda adestrado no software OralImmunoAnalyser.

Palabras chave: cancro oral, imaxes immunohistoquímicas, visión por computador, segmentación de imaxes, detección de células, aprendizaxe automática, aprendizaxe profundo, clasificación, máquina de vectores de soporte.


USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

English summary

This PhD thesis can be categorized under computer vision and machine learning research. Its objective is to design computational methods or solutions for the quantitative analysis of immunohistochemical images in the context of dentistry in relation to oral cancer studies. This PhD thesis is guided by Dr. Eva Cernadas Garcia and Dr. Pilar Gandara Vila. The former supervisor specializes in artificial intelligence studies, and the later supervisor specializes in dentistry-related research. The chapters of this thesis are organized as follows :

1. **Introduction**: this chapter describes the problem of oral cancer and its challenges in the diagnosis process, in addition to giving a brief overview of the research conducted in the context of digital pathology in oral cancer.
2. **Objective and hypothesis**: this chapter cover the formulation of the objective and hypothesis of this thesis.
3. **Materials**: this chapter discusses the collection of the oral tissue samples (immunohistochemistry microscopic images), their preparation and the description of the acquisition equipment used and the labeling process.
4. **Methodology**: before the introduction of new solutions to detects nucli cells in immunohistochemical images, a brief overview of state of art image segmentation techniques is provided. This chapter proposes solutions based on both deep supervised methods and traditional unsupervised methods.
5. **OralImmunoAnalyser software**: this chapter presents and discusses the developed software tool, OralImmunoAnalyser, developed to measure the pathological oral tissue images.
6. **Results**: This chapter provides a thorough comparison report of the cell detection algorithms discussed in chapter 4 based upon images of pathological oral tissues.
7. **Conclusion**: This chapter summarizes the key contributions of the thesis and future directions.

Appendix A lists the publications derived from the research presented in this thesis, while Appendix B contains a brief user manual for the OralImmunoAnalyser software.

Chapter 1 defines the problem of oral cancer and the challenges in the diagnosis process. Also, it focuses on the difficulties in research in the field of dentistry and the current proposals being suggested. Like other varieties of cancer, oral cancer is a significant health issue in the world. According to 2022 data, there are 389,846 novel cases and 188,438 annual deaths in the world due to oral cancer. This occupies 15^{th} position among all types of tumors. Melanoma tumors occupy 16^{th} position concerning incidence rates but only 22^{nd} in mortality rates. This underlines the significance of discovering oral cancer in the nascent stages. A considerable number of oral cancers occur from potentially malignant conditions like oral leukoplakia. Consequently, the detection of such conditions premalignant may be essential in combating the mortality rate or complications arising from oral cancer. Among the factors that predict malignant transformation in oral leukoplakia, epithelial dysplasia stands out. The analysis of samples using the immunohistochemical (IHC) technique is currently the most commonly accepted procedure in cancer diagnosis and prediction. The use of the prognostic indicator Ki-67 in cancer diagnosis fits under this category and has particular significance in cancer grading. The stained sections are visualized under a microscope by qualified personnel, who visually estimate the number of stained nuclei in the microscopic image, referred to as the Ki-67 proliferation index. Although this information has immense value in customized cancer therapy, this procedure consumes considerable time and involves expert judgment and tediousness. These reasons have driven the interest in recent years in the use of automatic algorithms and software solutions either for the automatic execution or assistance in the identification and quantification of cell nuclei in IHC images. Today, although there are various commercial and non-commercial image analysis solutions proposed in the literature regarding the identification and quantification of Ki-67 positive cells, the accuracy level of such solutions has not been satisfactory yet. Apart from the natural variability between tissues in different individuals, another type of variability that has been taken into consideration is the variability developed in the process of tissue staining and image acquisition. This has made the detection of cellular nuclei one of the most important steps in the quantitative analysis of IHC images. The objective of this thesis is the development of software tools that can help in quantifying the images obtained through oral tissue IHC analysis, in addition to working towards automatic algorithms that can identify cell nuclei in such images. The oral tissue samples processed in this research come from patients in the Department of Oral Medicine, Oral Surgery and Implantology in the Dentistry faculty of the University of Santiago de Compostela. All patients signed consent forms before participating in this study.

Chapter 3 describes the materials used in this study. Patients were selected to capture a broad diagnostic spectrum by maximizing variability, thus yielding a wide range of IHC images. The chosen images were from individuals with oral leukoplakia, with biopsy reports confirming the diagnosis and the level of dysplasia. Surgical samples were processed in the Pathology Department of the University Hospital of Santiago de Compostela. Images were taken with an Olympus BX51 microscope linked to an Olympus DP70 camera. A total of forty-one images were annotated by pathology experts using the OralImmunoAnalyser software developed within this thesis. The pathological experts marked the position of nuclei and their staining intensities. These annotations provide a ground truth for the statistical validation of automatic nucleus detection methods. Another important outcome of this work is the annotated image dataset, called OIADB, which is publicly available at https://gitlab.citius.gal/analyser/oiadb. OIADB consists of 41 high-resolution IHC images of oral tissue affected by oral leukoplakia, representing different degrees of epithelial dysplasia. The average staining distribution within the dataset was highly stained nuclei, 14.8%; low-stained nuclei, 16.4%; and unstained nuclei, 68.7%. Moreover, expert pathologists independently labeled the spatial positions of cell nuclei and provided reliable ground truth data for developing and evaluating automatic nucleus detection and quantification algorithms. As noted, automated quantification of oral tissue IHC images relies upon detecting cell nuclei and classifying them at different levels of staining.

Chapter 4 describes the strategy adopted for the quantitative analysis of IHC images taken from oral tissue. It focuses on a methodological framework that has been developed for automatic and semi-automatic analysis of IHC images, aiming to deliver trustworthy quantitative output. Since the goal is only to count objects-cell nuclei-across a range of staining intensities, the most important part of the pipeline in question is nucleus localization, i.e., detection. The precision in locating nuclei is quite important because any error introduced at this stage propagates to further steps such as classification and final counting. At the early stage of this research, there were lake of annotated immunohistochemical images of oral tissue in any public repository. That absence of labeled data presented a great barrier, especially for supervised methods that rely on such annotated datasets for training and validation. Hence, this work adopted a classical approach in computer vision. This strategy does not require previous annotations and can be applied directly to real images of IHC obtained in biomedical labs. In the proposed classical computer vision framework, the entire processing pipeline goes through the following successive stages:

1. Noise reduction and IHC image pre-processing: The aim here is to prepare images for subsequent processing steps. This step will help in reducing common noise and artifacts in microscopic images and normalize image characteristics arising from different staining and acquisition conditions. Proper pre-processing will make subsequent detection more robust and achieve more stable results across images.

2. Detection of cell nuclei by unsupervised image segmentation methods: Traditional unsupervised segmentation methods are reviewed and adapted to the specific traits of oral IHC images. In addition, a novel algorithm for pixel grouping to detect nuclei was proposed, termed the Region Detection Algorithm (RDA). RDA seeks image regions corresponding to nuclei by leveraging similarities in pixel intensity and color created by the staining.

3. Classification of cells according to staining level, usually by supervised learning. After detection, the nuclei are grouped according to the intensity of staining into biologically meaningful groups such as highly stained, weakly stained, or unstained. Such a classification step transforms raw detections into quantitative, biomedically interpretable information.

4. Finally, counting cells at each level of staining by designing an algorithm that partitions the epithelium into three layers and counts within each layer. Rather than a single global count, the epithelium is segmented into three distinct strata, and counting is performed separately in each. This stratified approach mirrors histopathological practice and allows analysis of the spatial distribution of cell proliferation within the epithelial tissue.

In the last couple of decades, deep learning has emerged as a powerful tool for image supervised segmentation and object detection. Deep learning methods can learn complex visual patterns present in biomedical images and handle variability when annotated data are available. Among these techniques, variants of U-Net and YOLO have been especially prominent in recent years for segmentation and object detection, respectively. This work adapted U-Net and YOLO to address the problem of detecting cell nuclei in oral cancer IHC images. Concretely, U-Net performs heatmap regression to find the nuclear centers of mass, thus allowing for accurate nucleus localization, while YOLO was adapted by using synthetic bounding boxes to directly perform object detection in the images. Although joint use for the tasks of detection and classification is possible, in this work the deep learning models are used only for object detection. This is to provide a fair comparison with the unsupervised segmentation techniques that have been described earlier in the chapter.

Chapter 5 presents and discusses the OralImmunoAnalyser (OIA) software as it operates within the biomedical lab of the Faculty of Dentistry. OIA is a desktop program running on standard computers under Linux and Windows. The chapter is designed to practically validate these methods by setting them into a coherent, ready-to-use system for everyday work in a real biomedical setting. It looks beyond pure algorithmic performance to consider usability, the workflow and how automatic methods interact with expert oversight. OIA is implemented in C++ with the GTK+ toolkit for the graphical user interface and OpenCV for the automatic image processing routines. It is freely available via the CiTIUS repository for research purposes. These choices strongly support portability, efficiency, and accessibility, enabling deployment in typical lab environments without specialized hardware or proprietary platforms.

The architecture follows a classic three-layer model: a user-interface layer, an application layer containing the image processing modules for IHC, and a persistence layer for storing analysis and quantification data. This separation clarifies responsibilities, supports modular development, and eases maintenance and future expansion. The modules within the application layer are:

1. Cell nucleus detection module: This will implement the RDA and EDA described in Chapter 4. These methods can either be used alone or together, giving users flexibility to match nucleus detection to the visual traits of each IHC image and to leverage complementary strategies when needed.

2. Cell nucleus classification module assigns nuclei to staining levels. OIA performs the classification by using a supervised Support Vector Machine classifier through the library libsvm with a C++ interface. The nuclei are divided into high staining, low staining, and no staining. Retraining of the classifier can be done through the interface of OIA at any time, thereby making it adaptable to new datasets, variations in staining, or lab-specific conditions with which expert annotations may be aligned.

3. Counting module: enables global counts by staining level and stratified counts for each epithelial layer. This underpins the quantitative analysis crucial to IHC studies and allows a more granular review of the staining distribution across the tissue layers. Complete operational details of all OIA features are provided in the user's manual (Appendix A). Accordingly, this supports the proper and constant use of the software by biomedical personnel.

A typical workflow for biomedical users is to open an IHC image; manually outline the region of interest where analyses should be performed; automatically detect nuclei, classify cells, review and correct automatic results by adding, removing, or editing cells; and finally export quantitative results on the percentage of nuclei at each staining level for each epithelial layer. This follows routine practice in the lab and emphasizes expert supervision as a core part of the process. Similarly, OIA's performance was assessed within the routine workflow used by expert biomedical staff. It logs user interactions in XML files, allowing detailed analysis of user behavior, time spent, and correction patterns to gauge practical effectiveness. Performance of automatic image processing was:

- Detection of nuclei: sensitivity 64.4% on average, specificity 93%, accuracy 60.7%.

- Nuclei classification into three levels of staining: average accuracy 79.8%.

These figures point out the challenge of fully automatic analysis, keeping expert supervision necessary for precise quantification. Tissue variability and staining complexity make fully automated IHC analysis difficult. Nevertheless, fully automated processing of one IHC image takes less than a minute, while the total analysis time depends on supervision level. Using the OIA, biomedical staff reported that this speeds up workflows about 2.3 times compared with traditional manual counting and reduces errors by allowing them to verify at any step. This represents a meaningful enhancement in efficiency and reliability compared to wholly manual methods. Consequently, although not fully automated, OIA significantly accelerates routine clinical analysis and is favorable compared to other existing solutions. Experts rated OIA at an average SUS (*system usability scale*) score of 80.9, which indicates good to excellent usability. A high usability score reflects positive user experience and alignment with practical lab needs. Besides being easy to use and simple to install, OIA allows for supervised review before quantification and supports dividing the epithelial layer into three thirds: basal, middle, and superior, thus comparing Ki-67 positivity across layers and relating it to epithelial dysplasia grades. This stratified approach enables more nuanced interpretation of the IHC results. One outcome of using OIA in the lab was the development of the openly available dataset, OIADB, for future research; this constitutes another important outcome from the deployment of the software and further methodological development. While the detection of nuclei in IHC images continues to be a challenging task with respect to automated histopathology, the design and evaluation of OralImmunoAnalyser constitute an important step toward more efficient and robust quantitative analyses in oral pathology.

In **Chapter 6**, the terms OIA-EDA and OIA-RDA are used to indicate that those methods live inside the OIA software. The chapter provides a detailed comparison between the unsupervised segmentation techniques and deep-learning–based supervised methods. OIA-RDA is distinguished by a high precision of about 87.6%, but it reveals relatively low recall and F_1, approximately R = 24.4% and F_1 = 35.9%, with significant variations depending on the image. Active contours work similarly to OIA-RDA but run altogether at a lower level, averaging F_1 = 21% with P = 89.4% and R = 12.6%. In contrast, OIA-EDA yields a more balanced mix of precision and recall (P = 46.7%, R = 44.1%) and the highest F_1-score of 41.0%, being the best performer among the unsupervised methods. The graph-based approach by FH mirrors the pattern of OIA-EDA but underperforms in general with an F_1=34.5% or P = 40.9%, R = 34.6%. Cluster-based methods seem to generally bring up the rear, with K-means in RGB and LAB colors yielding R around 12–13%, P about 40%, and F_1 in the 16–17% range. SLIC clustering can be considered very poor, where R = 0.3% and P = 4.2%. MeanShift is not listed because of its weak results. The key takeaway is that the performance of classical methods can be good in favorable conditions. For instance, OIA-RDA can correctly detect most highly stained cells. These are very fast methods, running in one to two seconds per image, without the need for special hardware. However they struggle with the more complex IHC images, especially those containing artifacts or preparation defects. On average, the best unsupervised performance is only $F_1 = 46.4\%$, achieved by combining OIA-RDA and OIA-EDA, i.e., OIA-RDA-EDA.

Regarding the detection of different levels of staining, the region-based methods, such as OIA-RDA, are much more sensitive to highly stained cells, about 63.9%, but considerably less so for other levels of staining. Evidence is emerging that the intensely stained cells colocalize with specific cellular phases and appear more frequently in invasive carcinoma or carcinoma in situ. Thus, it is particularly important to detect these cells. The unstained or lightly stained cells tend to map to normal or hyperplastic phases. On the other hand, edge-based methods such as OIA-EDA perform more consistently across all levels of staining. As might be expected, deep-learning cell detection significantly outperforms unsupervised segmentation, since these models learn task-specific cues during training. U-Net yields the best performance, $F_1 = 75.3\%$, $R = 74.4\%$, $P = 76.8\%$. YOLO is somewhat lower in F_1, $F_1 = 74.0\%$, $R = 69\%$, $P = 81\%$. A Wilcoxon test comparing F_1-scores shows U-Net is not statistically different from YOLO ($p = 0.358$), but is significantly better than all other methods ($p \sim 10^{-14}$), confirming deep learning's edge. Both U-Net and YOLO show consistent performance across all staining levels.This underlines how powerful deep learning can be in robust cell detection in histopathology.

It is not easy to compare the approaches between supervised and unsupervised. Conventional methods based on edge detection or clustering run in a couple of seconds per image, which could allow real-time use on conventional machines. These are accessible for institutions with limited resources or smaller studies. On the other hand, training deep-learning models involves significant resources: usually requiring GPUs (*graphical processing units*) and large amounts of labeled datasets, which have become quite costly to produce, as labeling medical images is a very time-consuming process requiring qualified personnel. Human costs and data may, therefore, hamper the diffusion of deep learning, at least in under-resourced contexts. Overall, while deep learning results are superior in most cases, especially in complex scenarios, the resource and hardware demands become higher, hence sometimes impractical in certain settings.

In future work, we will extend YOLO and U-Net towards the classification of the staining levels of detected cells, and we plan to include such a U-Net model that is trained directly within OralImmunoAnalyser.

Keywords: oral cancer, immunohistochemical images, computer vision, image segmentation, cell detection, machine learning, deep learning, classification, support vector machine.

CHAPTER 1

INTRODUCTION

Immunohistochemistry (IHC) analysis is a gold standard method used by pathologists to help in the diagnosis and prognosis of cancer [4, 6]. The Ki-67 expression is important to both as a reliable prognostic immunomarker [69] or an useful factor for determining tumor grade of various cancer types, including breast cancer [65, 75], bladder carcinoma [22], colorectal cancer, oral cancer [2] and lung cancer. Therefore, a precise and accurate assessment of Ki-67 IHC analysis will provide valuable information about tumor proliferation and aggressiveness. The stained tissue in IHC slices is examined under a microscope, where the pathologist manually estimates the proportion of stained cell's nuclei within the image, known as the Ki-67 proliferation index. This process is very time-consuming and depends on the expertise of pathologists [89], a fact that motivated the development of image analysis methods to automate Ki-67 proliferation index scoring. Currently, several commercial and non-commercial digital image analysis methods have been proposed to identify and quantify the Ki-67 positive cells. But, they have limited accuracy and more reliable methods for Ki-67 estimation in IHC images are stil demanded [4]. The variability in histopathological slides, including differences in tissue specimens and variations in staining, presents significant challenges for automatic image analysis [36]. To address these challenges, it is necessary to design a robust cell detection step that enhances the reliability of image quantification.

1.1 Oral cancer

Oral cancer represents the sixteenth type of cancer worldwide and represents a serious and growing public health problem [41], with an incidence of 389.846 new cases in 2022 and a mortality of 188.438 deaths, ranking 15th among all types of tumours, these figures, in a region so readily accessible to clinical examination, are a matter of considerable for concern. Since another very accessible tumor, such as melanoma, ranks just one place below in incidence (16th), but drops to 22^{nd} place in terms of mortality, this clearly highlights the importance of seeking ways to achieve early diagnosis of incipient oral cancer lesions.

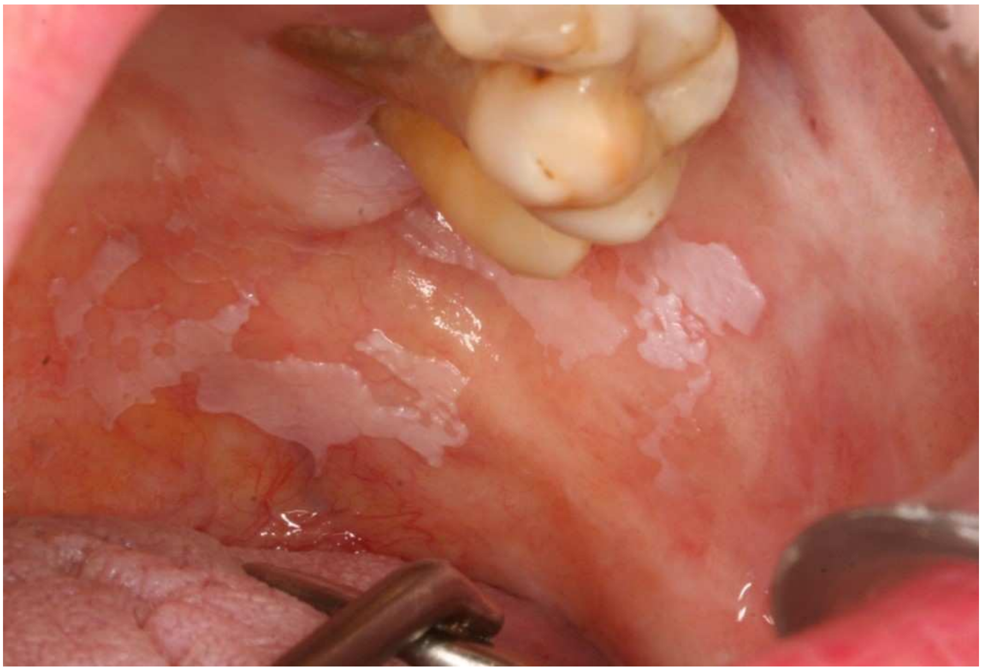


Figure 1.1: Example of homogeneous leukoplakia affecting soft palate and left buccal mucosa.

Figure 1.2: Example of eritroleukoplakia (non homogeneous form) on right lateral border of the tongue/ventral surface, red area indicated by arrow is an in situ carcinoma.

A significant part of oral cancers develop from potentially malignant disorders such as oral leukoplakia defined as “a predominantly white plaque of uncertain risk, having excluded other known diseases or disorders that carry no increased risk for cancer”[81].

Clinical and histopathological analysis by exclusion of other disorders is the conventional diagnosis of oral leukoplakia [77, 80]. It is included among the potentially malignant disorders and is considered the most common lesion within this group, with a global prevalence of 1.39% [87]. Its main and only recognized etiological factor is tobacco in all its forms. Clinically, it is classified as either homogeneous or non-homogeneous. Homogeneous leukoplakia appears as a uniformly thin white plaque, well-defined, and not raised from the surrounding mucosa. It often shows fissures on its surface (see Figure 1.1). Non-homogeneous leukoplakias present clinically in various forms (see figures 1.2, 1.3 and 1.4). Within this group, three types can be distinguished: erythroleukoplakia (a mixed red and white lesion), nodular (a white patch with small polypoid projections), and verrucous, when the surface becomes corrugated [82]. Leukoplakia has a malignant transformation rates ranging from 0.13% to 34.0% [62]. In a study conducted at the University Hospital Complex of Santiago de Compostela on the clinical, epidemiological, and prognostic characteristics of 243 patients with oral squamous cell carcinoma (OSCC) over a 5-year period, 40% were found to be in the most advanced stage, and 24% of the patients had previously had potentially malignant lesions. The 5-year survival rate for patients in the most advanced stage was 28%, compared to 53% for those in the early stages [5]. In another Latin American study, it was also found that 19% of the cancers studied originated from potentially malignant oral lesions, and among this 19%, 55% were cases of oral leukoplakia [32].

So, its identification and intervention predictor of malignant transformation in premalignant stages could be key in reducing mortality, morbidity and the cost of treatment associated with oral cancer [40].

A recent meta-analysis whose objective was to assess the available evidence on the expression of the hallmarks of cancer and the malignant transformation (MT) probability of oral leukoplakia evaluated the clinical and molecular factors that could alert clinicians to an imminent malignancy. It concluded that there is a higher clinical risk in hyperproliferative lesions that develop mesenchymal–epithelial transition phenomena and molecular mechanisms to evade immune response. As clinical markers, it identifies non-homogeneous, large lesions located on the tongue in smoking patients, and as a histopathological marker, the presence and severity of dysplasia [35].

The presence of dysplasia, recognized as one of the main predictors of malignant transformation (MT) of leukoplakia [79, 64, 30], is described as a series of histological changes that have no equivalent correlation in clinical morphology and involve cellular and architectural alterations of

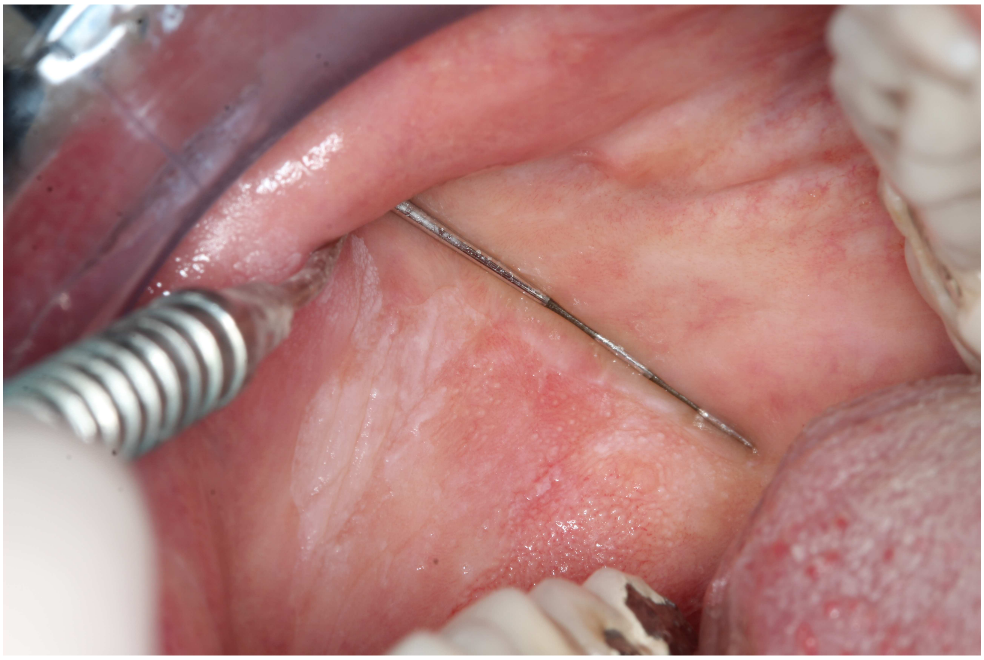


Figure 1.3: Example of Nodular leuokoplakia (non homogeneous form) on right buccal mucosa.

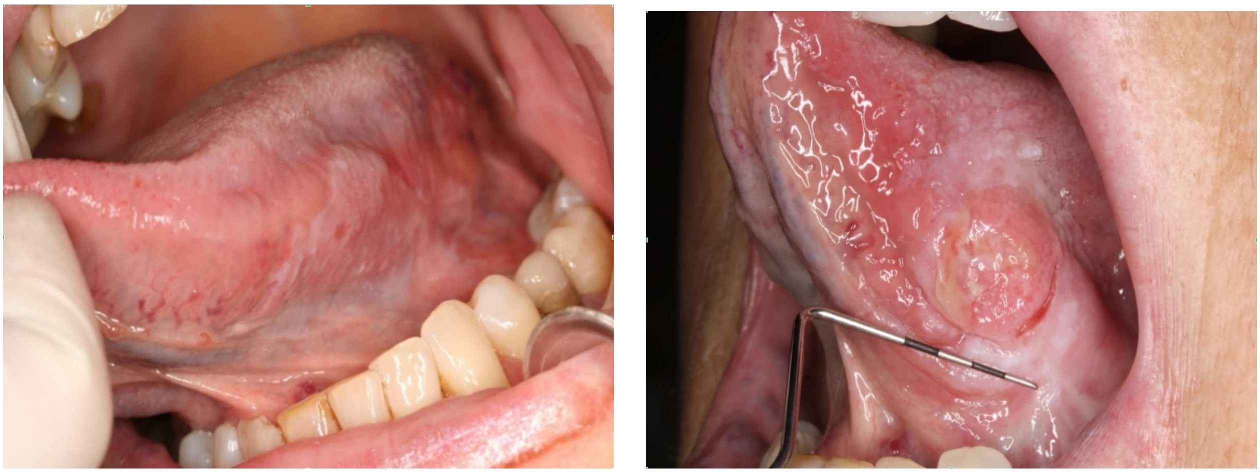


Figure 1.4: Examples of homogeneuous leukoplakia on left lateral border of the tongue (left panel) which transformed over 8 years into an 18 mms OSCC (Oral Squamous Cell Carcinoma) in a patient (right panel) who did not attend follow-up appointments.

the epithelium. As these alterations accumulate, the degree of dysplasia increases [79]. It may persist for long periods without progressing to cancer and can even resolve spontaneously [58].

Dysplasia is classified by the WHO into three levels—mild, moderate, and severe—based on the extent of epithelial involvement. The epithelium is divided into three layers: mild dysplasia is limited to the lower third, moderate dysplasia involves the lower and middle thirds, and severe dysplasia affects the full thickness of the epithelium. Cytological changes present in a mild dysplasia may also modify its grade to severe [81].

However, this diagnosis is based on a static image and it has been shown that there is great inter- and intra-examiner variability when evaluating the presence or absence of dysplasia, as well as its degree [79, 45]. For this reason, the implementation of a binary classification system has been proposed to help improve agreement among pathologists, considering two levels: low-grade and high-grade [21]. Machine learning was used in clinical images to predict the high risk of dysplasia and evolution to cancer [28].

However, changes at the molecular level occur before this histological evaluation [50], so the use of immunohistochemical staining that reveals the expression of the cell proliferation biomarkers, such as ki67, could be a complementary technique to improve diagnosis and prognosis [64, 35]. Ki-67 is a biomarker of tumor growth that is available in most hospitals worldwide due to its widespread use in the classification of breast tumors, which makes it easy to use and suitable as a routine marker.

There are different approaches to measuring Ki-67 described in the literature. Typically, the area of the epithelium showing protein overexpression—evidenced by stronger cellular staining—is selected, and the involvement is assessed based on the layers (basal/parabasal/spinous). Additionally, the number of stained cells can be counted using different methods, or the staining intensity can be evaluated. Staining assessment can be performed quantitatively, semi-quantitatively, or qualitatively. Quantitative systems involve calculating an index or percentage of Ki-67–positive cells relative to the total number of cells in the field of view [76]. Semi-quantitative methods rely on visual assessment of Ki-67 staining intensity and/or the percentage of positive cells, assigning a predefined numerical scale; the measurement is not precise but visually estimated [53]. Finally, qualitative systems are based on subjective visual evaluation of the presence or absence of Ki-67 in the cells, without precise quantitative or semi-quantitative measurement, nor evaluation of staining intensity—only presence or absence is recorded [86]. Ki-67 expression can also be classified according to the epithelial layers involved. In normal epithelium, Ki-67 expression is usually confined to the basal or suprabasal layer, but in dysplasia, positive cells appear in the upper layers, as described in the literature [76].

Some studies show that the expression of ki67 staining could be used to estimate the degree of dysplasia in oral leukoplakia and the risk of malignant transformation in oral potentially malignant disorders [34, 53, 51].

1.2 Digital pathology

The cell counting of these immunohistochemical images is classically carried out visually and manually with the help of devices designed for this purpose. These manual counting techniques are very time consuming, and the experts only count a reduced number of cells (normally 100). Although immunohistochemistry is routinely used in pathology diagnosis, its standardization still has shortcomings. The interpretation of these images is based on staining intensity, but there are differing criteria for defining positive cases [71]. These limitations could be alleviated by using image analysis software, as in the case of breast cancer [55]. The development of computerized methods to analyse biopsies in order to make diagnostic and prognostic assessments, mainly based on cell morphology and architecture, is an open challenge [42]. The Aperio system[1] [54] is a comercial solution, which has been used in the immunohistochemical study of oral lesions, and offers solutions for slide glass scanning and automatic analysis of immunohistochemical staining. ImageJ [67] and QuPath[2] are two open-source digital image analyzers notable for its user-friendly design, cross-platform compatibility, and customizable functionality. ImageJ has also been used in the immunohistochemical study of oral lesions [61]. QuPath has been used for quantifying other type of IHC images [57, 38] but, to our knowledge, there is not used in the analysis of oral leukoplakia. They provide utilities to create macros or plugins, but this process requires specialized knowledge on computer language and programming skills. Other works propose automatic algorithms to analyse the image, normally counting cells or other meassures, like human lung cancer [47] or breast cancer [55]. In general, the available alternatives to analyse the immunohistochemical samples in the case of oral samples [60] have some of the following drawbacks:

1. They usually have a high cost and a low flexibility to face with artifacts in the images or differences among samples.
2. They do not allow the expert supervision before the quantification.

These limitations are also commons to other types of cancers. Hence, a recent comparative analysis of Ki-67 biomarker in IHC images of stomach and breast cancer using deep learning, conventional image analysis, and manual counting indicate the potencial of image analysis, but also the current shortcomings, evidencing the need of further research [4].

In previous works, we developed software tools that combine the automatic processing of the image with a friendly graphical user interface (GUI) to review the recognition process before image quantification in other problems. STERapp [49] performs stereological analysis from

---

[1] https://www.leicabiosystems.com/

[2] https://qupath.github.io/

histological images of fish gonads to estimate its fecundity. CystAnalyser [19] studies histological images of cystic liver and kidney in order to provide their cystic index, number of cysts and cysts profile according to their size. Both softwares use sophisticated algorithms of image analysis and machine learning to automatically recognize and classify the objects of interest in the image. When the automatic recognition provided is not suitable for experts, due to the inherent complexity of microscopic immunohistochemical images, they provide a friendly GUI, that allows the experts to review the recognition before measuring and counting them.

CHAPTER 2

OBJECTIVE AND HYPOTHESIS

Recent advances in microscope technology over the past few decades [44] have provided higher quality images and prompted the development of computerized methods for their analysis [37, 42]. In previous research, we developed the STERapp [49] and CystAnalyser [19] software and algorithms [48] to process histological images of fish gonads and mouse kidney and liver respectively. This thesis focus on the analysis of immunohistochemical (IHC) images of mouth tissue, in order to provide new skills to the diagnosis and prognosis of oral cancer.

As mention, tumor proliferation and aggressiveness is assessed by precise and accurate counting of the proportion of stained cell's nuclei within the IHC image, which can be considered as a cell segmentacion and classification problem from the computer vision perspective. This cell counting process is traditionally carried out manually by expert's pathologist, which is very time-consuming and depends on the expertise of pathologists [89]. Although, several commercial and non-commercial digital image analysis methods have been proposed to identify and quantify the Ki-67 positive cells, they have limited accuracy and more reliable methods for Ki-67 estimation in IHC images are stil demanded [4].

The objectives of the thesis include these two partial objectives:

1. To develop a new software tool to facilitate the daily work of experts in their biomedical laboratories. Chapter 5 describes and evaluates the OralImmunoAnalyser (OIA) software developed.

2. Research and proposal of new algorithms for detecting cells in IHC images, since segmentation/detection of cell nuclei in IHC images is the challenging step towards fully automated digital pathology.

Since experts do not have specialized computer knowledge, the software should meet the following requirements:

1. Provide an easy-to-use interface to draw the region of interest where you want to count cells.

2. Include image analysis algorithms to automatically detect and classify cell nuclei according to their level of staining, allowing the expert to review and modify the analysis before image quantification.

3. Count the number of cells for each staining level in each epithelial strata: basal, medium and superior.

4. Allow sharing of image analysis with other experts.

5. Export the quantification analysis of images for futher treatment with other applications.

6. The software will be accurate, trustworthy, and easy to install and learn how to use for non-expert users.

On another hand, it is known the frequently high variability present in pathological slices, due mainly to the differences in tissue specimens and variations in sample processing in the biomedical labs (staining, image acquisition process, etc), which presents significant challenges for automatic image analysis. To address these challenges, it is necessary to design a robust cell detection step that enhances the reliability of image quantification. The chapter 4 describes the **Region-Detection Algorithm** (RDA) and adapt other state of the art segmentation algorithms to detect cells in IHC images and the chapter 6 presents a fair comparison of all the cell detection methods.

CHAPTER 3

MATERIALS

The content of the following paper is partially reproduced in this chapter. The publication details are:

- Zakaria A. Al-Tarawneh, Maite Pena-Cristóbal, Eva Cernadas, José Manuel Suarez-Peñaranda, Manuel Fernández-Delgado, Almoutaz Mbaidin, Mercedes Gallas-Torreira Mercedes and Pilar Gándara-Vila. **OralImmunoAnalyser: a software tool for immunohistochemical assessment of oral leukoplakia using image segmentation and classification models**. *Frontiers in Artificial Intelligence*, Volume 7, 2024. DOI: 10.3389/frai.2024.1324410. ISSN=2624-8212

whose authors are:

1. Zakaria A. Al-Tarawneh and Almoutaz Mbaidin from the Computer Science Department, Mutah University, Karak 61711, Jordan.
2. Maite Pena-Cristóbal, Mercedes Gallas-Torreira and Pilar Gándara-Vila from Oral Medicine, Oral Surgery and Implantology Unit, MedOralRes Group of University of Santiago, Santiago de Compostela, Spain.
3. Eva Cernadas and Manuel Fernández Delgado from Centro Singular de Investigación en Tecnoloxías Intelixentes da USC (CiTIUS), Universidade de Santiago de Compostela, Spain.
4. José Manuel Suarez-Peñaranda from Pathological Anatomy Service, University Hospital Complex of Santiago (CHUS), and Department of Forensic Sciences and Pathology, University of Santiago. Both in Santiago de Compostela, Spain.

This reseach is done in colaboration with Odontology Faculty of the University of Santiago de Compostela, who provided the samples and knowledge about the pathology images. The tissue samples correspond to patients from the Department of Oral Medicine, Oral Surgery and Implantology, where they were diagnosed, clinically and histologically, of oral leukoplakia. The participants consent to the use of their clinical, histological and photographic data, treated anonymously, in the present study. This study has the approval of the Santiago-Lugo Research Ethics Committee, with registration code 2020/470.

This chapter describes the biomedical materials used in this thesis, including the selection of patients and tissue preparation (section 3.1) and the image annotation by an expert pathologist (section 3.2).

3.1 Patient's characteristics and tissue preparation

The patients to participate in this research were selected trying to achieve a maximum variability in their diagnosis, and then a wide variability among IHC images. Images were selected from patients diagnosed with oral leukoplakia for whom a biopsy report was available to confirm the diagnosis and the degree of dysplasia. Only cases with immunohistochemical Ki-67 staining documented in their records were included and their corresponding biopsy slides were accessible in the Pathology Laboratory of the University Hospital Complex of Santiago (CHUS) for photographing.

A total of 41 samples was collected distributed in the following types based on the latest classification recommended by the World Health Organization [24]: 28 cases of leukoplakia without epithelial dysplasia, four of mild dysplasia, one case of moderate, two of severe, three carcinomas in situ, two infiltrating carcinomas, one verrucous carcinoma. Patients clinically diagnosed with oral leukoplakia underwent incisional biopsy according to routine clinical practice, in which a representative tissue sample was obtained. Surgical specimens were fixed in 10% buffered formalin and sent to the Pathology Service of the University Hospital Complex of Santiago de Compostela. Once there, the samples were embedded in paraffin and stained with hematoxylin and eosin for histopathological evaluation by a pathologist (see figure 3.1).

After the study with the usual hematoxylin-eosin techniques, new sections of 4μ of different degrees of epithelial dysplasia were made for the immunohistochemical study using monoclonal mouse anti-human ki67 antigen (clone MIB-1) (Dako, Denmark), following the manufacturer's recommended instructions. Cells labeled by the antibody show a nuclear staining pattern in brown colour. Following staining, the slides were examined by an observer to select the region exhibiting the highest staining intensity. Once the area with the highest number of stained cells has been selected, a photograph was acquired using an Olympus BX51 microscope connected to an Olympus Camera DP70, using a magnification of 20X but the same can be made with a

Figure 3.1: Example of histological image (Hematoxiline-Eosine stain) presenting high grade of Dysplasia .

major magnification (see figure 3.2). The size of the acquired images is 4080 × 3072 pixels. This image was archived as TIF format. Figure 3.3 shows examples of original color images.

3.2 Cell annotation by expert pathologists

One of the objectives of this thesis is the creation of an annotated dataset of IHC images of oral leukoplakia annotated by experts. This dataset of images, called OIADB dataset, is publically available in the CiTIUS repository[1] and it will be used to measure the statistical performance of image segmentation methods in the chapter 6.

The OIADB dataset was annotated using the OralImmunoAnalyser software, developed by us, and it will be described and evaluated in the chapter 5. OralImmunoAnalyser (OIA) software provides a friendly graphical user interface to load, manage and annotate image. This OIADB dataset includes 41 original IHC images of mouth tissue with oral leukoplakia with different

---

[1] https://gitlab.citius.gal/analyser/oiadb

Figure 3.2: Olympus BX51 light microscope and Olympus DP70 camera used for imaging the biopsy slides.

levels of dysplasia, in which experts manually draw a region of interest (ROI) to count the cells. The positions of the cell's nuclei were manually annotated by expert pathologists as ground-truth data. Each cell nucleus was marked with its centroid coordinates (x, y position), and classified according to staining intensity: highly stained, low stained, or unstained nuclei. Figure 3.4 shows examples of IHC images from the dataset annotated by one expert pathologist, where the region designated for counting is highlighted in red, and the nuclei are marked as circles in different colors to indicate their staining levels: yellow for highly stained nuclei, green for lower stained nuclei, and blue for unstained nuclei.

As mentioned in the above section, patients are selected to address maximum variability in their diagnosis. Then, the OIADB image database exhibits high variability, primarily due to: the high variability inherent of the immunohistochemical process; patients present oral leukoplakia with varying levels of dysplasia; and the inclusion of images with varying cell counts for each staining level. These variability can be observed visually in the images of figure 3.4 and numerically in the table 3.1 for the selected four images, which shows the percentage of cells for each staining level.

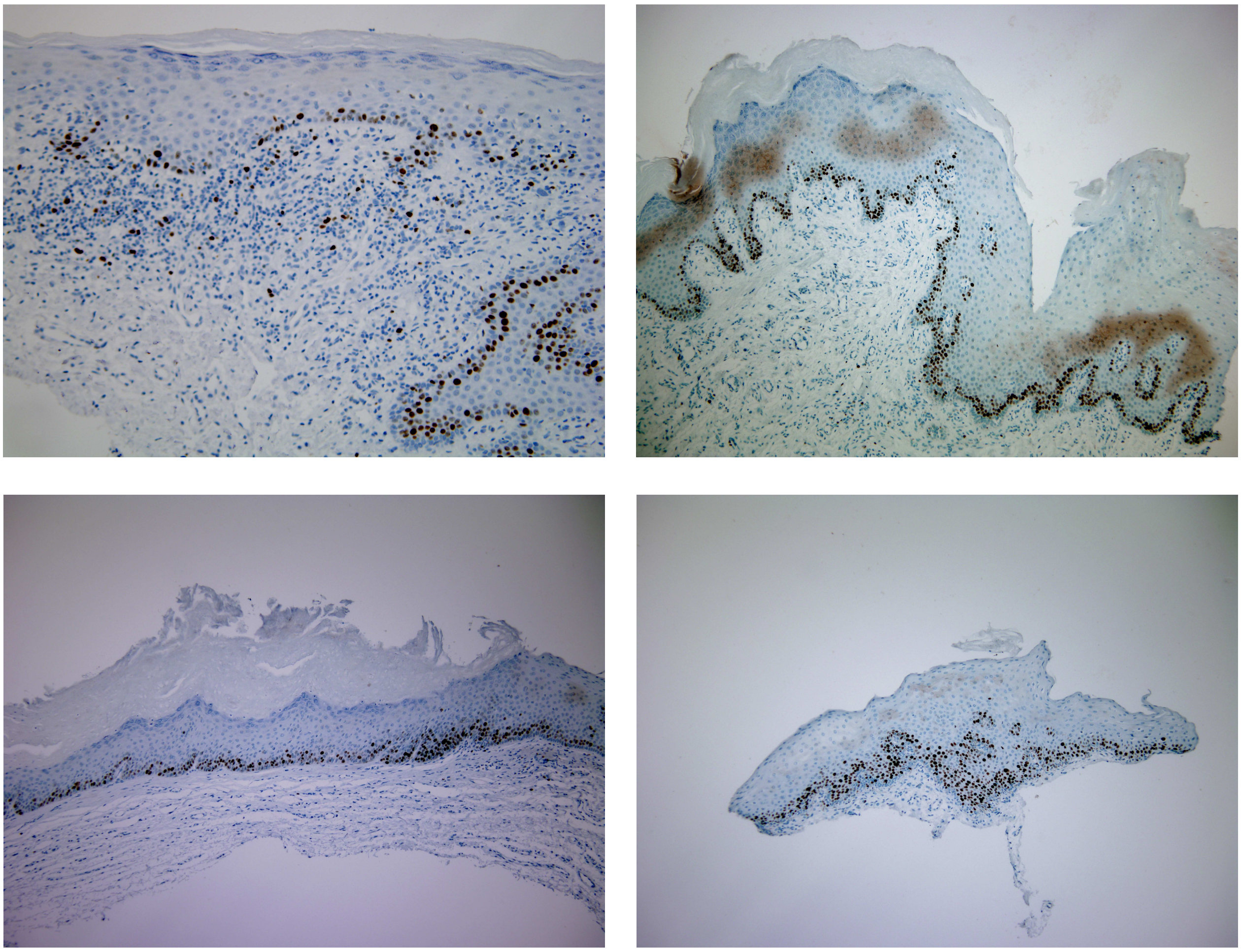


Figure 3.3: Examples of original immunohistochemical images of mouth tissue with leukoplakia.

Table 3.1: Total number of cells and percentage of them for each staining levels for some selected images of OIADB dataset.

Image	No. cells	High (%)	Low (%)	WS (%)
Image4	1378	19.3	12.1	68.6
Image12	1545	8.8	14.5	76.7
Image17	1054	59.1	25.3	15.6
Image34	1906	14.7	9.4	75.9

Globally, the images in the OIADB dataset contain between 334 and 2252 cells per image within the region of interest (drawn in red in images of figure 3.4) and an average of 1043 cells per image. According to the level of nuclei staining, the IHC image contains:

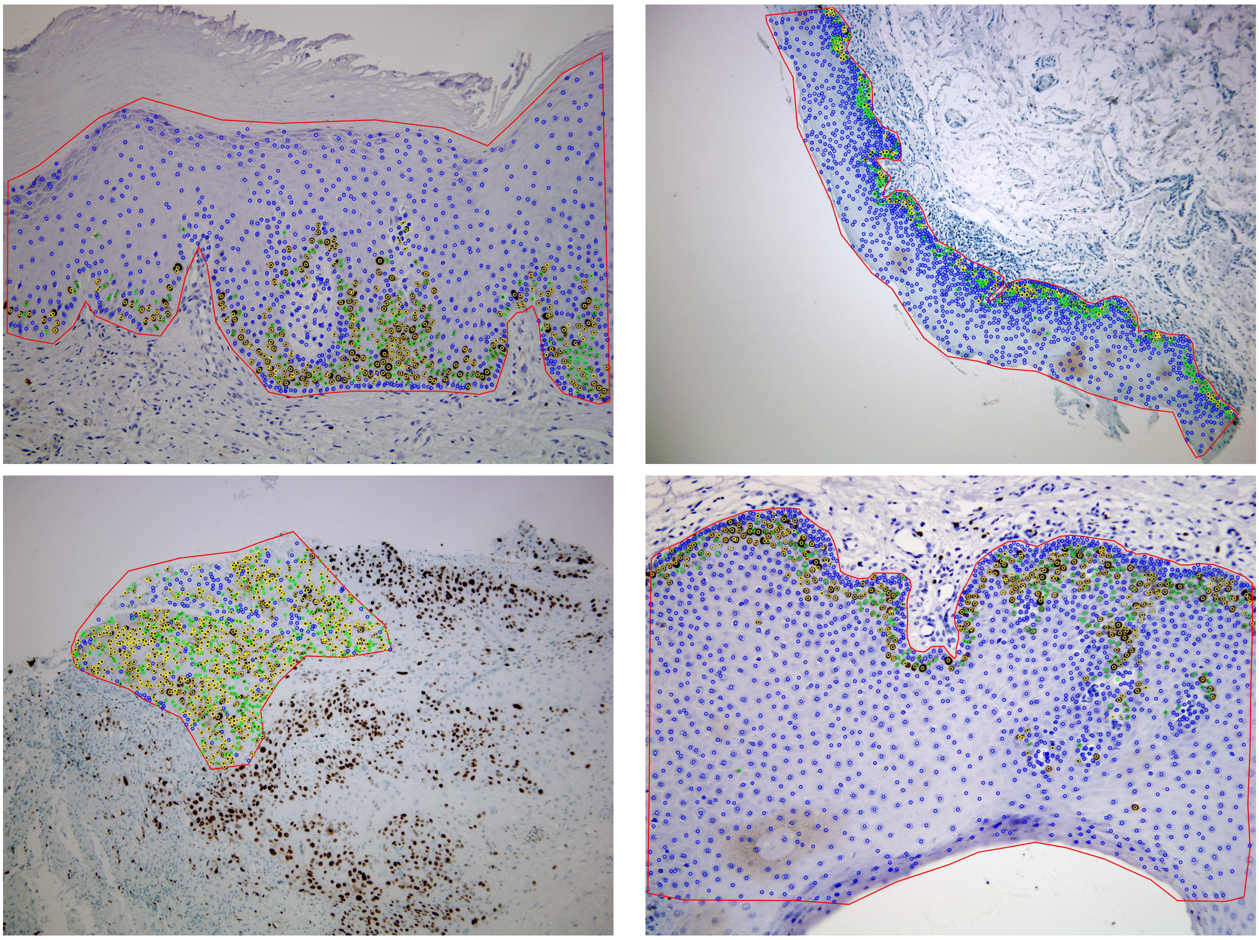


Figure 3.4: Examples of the original IHC images with the pathologist annotation overlapped for `image4` (upper left panel), `image12` (upper right panel), `image17` (lower left panel) and `image34` (lower right panel). The red color delineates the region of analysis and the yellow, green and blue are respectively the position of the cells highly stained, low stained and without staining.

- **Nuclei highly stained**: an average of 152 (14.8%) cells per image, ranging from 16 (2.6%) to 623 (59.1%).
- **Nuclei low stained**: 173 (16.4%) cells per image, ranging from 37 (3.9%) to 553 (38.8%).
- **Nuclei without staining**: 720.9 (68.7%) cells per image, ranging from 156 (15.6%) to 1542 (89%).

Detailed information on the proportion of cells belonging to different staining levels is shown in figure 3.5 for each IHC image, showing the high variability of the data set.

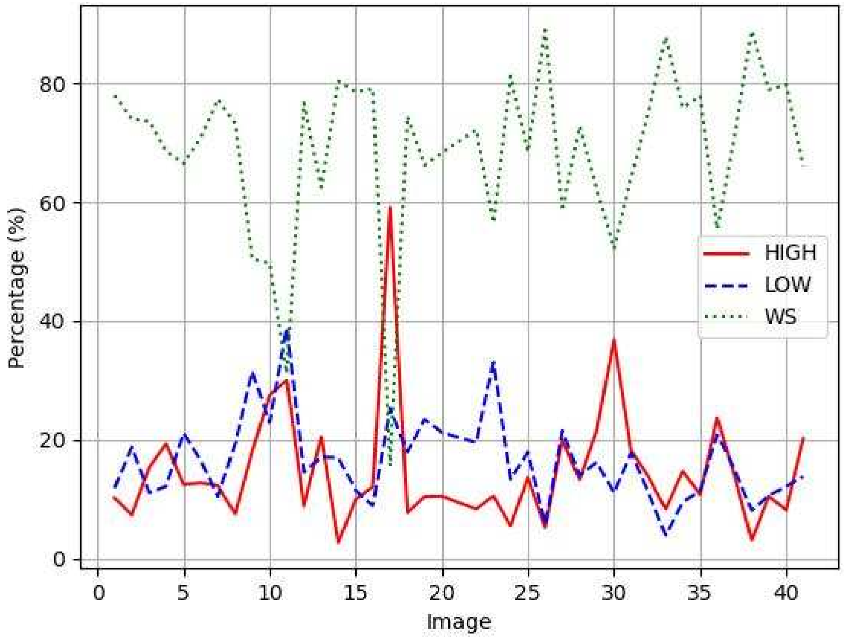


Figure 3.5: Percentage of cells for each staining level for all IHC images in OIADB dataset: Nuclei highly stained in red, nuclei low stained in blue, and nuclei without staining in green.


USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

CHAPTER 4

METHODOLOGY

The content of the following papers is partially reproduced in this chapter. The publication details are:

- Zakaria A. Al-Tarawneh, Maite Pena-Cristóbal, Eva Cernadas, José Manuel Suarez-Peñaranda, Manuel Fernández-Delgado, Almoutaz Mbaidin, Mercedes Gallas-Torreira Mercedes and Pilar Gándara-Vila. **OralImmunoAnalyser: a software tool for immunohistochemical assessment of oral leukoplakia using image segmentation and classification models**. *Frontiers in Artificial Intelligence*, Volume 7, 2024. DOI: 10.3389/frai.2024.1324410. ISSN=2624-8212
- Z.A. Al-Tarawneh, A.S. Tarawneh, A. Mbaidin, M. Fernández-Delgado, P. Gándara-Vila, A. Hassanat, E. Cernadas. **Cell Detection in Biomedical Immunohistochemical Images Using Unsupervised Segmentation and Deep Learning**. *Electronics*. Special issue *Applications of Computer Vision, 3rd Edition*, 2025, Vol. 14(18), p. 3705, ISSN 2079-9292. DOI: 10.3390/electronics14183705.

whose authors are:

1. Z.A. Al-Tarawneh, A.S. Tarawneh, A. Mbaidin and A. Hassanat from the Computer Science Department, Mutah University, Karak 61711, Jordan.

2. Maite Pena-Cristóbal, Mercedes Gallas-Torreira and Pilar Gándara-Vila from Oral Medicine, Oral Surgery and Implantology Unit, MedOralRes Group of University of Santiago, Santiago de Compostela, Spain.

3. Eva Cernadas and Manuel Fernández Delgado from Centro Singular de Investigación en Tecnoloxías Intelixentes da USC (CiTIUS), Universidade de Santiago de Compostela, Spain.

4. José Manuel Suarez-Peñaranda from Pathological Anatomy Service, University Hospital Complex of Santiago (CHUS), and Department of Forensic Sciences and Pathology, University of Santiago. Both in Santiago de Compostela, Spain.

The problem of quantifying pathological images, understood as the cell count for each level of staining, is a computer vision issue that can be addressed following different perspectives. Cells should be recognized and labeled in the image in order to be counted. The recognition of cells normally implies a segmentation process. Image segmentation is often the first essential and critical step in achieving quality outcomes in image analysis systems. The results of a segmentation process are normally represented either by: 1) a set of connected pixels to define a region or object, being visualized as binary or labeled images; 2) a set of connected and sorted pixels to define the outline of objects or boundaries; or 3) in some applications the segmentation results are also provided as lines, corners, or keypoints.

Although segmentation results can be provided by regions or edge segments, both representations are equivalent because it is easy to construct a region from its contour and obtain the contour of an existing region. From the supervision point of view, segmentation techniques can be classified into unsupervised and supervised. In the last two decades, supervised image segmentation and object detection have been developed especially using deep learning (DL). Among the DL techniques proposed in the literature [88], the different variants of the U-Net and YOLO (You Only Look Once) architectures are the most prominent for image segmentation and object detection [17, 66]. There are also proposals to integrate classical unsupervised segmentation techniques into deep learning approaches [84, 63].

Figure 4.1 graphically shows the different stages of an image analysis system for quantifying IHC images, following a fully supervised approach using deep learning techniques or a semi-supervised approach, where cell detection is performed directly on the IHC image, without supervision. The classical approach includes the following stages: 1) the original IHC image is pre-processed in order to attenuate noise or adequate the source image to the folloging step (block *image pre-processing* in figure 4.1); 2) the cells are segmented or detected in the IHC images using unsupervised segmentation techniques (block *cell detection* in figure 4.1); 3) cell is classified into different staining levels, normally using supervised techniques (block *cell classification* in figure 4.1); and 4) the cells are counted (block *cell counting* in figure 4.1). The deep learning techniques can be used to detect the cells in the IHC images (block *deep learning* in figure 4.1) and then employ the same module to classify the cells as in the classical approach. But, they can also be used both to cell detection and classification. However, this thesis focuses solely on the comparison of unsupervised and supervised segmentation techniques to the task of cell detection and it leaves the use of deep learning to jointly detect and classify cells for future work.

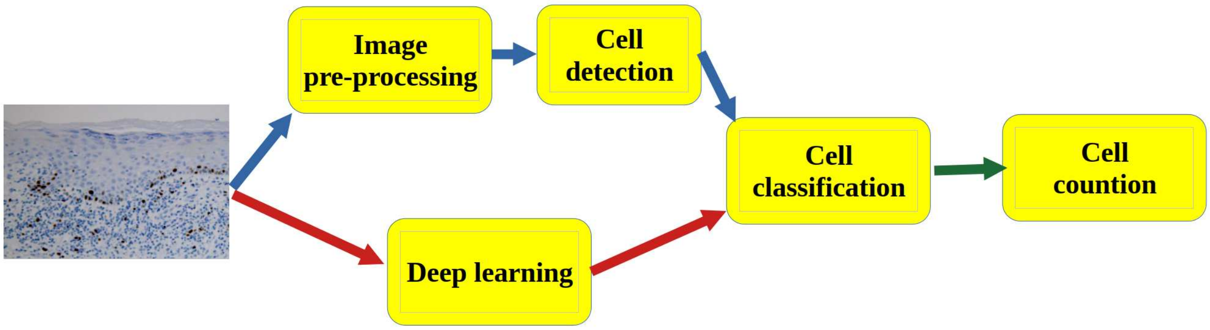


Figure 4.1: Typical block diagram of a computer vision system for IHC image quantification.

Given the objectives of the thesis, this chapter is organized as follows: section 4.1 reviews the state-of-art of unsupervised segmentation techniques and proposes new algorithms to the specific case of IHC images of oral cancer; section 4.2 address the design of the cell classification module; section 4.3 describes the method to count the cells into each stracta of epithelium; and section 4.4 details the application of deep learning techniques to our problem.

4.1 Unsupervised segmentation methods

The classical segmentation approaches can be categorized into edge-based, region-based, and hybrid strategies [33]. Region-based segmentation involves partitioning an image into homogeneous areas of connected pixels by applying homogeneity or similarity criteria among the image pixels. Pixels are considered similar on the basis of certain computed properties, such as color, intensity, and/or texture, among others. These techniques assume that the partitions formed represent regions, objects, or meaningful parts of the image. An edge is considered the boundary of the object, which is manifested as a change or discontinuity in a number of pixels along a certain direction. Edge-based segmentation involves identifying the locations of pixels that correspond to the boundaries of objects present in the image. Then, subsection 4.1.1 reviews the state-of-art of unsupervised segmentation algorithms and subsection 4.1.2 describes the proposed algorithm to detect cells in IHC images of oral cancer.

4.1.1 State of the art methods

The common state-of-art region-based techniques includes thresholding, region growth, and clustering. Thresholding is the simplest and fastest technique in which the image is partitioned into sets of pixels with values lower than or higher than a specified threshold [9]. Also, it can be applied locally or globally, but it often yields unsatisfactory segmentations in images with complex backgrounds. Although different methods to calculate the optimal threshold have been proposed in the literature [43], the outcome is normally unsatisfactory. The region growth methods group pixels or subregions into larger regions based on predefined growth criteria [25]. Among other drawbacks, these methods can be very time-consuming when segmenting large images.

Clustering methods also divide an image into homogeneous regions based on some image property such as gray level, color, texture, and others [8]. The K-means algorithm is the most popular method due to its simplicity and its very low execution time [33]. Each pixel in the image is considered a pattern, and the K-means algorithm subdivides these patterns into K components by calculating the average of patterns for each component. Its main drawbacks are its sensitivity to noise in images and the fact that the selection of the K value affects the quality of the segmentation. The use of image pre-processing techniques may reduce image noise and improve segmentation results. Different improvements on the original algorithm have been addressed in the literature [72]. For example, the MeanShift algorithm [18] does not require the number of clusters as input and calculates this value indirectly using a parameter called the kernel bandwidth.

Among the edge-based segmentation methods, first and second derivatives are normally used to determine the edge position, marking the pixel as white if it belongs to the boundary of an object and black otherwise. Many edge detectors are available in the literature. They are very sensitive to the presence of noise in the image and cannot detect edgeless objects. However, they have the advantage of requiring a low computation time for application. Canny filter [11] is the most popular due to its trade-off between edge detection and noise attenuation. Edge detectors find the position of pixels in the boundaries of objects, but these pixels must be linked into closed chains of pixels to identify the objects. Recently, Multi-Scale Canny Filter (MSCF) is a multi-scale approach, based on Canny filter, that provides closed contours, specifically proposed to detect cells in microscopy images [48].

The watershed transform is a segmentation method that interprets a grayscale image as a topographical relief, where each pixel is associated with an altitude that corresponds to its intensity. Thus, the pixel values are equivalent to the surface heights defined on a lattice. The main idea is that, as rainwater naturally flows down the steepest paths, the watershed algorithm will eventually fall into a number of domains of attraction. The watersheds are the dividing lines of these domains, called catchment basins, which are the natural segmenting contours of the landscape. The watershed transform is applied to the gradient image to achieve image segmentation. Various algorithms have been proposed to compute the watershed based on different principles. Some popular approaches are the immersion approach and toboggan simulation [46].

Active contour-based models, also called snakes, have long been a popular group. They transform the image segmentation problem into an energy minimization problem, where the energy functional specifies the segmentation criterion, and the unknown variables describe the contours of different regions [15]. These methods can be divided into parametric and geometric snakes, depending on whether their representation is explicit using parametric curves or implicit using level-set algorithms [20]. Snakes are considered edge-based models when the parameterized curve of the snake evolves until it reaches high-gradient areas or edges. Since edge-based approaches use gradient functions to stop the evolution of the curve, they struggle with identifying weak edges, edges in noisy images, or capturing regions separated by textures. In contrast, geometric snakes are considered region-based models as they employ homogeneity properties to decompose the image domain into different regions, thereby overcoming the aforementioned drawbacks of edge-based approaches. Many of these methods require an initial curve, and its placement on the image plays an important role in the final segmentation, which is a drawback for their application to our problem. The Chan-Vese model is the most representative model, as it can be applied as a region-based and edge-based technique [15].

Image segmentation can also be tackled using graph cuts models, in which the image pixels are the nodes of a weighted graph, and the segmentation problem is translated to finding minimum cut in the constructed velocity graph via a maximum flow computation. Several graph-based methods have been proposed for unsupervised segmentation, including: ratio cut [78], normalized cuts [73], and minimum spanning tree-based methods like the efficient Felzenszwalb-Huttenlocher segmentation algorithm [26].

Some recent hybrid segmentation approaches combine different unsupervised segmentation methods, typically belonging to various families of algorithms. For example, there are combinations of clustering and watershed to segment white blood cells from microscopy images [31], as well as the integration of region-growing and watershed for cell segmentation in fluorescence microscopy images [29], among others. However, such a combination may increase segmentation accuracy, but will likely consume more time.

Taking into account the large size of IHC images and the availability of public code, this thesis compares the following image segmentation methods:

1. **OIA-EDA**: It use the Multi-Scale Canny Filter (MSCF) [48], that applies the Canny filter at different scales, i.e., using various values of Gaussian spread (σ), and different values of hysteresis thresholds, followed by an edge linking step to produce closed contours. MSCF was configured to be included in the software OralImmunoAnalyser with one scale ($\sigma = 4$) and one pair of thresholds, the rates 0.3 and 0.7 to the low and high thresholds respectively. This method is called **EDA** in the OralImmunoAnalyser software (see chapter

5) and **OIA-EDA** in the results chapter 6. The minimum diameter of the cells to detect is also set to 20 pixels for the OIDB dataset.

2. **FH**: The Felzenszwalb-Huttenlocher algorithm [26] was selected among the graph cut models due to its efficiency and the availability of an implementation in the segmentation module of the Python scikit-image library[1]. We use the `felzenszwalb` function, configured with default values (scale=1.0 and sigma=0.8), except for the minimum component size min_{size}, which is set to 400 pixels. Since the minimum diameter of the cells provided by the experts is 20 pixels for the OIADB dataset, we assume the minimum size of the components as a square of size minimum diameter, $20 \times 20 = 400$ pixels.

3. **ChV**: Among the active contour models, the Chan-Vase algorithm [15], implemented also in the Python scikit-image library, is selected and which does not need to provide initial seeds to its application. Specifically, we use the `morphological_chan_vese` function to apply Morphological Active Contours without Edges, called MorphACWE, and `Morphological_geodesic_active_contour` function to apply Morphological Geodesic Active Contours, called MorphGAC. Both methods are applied using default configurations (using 500 iteractions), and the original images were pre-processed with a Gaussian filter of size (3×3 pixels) to reduce noise in the images.

4. **Clustering**: The available clustering algorithms used are: 1) the popular K-means algorithm implemented by the `kmeans` function in the OpenCV library[2] using the original image in the RGB and Lab color spaces, designed as kmeans-RGB and kmeans-LAB respectively; and 2) the `slic` function of the segmentation API of scikit-image library, which implements the simple linear iterative clustering (SLIC) superpixels method [1], called SLIC. In the kmeans-RGB approach, the `kmeans` function is applied on patterns including the RGB signature of each pixel in the image using four clusters (one for each staining level of nuclei and another for sample background). Initial centers for the clusters are taken randomly, and we use the default configuration for the remaining parameters (the algorithm stops when ten iterations of algorithm is ran or an accuracy of epsilon = 1.0 is reached). A signature containing the three values of the channels of the Lab color space is used for the kmeans-LAB approach.

Although there is available code to implement the watershed algorithm in the OpenCV library, we rejected its application to our problem because it is an iterative algorithm that requires

---

[1] https://scikit-image.org/

[2] https://opencv.org/

placing markers on the local minima in the image to set the starting points, which must correspond precisely to the nuclei positions.

4.1.2 Region detection approach (RDA)

Since the area of cells is not relevant for our objectives, the segmented regions are managed as points. We propose a region-based algorithm, called **Region Detection Algorithm** or RDA, containing the following stages: 1) detect the highly stained regions in the image; 2) split highly stained regions when some cells are joined; and 3) detect low stained and no-stained cells. We use the color space Lab, also known as CIELAB [33], to process the image due to: a) its better perceptual linearity, compared to the color space RGB of image acquisition; and b) its robustness to illuminance variances [13].

Let I_{in} be the RGB immunohistochemical image, d_{min} and d_{max} the minimum and maximum diameters of the cells to be detected and TP an external option to indicate the interested stained levels. Valid values to TP argument are **HIGH**, **LOW** or **WS** to highly stained, low stained and without staining respectively. Algorithm 1 summarizes the proposed RDA method and figure 4.2 shows visual examples of intermediate stages in the image processing.

As mentioned, the original RGB image I_{in} is converted to the Lab color space giving the I_{lab} image. Then the channels L and b of I_{lab} are multiplied and the product is normalized to the range [0,255], giving the image I_{grey}. The L channel of Lab image represents the image intensity and the b channel the image yellowness. The multiplication of both channels gives an image with the lowest levels in the more stained regions (I_{grey} in figure 4.2). Thresholding is applied on I_{grey} in order to create the binary image I_{bin}. The selection of the optimal threshold is critical, and it is commonly selected by trial and error in many applications. In our case, this value is selected from the statistical properties of each image using the multi-level Otsu's method [59], implemented by the `Otsu3Thres` function in algorithm 1. The optimal threshold considered was the lowest value t_1 of the three calculated by `Otsu3Thres`(I_{grey}). The binary image I_{bin} (see figure 4.2), containing white in the more stained pixels and black in the remaining pixels, is calculated thresholding the image I_{grey} using t_1 (in this case, $t_1 = 69$) and taking the values less than t_1 as white. Black and white pixels in the I_{bin} image represent background and highly stained cells, respectively.

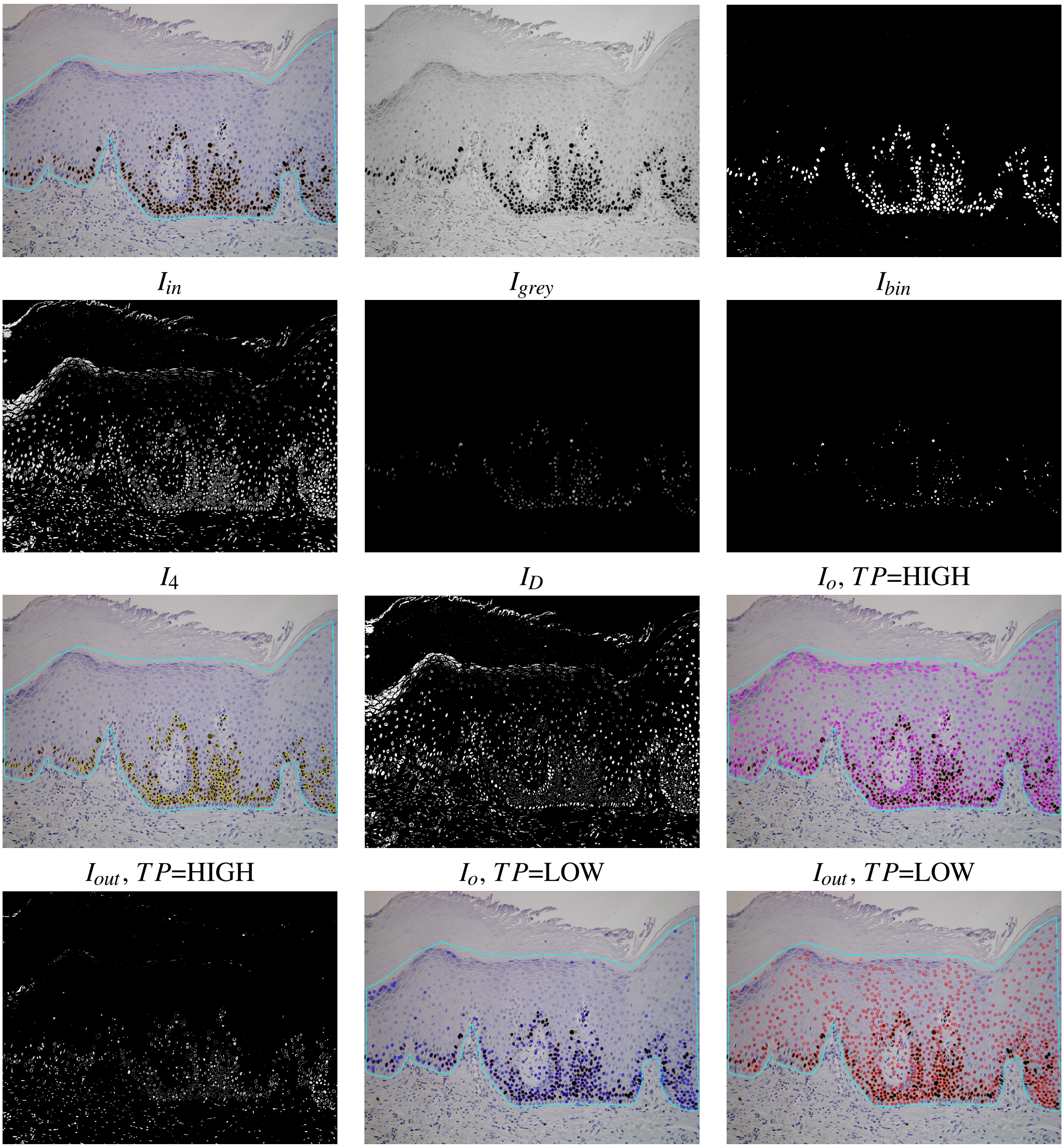
Afterwards, k-means clustering [23], implemented by the `kmeans` function in algorithm 1, is used to group the pixels in the color image I_{lab} according to color similarity into four clusters. The I_{bin} image is used as seed for the cluster prototypes, representing the two extremes that must be discriminated. The application of k-means returns the image I_4 whose pixels are labeled by their cluster (4 labels in our case), shown in figure 4.2. The 0-th cluster represents

Algorithm 1: Pseudocode of `RDA` algorithm to detect cells in immunohistochemical images.

```
Algorithm: $\mathcal{P}^{TP}$=RDA($I_{in}, d_{min}, d_{max}, TP$)
Data: $I_{in}$: original RGB image; $d_{min}/d_{max}$: min./max. cell diameter; $TP$: user option
Result: $\mathcal{P}^{TP}$: set of cells detected
$L_{lab} \leftarrow$ Lab image of $I_{in}$
$I_L, I_b \leftarrow L$ and $b$ channels $I_{lab}$
$I_{grey} \leftarrow I_L \cdot I_b$ normalized to range [0,255]
$t_1, t_2, t_3 \leftarrow$Otsu3Thres$(I_{grey})$
$I_{bin} \leftarrow 255$ if $I_{in}(x,y) < t_1$ and 0 otherwise
$I_4 \leftarrow$ kmeans$(I_{lab}, 4, I_{bin})$
if $TP$ = HIGH then
    $I_{clu1} \leftarrow 255$ if $I_4(x,y) = 1$ and 0 otherwise
    $I_D \leftarrow$ normalize to [0,255] the image transformDistance$(I_{clu1})$
    $t'_1, t'_2, t'_3 \leftarrow$ Otsu3Thres$(I_D)$
    $I_o \leftarrow 255$ if $I_D(x,y) > t'_3$ and 0 otherwise
else if $TP$ = LOW then
    $I_{clu2} \leftarrow 255$ if $I_4(x,y) = 2$ and 0 otherwise
    $I_o \leftarrow$ close filter of $I_{clu2}$ using masksize 3
else if $TP$ = WS then
    $I_{clu3} \leftarrow 255$ if $I_4(x,y) = 3$ and 0 otherwise
    $I_o \leftarrow$ close filter of $I_{clu3}$ using masksize 3
end
$\{r_i\} \leftarrow$suzuki$(I_o)$; $\{d_i\} \leftarrow \{$diameter$(r_i)\}$
$\mathcal{P}^{TP} \leftarrow \{p_i \mid d_{min} < d_i < d_{max}\}$ // $p_i$: mass center of $r_i$
```

the background and includes pixels in I_{lab} with color similar to the pixels in I_{lab} that are black in I_{bin}. Analogously, the 1st cluster represents the highly stained regions and contains pixels in I_{lab} with color similar to the pixels in I_{lab} that are white in I_{bin}. Finally, the 2nd and 3rd clusters may represent low or non-stained regions, including pixels with intermediate color values in I_{lab} that are less similar to the seed pixels (both black and white). The binary images I_{clu1} (highly stained regions) and I_{clu2}, I_{clu3} (low and non-stained regions), are created with values 1 in the pixels of the corresponding cluster and 0 in the remaining pixels. Each cluster is associated to a different value of the option TP in `RDA`: I_{clu1} with TP=HIGH, I_{clu2} with TP=LOW and I_{clu3} with TP=WS. These binary images are the I_O images in figure 4.2 for TP=LOW and TP=WS options, after applying a close morphological filter with masksize three to remove small holes.

Since the highly stained cells are frequently joined, we apply the distance transform algorithm (function `transformDistance` in algorithm 1) to the image I_{clu1}, which provides the derived representation of the binary image, where the value of each pixel is replaced by its dis-



I_{in} I_{grey} I_{bin}

I_4 I_D I_o, TP=HIGH

I_{out}, TP=HIGH I_o, TP=LOW I_{out}, TP=LOW

I_o, TP=WS I_{out}, TP=WS I_{Out}

Figure 4.2: Examples of the automatic processing of immunohistochemical images using the proposed `RDA` for different types of processing (see algorithm 1 for the meaning of I_{in}, I_{grey}, I_{bin}, I_4 and I_O). The I_{out} images show the cells detected overlaped to the original image for each option TP and the cell detection of all options merged (lower right image I_{Out}).

tance to the nearest background pixel. This resulting image I_D (shown in figure 4.2) is normalized to the range [0,255] and thresholded using the highest value provided by `Otsu3Thres`(I_D) t_3'

(in this example t'_3=101). The I_o image for TP=HIGH is 255 if $I_D(x,y) > t'_3$ and 0 otherwise (see figure 4.2). The contours of the cells, or external regions, are extracted from the thresholded image I_o (see figure 4.2) using the algorithm proposed in [68], implemented by `suzuki` function in algorithm 1. A detected contour r_i is considered a true cell if its diameter d_i is between the minimum and the maximum diameter of the cells, i.e. $d_{min} < d_i < d_{max}$. In this case, the detected cell is the centroid (mass center) of the region. Finally, the sets $\mathcal{P}^{HIGH}$, $\mathcal{P}^{LOW}$ and $\mathcal{P}^{WS}$ for TP=HIGH, TP=LOW and TP=WS, respectively, are the sets of cells (represented as points) detected applying the Suzuki algorithm that verify $d_{min} < d_i < d_{max}$. The I_{out} image in figure 4.2 shows the set of cells $\mathcal{P}$ overlapped to original image inside the region of analysis drawn by user. In this example, the 96.25% of cells was correctly detected.

4.2 Cell classification

Once the cells are detected on the image, they must be classified according to their staining level. Since only a limited number of staining levels are normally used in these type of problems, the supervised machine learning model will be a classifier. Although many classifiers have been proposed in the literature, the support vector machine (SVM) with radial basis function (RBF) kernel is one of the best-performing machine learning models for classification [27].

In the training process, SVM model learns from a collection of examples (cells) to predict their categories. Typically, images are not used directly to train the SVM. Instead, features are extracted from the image patches that contain the cell. So, each cell is represented by a set of numeric characteristics calculated in the following way:

1. A square region centered in the cell, of size the minimum diameter, is extracted from the original I_{lab} image.
2. The average value of the three channels L, a and b of the I_{lab} image over this region is calculated. In the training, the model learns to predict the cell category using these three average values over the cell square for all the cells in the collection.

The performance of the classifier is evaluated by the Cohen kappa statistic [12], which measures the agreement between the true and predicted category excluding the agreement by chance. Kappa (in %) is defined as:

$$
\begin{aligned}
\text{kappa} &= 100\frac{p_a - p_e}{s - p_e} \qquad (4.1)\\
p_a &= \sum_{i=1}^{C} N_{ii}\\
s &= \sum_{i=1}^{C}\sum_{j=1}^{C} N_{ij}\\
p_e &= \frac{1}{N^2}\sum_{i=1}^{C}\left(\sum_{j=1}^{C} N_{ij}\right)\left(\sum_{j=1}^{C} N_{ji}\right)
\end{aligned}
$$

where N_{ij} is the number of cells of category i and that are assigned by the SVM to category j, C is the number of categories and N is the number of cells.

SVM is a classifier that has tunable hyper-parameters, i.e., parameters whose values must be specified previously to training, that often have a strong influence on the classifier performance. Then, it is a good practice to try several values for each hyper-parameter in a trial-and-error procedure, and to select the value that provides the best performance on a separate data set. This method is called "grid search" tuning. In the SVM with RBF kernel classifier, it is often tunned the values of the parameters regularization, λ, and kernal spread, σ. Specifically, the values of $\lambda = \{2^{2i-7}\}_{i=1}^{10}$ and $\sigma = \{2^{-(i+1)/2}\}_{i=-15}^{0}$ are used. For each combination of hyper-parameter values, the SVM is trained using the K-fold cross-validation methodology with $K = 4$, so that $K - 1 = 3$ folds are used to train the SVM, and the remaining fold is used to calculate the kappa of the trained SVM. All k folders maintain class populations. The training and prediction are performed K times, rotating the folds each time. For example, trial 1 uses folds 1 and 2 for training, fold 3 for validation and fold 4 for test (see table 4.1). Trial 2 uses folds 2 and 3 for training, fold 4 for validation and fold 1 for test, and analogously for trials 3 and 4. Therefore, the training, validation, and test sets in all trials include data from all classes, ensuring the representativeness of each class. Then, Kappa value is averaged over the K test folds. The process is repeated for all the combinations of hyper-parameter values, and the one that achieves the highest average kappa is selected. Finally, the SVM is trained over the whole collection of cells, using the selected combination of hyper-parameter values, and then it is ready to predict the category for new cells.

K=4	Trial 1	Trial 2	Trial 3	Trial 4
Train	T_1={1,2}	T_1={2,3}	T_1={3,4}	T_1={4,1}
Validation	V_1={3}	V_1={4}	V_1={1}	V_1={2}
Test	S_1={4}	S_1={1}	S_1={2}	S_1={3}

Table 4.1: Distributions of folds in training, validation and test for each trial.

4.3 Cell counting

Once the image is analysed (drawn the region of analysis, detected and classified the cells in that region), we want to count the cells. Although the experts are interested in counting the number of cells for each staining level in each epithelial strata, we are not aware of any available tool that performs this count. The systems normally count the cells globally in a whole IHC image, a rectangular area or an user-defined region. However, these measurements allow clinicians to compare the positivity of ki67 among layers and to study its relationship with the degree of epithelial dysplasia. Specifically, it allows clinicians to compare the cell proliferation among epithelium layers, similarly to the analysis of epithelial dysplasia (mild-moderate-severe), where architectural and cytological changes are analyzed layer by layer. The basal expression of ki67 can be related to a physiological proliferative activity. In cases of dysplasia, the expression of this biomarker increases and can manifest beyond the basal third.

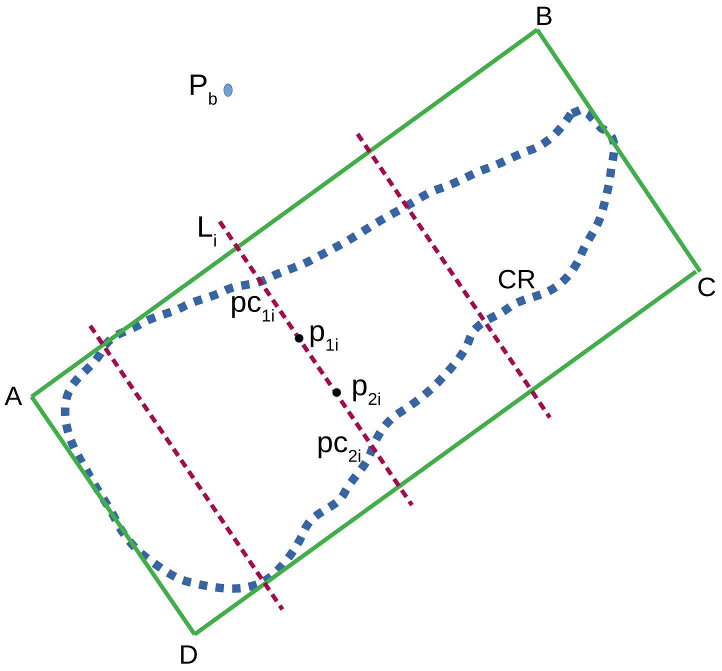


Figure 4.3: Scheme of the calculation of the epithelial strata regions.

In this thesis, a method to count the cell within three layers (basal, medium and superior) was designed. Let *CR* be the set of points (x_i, y_i) that define the region of analysis drawn by the expert and let $P_b = (x_p, y_p)$ be a point marked by the user to indicate the side of the region *CR* where the basal layer is located (outside *CR*). The algorithm to calculate the layer regions, illustrated in figure 4.3, includes the following steps:

1. Calculate the minimum enclosing rectangle R of the region CR defined by the points A, B, C and D.
2. Calculate the distance between P_b to each segment of the rectangle (line segments AB, BC, CD and DA) in order to determine the closest segment to P_b.
3. Let, for example, AB be the closest segment, then we take the sides of rectangle that are perpendicular and adjacent to AB, in our example are BC and DA.
4. Divide the line segments BC and DA into three equal parts, corresponding to the basal, medial and superior layers.
5. Create the list of adjacent rectangle vertices, in our example (A,B,C,D), and calculate the four points on the contour CR that are closest to these four vertices, saving the indices of these points in the contour CR.
6. Create the subregions basal R_b, media R_m and superior R_s and copy the contour points CR to the subregions.
7. Generate the boundary points among subregions taking a set of N_L lines L_i, with $i = 1, \ldots, N_L$, parallel to the line segment BC between the line segments AB and CD. For each line L_i, calculate its two cross points, pc_{1i} and pc_{2i} in the figure 4.3, with contour CR. Then, calculate the two points, p_{1i} and p_{2i}, of L_i that divide the segment between pc_{1i} and pc_{2i} in three pieces of equal length. The points p_{1i} and p_{2i} are added to the contours of the subregions basal R_b (only p_{1i}), media R_m (between p_{1i} and p_{i2}) and superior R_s (only p_{i2}).

4.4 Deep learning

The most prominent deep learning models to image segmentation or object detection are U-Net architecture and YOLO model. The application of these models to our problem of cell detection are detailed in the subsections 4.4.1 and 4.4.2 respectively.

4.4.1 U-Net architecture

Cell segmentation and detection on histopathology [85] or RGB immunofluorescence [52] images are commonly performed using the U-Net deep neural network [66] and its variants. U-Net consists of two primary components: an encoder and a decoder. The encoder processes the input image, gradually reducing its spatial dimensions while extracting higher-level features, effectively compressing the image into a lower-dimensional representation. The decoder then reconstructs the image from this compressed representation, upsampling it to produce the final output, which varies depending on the task. For example, in semantic segmentation, the input is an image, and the output is a segmentation mask that highlights the regions of interest [66]. Figure 4.4 illustrates the typical diagram of the U-Net architecture.

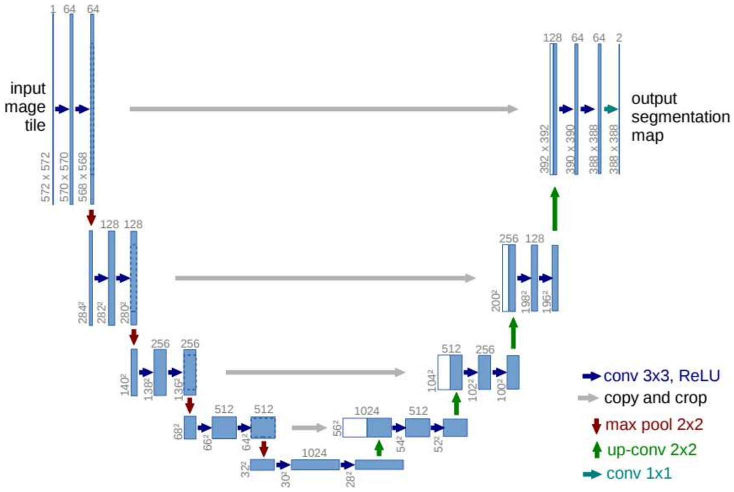


Figure 4.4: Typical diagram of U-Net architecture.

U-Net allows the selection of a specific encoder from a wide range of convolutional neural networks (CNN) available in the timm library [83], a popular repository that provides access to numerous CNN architectures pre-trained on large-scale datasets such as ImageNet. In our case, we use the segmentation-models-pytorch (SMP) module[3] with the pre-trained EfficientNet-B0 model as the encoder for our U-Net architecture. EfficientNet-B0 is known for its strong performance-to-parameter ratio, enabling efficient feature extraction with relatively low computational cost. Using pre-trained weights, the model benefits from transfer learning, which accelerates convergence and improves performance, especially in scenarios with limited labeled data.

Like most deep learning architectures, U-Net requires a substantial amount of training data to perform effectively. To mitigate data scarcity in our methodology, each image is divided into non-overlapping patches of size 224×224 pixels, which is the model input size. This process increases the number of training samples. These patches serve as input to the proposed U-Net model. The corresponding outputs, called targets, are heatmaps in which the cell locations are encoded using 2D Gaussian functions ($\sigma = 7$ and peak 5) centered on the ground-truth cell centroids. This heatmap representation enables the model to learn the spatial location of cells without relying on explicit segmentation. The U-Net model is trained using AdamW using weight decay = 0.001.

---

[3] https://pypi.org/project/segmentation-models-pytorch/0.0.3/

To further enrich the dataset and reduce the risk of overfitting, image augmentation techniques are applied [74]. These techniques introduce slight variations to the input patches before feeding them into U-Net, ensuring that it does not see the exact same images across training epochs. This strategy is a well-established practice in training deep learning models and helps improve generalization performance. In this work, the applied augmentations include color jittering (adjusting brightness, contrast, saturation, and hue), horizontal and vertical flipping, Gaussian blurring with a kernel size of 5 pixels and a variable sigma ranging from 0.01 to 2, and the addition of Gaussian noise with a standard deviation sampled from the range 0.01 to 0.2. These transformations are applied using the torchvision transforms and are composed as part of the training data pipeline. Figure 4.5 shows an example of a single input patch with its corresponding heatmap.

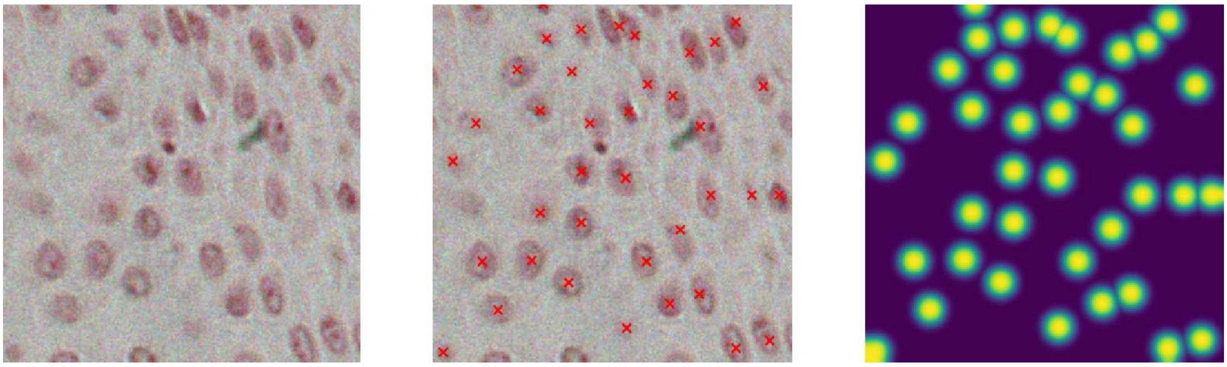


Figure 4.5: Original image patch (left panel); ground-truth centroids overlapped to the image in red (middle panel); and the corresponding target heatmap (right panel).

Both input patches and output targets are normalized to the range [0, 1] to ensure numerical stability during training and to facilitate faster model convergence [39]. In particular, normalizing the target heatmaps is essential for stable loss computation and contributes to more effective optimization by keeping the target values within a consistent and predictable scale.

To effectively train the U-Net network for accurate point detection using heatmap regression, we defined the following composite loss function, $\mathcal{L}_{\text{total}}$, that integrates a weighted Binary Cross-Entropy (BCE) loss with a spatially-weighted Mean Squared Error (MSE) loss:

$$\mathcal{L}_{\text{total}} = \lambda_{\text{BCE}} \cdot \mathcal{L}_{\text{BCE}}(z,y) + \lambda_{\text{WMSE}} \cdot \frac{1}{N} \sum_{i=1}^{N} (1+\alpha y_i) \cdot (\sigma(z_i) - y_i)^2 \tag{4.2}$$

where z_i denotes the predicted logit at pixel i, $y_i \in [0,1]$ is the corresponding ground-truth heatmap value, $\sigma(z_i) = \frac{1}{1+e^{-z_i}}$ represents the sigmoid activation applied to the logit at pixel i, N is the total number of pixels in the heatmap, λ_{BCE} and λ_{WMSE} are weighting coefficients for the BCE and weighted MSE loss components respectively, and α is the spatial weighting parameter. The coefficients were empirically chosen as $\lambda_{\text{BCE}} = 0.3$ and $\lambda_{\text{WMSE}} = 0.7$ to balance classification and localization performance. The BCE loss $\mathcal{L}_{\text{BCE}}$ operates directly on logits and incorporates positive class weighting to mitigate class imbalance between sparse foreground peaks and dominant background. The weighted MSE component applies spatial weighting where each pixel is weighted by $(1 + \alpha y_i)$ with $\alpha = 10$, increasing loss contribution from regions with higher target values to encourage sharper, more localized heatmap peaks centered on annotated points.

This parameter determines the contribution of the weighted MSE component. A higher value (0.7) emphasizes spatial accuracy, encouraging the model to produce precise and smooth heatmaps that match the ground-truth intensity distribution around annotated points. The α value controls the degree of spatial weighting in the MSE loss. The term $1 + \alpha y_i$ increases the importance of pixels with higher target values (typically near keypoint centers). A value of 10 strongly emphasizes these areas, helping the model focus on accurately localizing keypoints while reducing the influence of distant background pixels.

Despite the logical selection of the coefficients $\lambda_{\text{BCE}} = 0.3$, $\lambda_{\text{WMSE}} = 0.7$ and $\alpha = 10$, we acknowledge that these hyperparameters were not systematically optimized and should be systematically optimized in future applications of this loss function to potentially achieve better performance.

After training the model, a test image is passed as a set of 224×224 patches to get the predicted heatmap. Figure 4.6 shows visual examples of the performance of the U-Net model. To locate the centroids, the predicted heatmaps are first smoothed using a Gaussian filter to suppress spurious local maxima. Then, a maximum filter is applied to identify peaks in the heatmap. Only peaks with values above a predefined threshold (default value of 0.5 is used) are retained. To address multiple detections for the same cell, we remove redundant peaks that fall within a specified perimeter (20 pixels, which is the minimum diameter of the cells in the OIADB dataset). When multiple peaks are found within this perimeter, their mean position is taken as the final centroid.

4.4.2 YOLO model

Another neural network widely used for object detection is YOLO [14], renowned for its high speed and efficiency in detecting objects and predicting their corresponding bounding boxes. Its ability to perform detection in real time makes it particularly suitable for various time-sensitive applications [88, 17]. In the current work, we employ YOLOv8 (nano variant), the latest version of the YOLO architecture, implemented using the Ultralytics library [4] (version 8.3.0). The input

---

[4] https://github.com/ultralytics

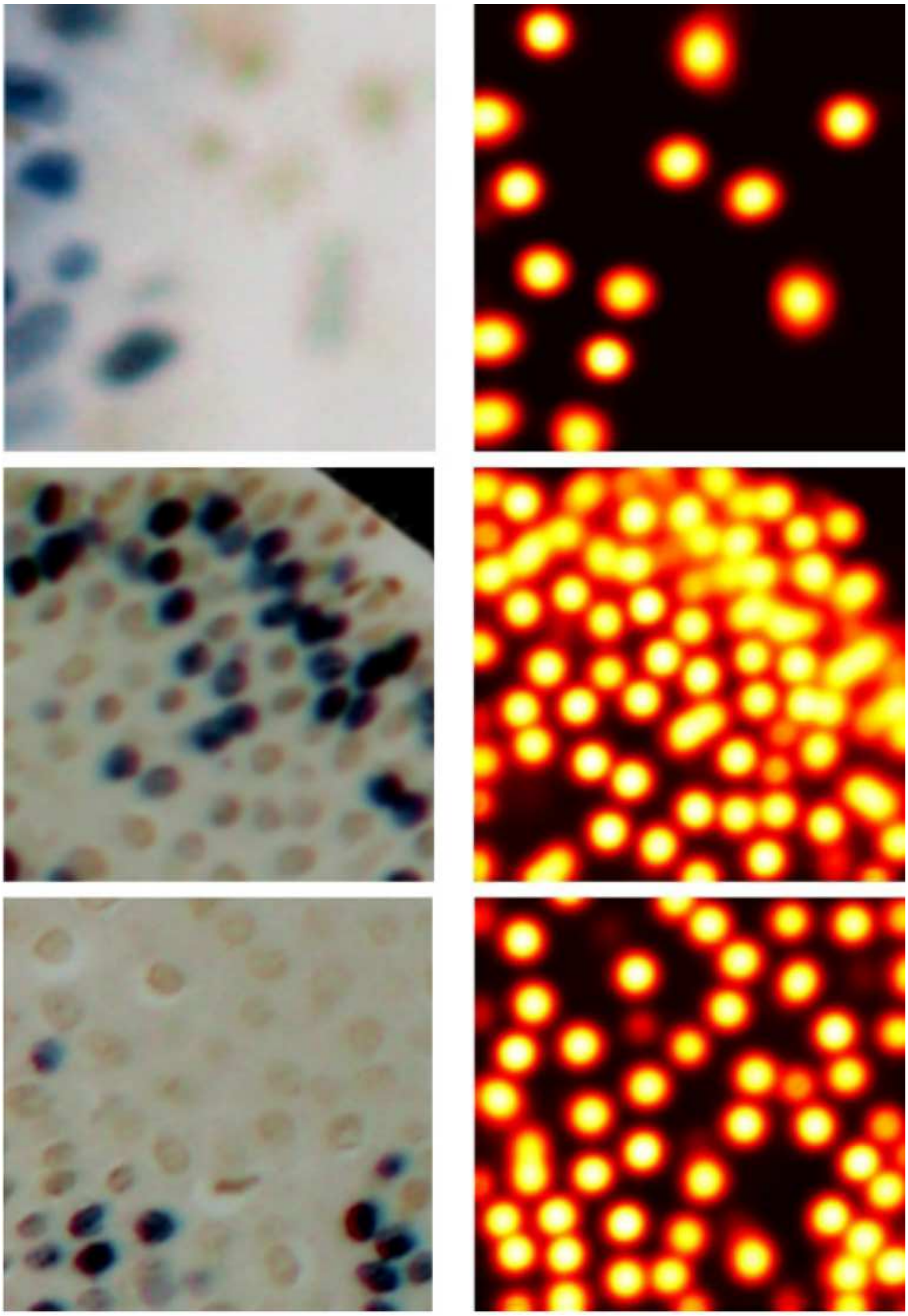


Figure 4.6: The first column has samples from the test set and the second column shows their corresponding predicted heatmaps.

images are divided into patches following the same strategy described in section 4.4.1 for U-Net model. In YOLO, the Ultralytics library handles image normalization internally, automatically scaling pixel values to the range [0,1] during both training and inference phases. Also, it uses internal augmentation techniques to enrich the training dataset.

Unlike U-Net, YOLO does not predict heatmaps, but bounding boxes. Since our dataset does not include bounding-box annotations, but only the coordinates of cell centroids, we generate synthetic bounding boxes by centering a fixed-size square around each centroid. Specifically, each bounding box extends 5 pixels equally in all directions from the centroid, resulting in a region of 11×11 pixels per cell. While this is smaller than the minimum cell diameter of 20 pixels, this approach is justified because:

- Our task focuses on centroid localization rather than full cell detection.
- YOLOv8 employs an anchor-free architecture with automatic anchor assignment, eliminating anchor box mismatch issues.
- YOLO's convolutional layers process contextual information through their receptive fields, allowing access to full cell information beyond bounding box boundaries.
- Smaller boxes reduce computational overhead and minimize overlap between adjacent cells.

The YOLOv8 model was initialized with pre-trained weights from the COCO dataset (transfer learning) and trained using its default composite loss, which combines classification, IoU-based bounding box regression, and objectness terms. During training, the same augmentations applied to U-Net were used, including flipping, color jittering, Gaussian blurring, and noise addition. Optimization was performed using the default optimizer, Stochastic Gradient Decent (SGD with momentum = 0.937), with a cosine annealing learning rate schedule, starting at 0.01 and decaying to 0.001 over 50 epochs.

YOLO typically makes classifications of the detected objects. However, in this work, we are concerned with the detection of cells without making a classification of the cell staining level. In this way, we make its results comparable to the remaining segmentation methods.

It worths mentioning here that no morphological filtering is done to any feature map in any of the used DL models, including the output feature map of U-Net.

CHAPTER 5

ORALIMMUNOANALYSER SOFTWARE

The content of the following paper is partially reproduced in this chapter. The publication details are:

- Zakaria A. Al-Tarawneh, Maite Pena-Cristóbal, Eva Cernadas, José Manuel Suarez-Peñaranda, Manuel Fernández-Delgado, Almoutaz Mbaidin, Mercedes Gallas-Torreira Mercedes and Pilar Gándara-Vila. **OralImmunoAnalyser: a software tool for immunohistochemical assessment of oral leukoplakia using image segmentation and classification models**. *Frontiers in Artificial Intelligence*, Volume 7, 2024. DOI: 10.3389/frai.2024.1324410. ISSN=2624-8212

whose authors are:

1. Zakaria A. Al-Tarawneh and Almoutaz Mbaidin from the Computer Science Department, Mutah University, Karak 61711, Jordan.
2. Maite Pena-Cristóbal, Mercedes Gallas-Torreira and Pilar Gándara-Vila from Oral Medicine, Oral Surgery and Implantology Unit, MedOralRes Group of University of Santiago, Santiago de Compostela, Spain.
3. Eva Cernadas and Manuel Fernández Delgado from Centro Singular de Investigación en Tecnoloxías Intelixentes da USC (CiTIUS), Universidade de Santiago de Compostela, Spain.
4. José Manuel Suarez-Peñaranda from Pathological Anatomy Service, University Hospital Complex of Santiago (CHUS), and Department of Forensic Sciences and Pathology, University of Santiago. Both in Santiago de Compostela, Spain.

This chapter describes and evaluates in a biomedical research lab a graphical software tool, called OralImmunoAnalyser (OIA), to quantify IHC images of oral leukoplakia. OralImmnuno-Analyser (OIA) is a desktop application that runs on a general purpose computer under the Linux and Windows operating systems. It has been written in the C++ programming language using the GTK+ library[1] to develop the GUI and the OpenCV library[2] to develop the automatic algorithms to process the images. OralImmunoAnalyser is available from CiTIUS repository[3] for research purposes. Figure 5.1 shows the GUI of OIA with a typical immunohistochemical image loaded, processed and reviewed by the expert, and with the lateral panel displayed. The main contributions of OralImmunoAnalyser software are as follows:

1. Its friendly graphical interface simplifies the expert daily work flow to count cells with precision.
2. The automatic image processing is fast enough to operate in real time.
3. Experts can review and modify the results of the automatic analysis before the image quantification step, overcoming other existing software packages.
4. Helps pathologists to annotate IHC images for further research work.
5. Allows us to train the classifier for another biomedical research lab.
6. Count the cells for each staining level in each epithelial strata of IHC image.
7. The software is accurate, trustworthy, and easy to install for non-expert users.

This chapter is organized in the following sections: section 5.1 describes the system architecture and software functionality, section 5.2 specifies the cell detection methods included in the tool, section 5.3 explains the cell classification process according to the staining level of cells, section 5.4 briefly mentioned the statistical measures used to quantify the performance of the cell detection and classification tasks, section 5.5 presents the results of the evaluation of OIA software from the performance and usability perspectives and, finally, section 5.6 summarizes the main conclusions reached.

5.1 System architecture and functionality

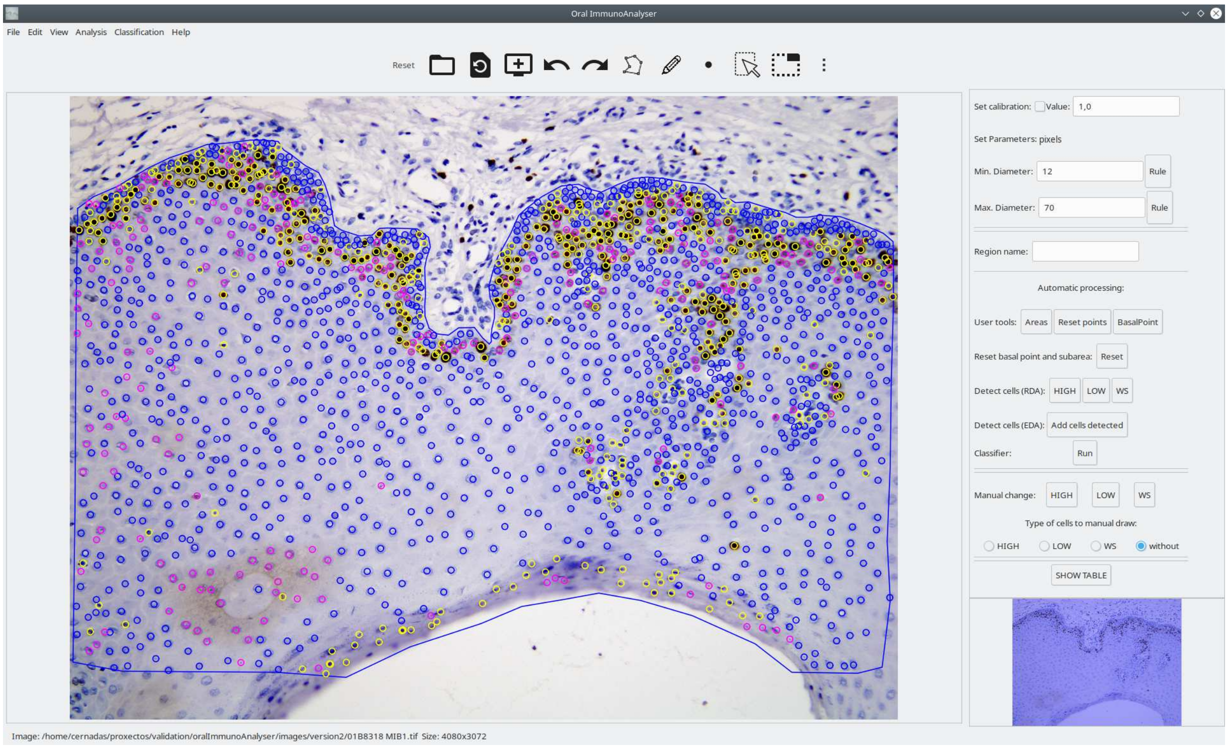
The architecture of OralImmunoAnalyser is modular and extensible, being composed by the classical three layers:

---

[1] https://www.gtk.org/

[2] https://opencv.org/

[3] https://citius.usc.es/transferencia/software/oralimmunoanalyser



Set calibration: ☐ Value: 1,0

Set Parameters: pixels

Min. Diameter: 12

Max. Diameter: 70

Region name:

Type of cells to manual draw:
○ HIGH ○ LOW ○ WS ● without

Image: /home/cernadas/proxectos/validation/oralImmunoAnalyser/images/version2/01B8318 MIB1.tif Size: 4080x3072

Figure 5.1: Screenshot of the software OralImmunoAnalyser. In the region of analysis (defined by the blue line), the color of the dots shows the category of the cells: yellow (highly stained), pink (low stained) and blue (no-stained).

1. The GUI layer interacts with the user with editing tools, including modules to draw and manage objects, set preferences or interact with the software.

2. The logic application layer contains modules to detect the cells, classify them, train the classifier and calculate the statistical results.

3. The persistent layer to store all the data needed and calculated by the software, including modules to save the overlays on the image and the statistical results. The overlays on the image, which contain the analysis supervised by the experts, are stored in the popular text format XML (Extensible Markup Language). The statistical results, calculated from the overlays, are stored in the known text format CSV (Comma-Separated Values), which is portable, and it can be imported from other spreadsheet software for further use.

Figure 5.2 shows a flowchart with the main functionality of OIA, which is accessible from its GUI. A typical working session for an expert should have the following actions:

1. Open or load a microscopic IHC image into OIA.

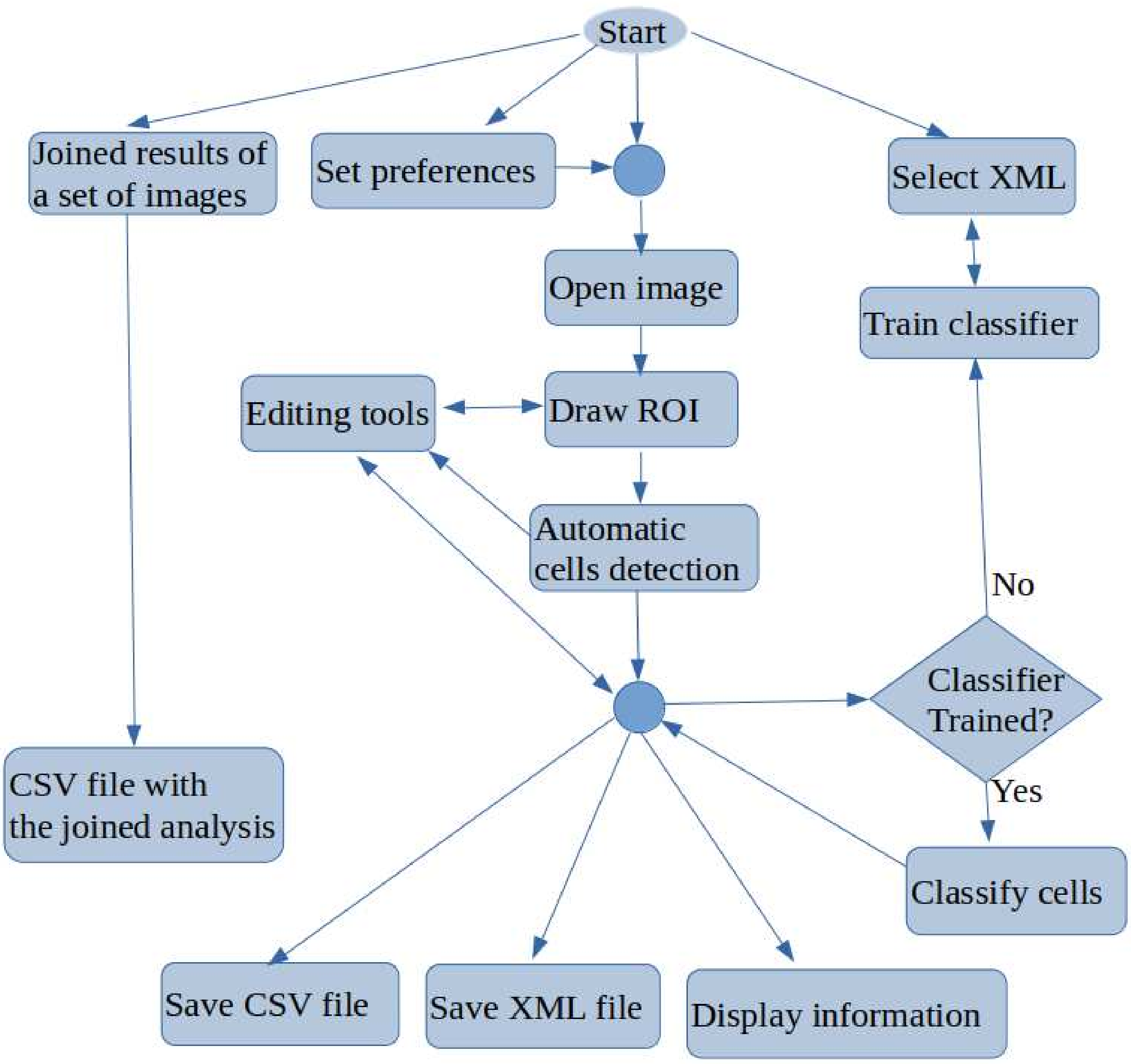


Figure 5.2: Flowchart containing the main tasks of OralImmunoAnalyser.

2. Once the image is loaded, a freehand region must manually be drawn using the OIA edition tools, in which the user wants to do the count.

3. After drawing and selecting that region of interest (ROI), the buttons of lateral panel can be used to automatically detect the cells in the ROI. The two algorithms included, RDA and EDA, are described in section 4.1.

4. Classifier can be ran in order to label each detected cell as “highly stained”, “low stained” and “no-stained”. To apply this process, the trained classifier is needed. Classifier can be

trained using the OIA functionality (see the details of classifier training in sections 4.2 and 5.3).

5. This automatic analysis may be not perfect due to the complexity of this type of images, so the expert supervises the detection and classification results using the editing tools in the following ways: i) deleting a set of detected cells; ii) changing the category of the detected cells; and iii) adding new cells specifying their category labels.

6. Save the overlays drawn on the image into a XML file. The overlays of the analysis must be saved into the XML file in order to review or export joined results in a future time.

7. Export the statistical measures and counts to CSV files. OralImmunoAnalyser allows to export joined results of a set of images going to the menu *Analysis → XML Files*, which opens a dialog screen to select the XML files and the output CSV file. To do this task, the images had to be analysed, supervised by the expert and saved the overlays in XML files (one per image).

8. Every time the user can do the following optional functionality: set preferences, set calibration and diameters, save joined results of a set of images, and train the classifier.

The working preferences of OIA can be set going to the menu *File → Set preferences*, in which you can set:

- The working directories for images, overlays and results.
- The width of points and lines.
- The color of the overlays for each cell category.
- The color of basal point.

OralImmunoAnalyser allows to set the spatial calibration, which is the relation between pixels in the image and real values (micrometers), or working in pixels units. If the calibration is set, the results are provided in real measures instead of pixels. The user must provide to OIA the minimum and maximum diameter of the cells to detect for an optimal operation of the automatic algorithms, which can be set writing in the *Preferences* dialog or graphically by drawing a straight line with the editing tools of the lateral panel. The preferences can be stored for future working sessions. OralImmunoAnalyser also allows to train the classifier going to the menu *Analysis → Train classifier*. More details of these functionalities can be read in the user guide, provided in the appendix A.

5.2 Cell detection algorithms

OIA includes two cell detection algorithms, called EDA and RDA, belonging to the edge-based and region-based segmentation paradigms respectively. Both algorithms ca be run from the lateral panel of OIA. The edge-based approach (EDA) is the Multi-Scale Canny Filter (MSCF) [48], parametrized with one scale and one pair of hysteresis thresholds (see section 4.1.1) and it is accessible from the lateral panel of the GUI in the label *Detect cells (EDA)*. The region-based approach (RDA) included in the software is described in section 4.1.2. It is accessible from the lateral panel of the GUI in the label *Detect cells (RDA)*, which encloses the three toggle buttons **HIGH**, **LOW** and **WS** to show the point sets $\mathcal{P}^{HIGH}$, $\mathcal{P}^{LOW}$ and $\mathcal{P}^{WS}$ respectively (see the algorithm 1 in section 4.1.2). The visualization of the detected cells is accumulated, applying an overlapping test to remove cells detected by different approaches. Two detected points are considered as two different cells if their distance is superior than the minimum diameter provided by user, i.e. the final set $\mathcal{P}$ of detected cells is $\mathcal{P} \leftarrow \{p_i \in \mathcal{P}^{TP}, p_j \in \mathcal{P}^{TP}, TP \in \{HIGH, LOW, WS, EDA\} | \texttt{distance}(p_i,\ p_j) > d_{min}\}$. A set of cells can be added or removed by clicking the previous toggle buttons. Finally, only the detected cells inside the region of analysis provided by the user are visualized in the software. The color of points means the cell staining level provided by the classifier (**Run** button after the label *Classifier*).

5.3 Cell classification

As it can be seen in the flowchart of figure 5.2, OIA includes modules to classify cells and to train the classifier. Some details of this process was described in section 4.2. Specifically, OIA uses the LibSVM implementation [16] of support vector machine (SVM), accessed through its C++ binding. Although the first version included a pre-trained SVM, the current version allows the SVM training. The module *Training panel*, through the submenu *Classification* → *Train classifier* allows the user to set the XML files (generated previously by OIA) that will be used to train the SVM classifier (see the details to train the classifier in the section A.11 of appendix A). These files must contain the cells recognized for a collection of images, alongside with their category.

Although some studies for the evaluation of ki67 in breast cancer [56] recommend labeling cells as positive and negative, OIA uses three categories: highly stained, low stained, and unstained cells to provide a more accurate quantification of the IHC oral leukemia image. The collection of images should be representative enough for the classification problem and must contain cells of all the categories. For the training, a maximum number of 1,000 cells is randomly selected from the XML files provided by the user, with similar numbers of cells for each category whenever possible (a minimum number of 10 cells is required for a category to be included in the training). OralImmunoAnalyser performs the tuning of the two hyper-parameters of the SVM (regularization λ and RBF kernel spread σ) using the grid-search method and the parameters value and experimental methodology mentioned in section 4.2.

5.4 Statistical analysis

To evaluate the cell detection algorithms, we define a true positive (TP) hit when a cell is correctly detected and a false positive (FP) whenever the user manually deletes the cell using the GUI. A cell is considered false negative (FN) if the user manually adds it. Once the TP, FP, and FN values are counted for an image, the sensitivity (Se), specificity (Sp) and average precision (AP), in %, are calculated as:

$$\begin{aligned} Se &= 100\frac{TP}{FN+TP} \\ Sp &= 100\frac{TP}{FP+TP} \\ AP &= 100\frac{TP}{TP+FP+FN} \end{aligned} \tag{5.1}$$

The performance of the SVM model in the prediction of the cell category $C \in \{$highly stained, low stained and without stain$\}$ is evaluated using the Cohen kappa, defined in eq. 4.2, and the accuracy (in %), whose value is 100 multiplied by the number of cells correctly classified by the classifier and divided by the total number of cells. The sensitivity and specificity of each category are also calculated considering that:

- The TP are the number of cells of category C_i correctly classified by the SVM into the category C_i.
- The FP are the number of cells classified into category C_i, but whose true category label is other.
- The FN are the number of cells of true category C_i that the classified assigned to other category.

5.5 Statistical results

This section presents and discuss the results achieved evaluating OralImmunoAnalyser softwate from differents perspectives: the ability to detect and classify cells in section 5.5.1, the cell counting process for each epithelial layer in section 5.5.2 and the evaluation of software tool from other perspectives in section 5.5.3. It is important to highlight that this software has been tested by the Odontology Faculty of the University of Santiago de Compostela since 2019 in its daily practice.

5.5.1 Detection and classification of cells

Two versions of OralImmunoAnalyser (OIA) was used in this research. The first version only includes the cell detection module with the EDA algorithm and it was used by experts to both quantify IHC images and annotated the staining level category for each detected cells. Regarding the classification module, it was trained by only two IHC images in the first version of OIA due to the lack of labeled cells. The second version of OIA also includes the RDA algorithm to detect cells. The first version of OIA was used to analyse 15 images distributed into 4 cases of leukoplakia without dysplasia, 3 of mild dysplasia, 2 of severe, 3 carcinomas in situ, 2 infiltrating carcinomas and 1 verrucous carcinoma. These images labeled by experts were used to train the current classification module of OIA.

Version	#images	#cells	Se	Sp	AP	Acc.	kappa
First	15	805.3	41.55	99.89	41.16	77.0	–
Current	26	1206.0	64.38	92.98	60.69	79.8	60.23

Table 5.1: Sensitivity (Se), specificity (Sp) and average precision (AP) in % of the two versions of OralImmunoAnalyser working in the lab to detect the cells. Average accuracy (Acc) and kappa for cell classification into three categories (highly stained, low stained and without stain).

Table 5.1 shows the results achieved for the two versions of OIA, obtained from the analysis of the XML files. The average number of cells in each image was approximately 1,000. The inclusion of the region-based approach (RDA algorithm) increases the average precision from the 41.2% of the first version up to 60.7% to the current version, and also increases sensitivity form 41.5% up to 64.4%. That means that, in average, the 60.7% of cells in the image are correctly detected and the remaining 39.3% of cells had to be supervised by expert (adding or deleting cells). Thus, the specificity is much higher than sensitivity, which means that the experts needed to add cells and deleted a small number of cells. The classification accuracy remains more or less constant for both versions, achieving the value 79.8% with the current version and a kappa value of 60.23%. In order to analyse the influence of the high variability among the immunohistochemical images on the system performance, figure 5.3 shows the boxplots of sensitivity, specificity and average precision for the detection of cells, and the accuracy of the cell classification. The boxes enclose the data between 25th and 75th percentiles, the red line is the median and the black wiskers extend the extreme data points. The sensitivity box is the largest one (about 20 points), so it is more affected by the image variability than the other measures, existing important differences among images in the proportion of cells that are not detected and must be added by the users. The specificity box is much smaller, so it varies less among images, and its median is much higher. The average precision box is also small, but its median is even lower than sensitivity. Finally, the accuracy box is also very small, so it is not very sensitive to the image variability.

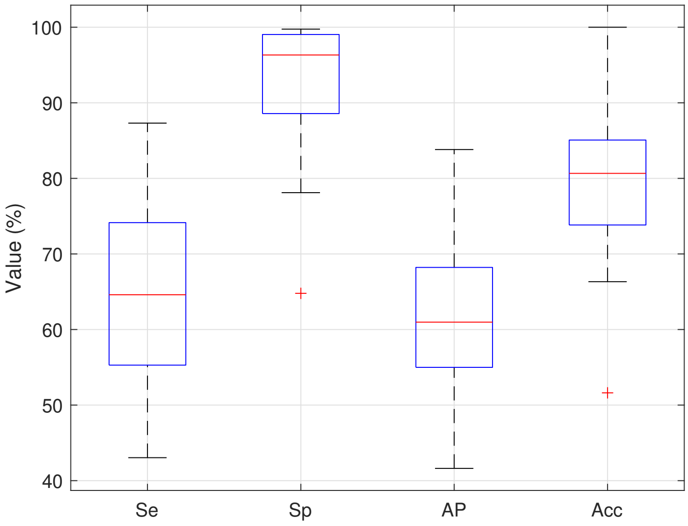


Figure 5.3: Boxplots of sensitivity (Se), specificity (Sp), average precision (AP) and classification accuracy (Acc) for the images analysed using the current version of OIA.

The current version was used to analyse 26 new IHC images distributed into: 24 did not present epithelial dysplasia, 1 case of mild dysplasia and 1 case moderate dysplasia. In relation to the cell classification, it classifies correctly the 79.8% of the cells with a kappa of 60.23%. More detailed view are in table 5.2, which shows the confusion matrix for the category prediction (the value in row i and column j is $100N_{ij}/N$, where N_{ij} and N are the same as in eq. 4.2). The diagonal numbers (in bold) give the percentage of cells correctly classified for each category, and the sum of the diagonal gives the classification accuracy. The best performance is provided for cells without staining achieving a high sensitivity (96.9%) and specificity (80.7%) and the worst results are for the low stained cells (sensitivity 42.6% and specificity 62.9%), because the system confuses cells with low and no staining. The reason is that the background staining, which is produced by a defect in the processing of the sample, is very similar to low-intensity brown, and the experts label the cells on these parts as "without stain". During training, the classifier learns to predict "without stain" for cells with this brown color, and therefore it wrongly classifies cells that are low stained as "without stain". For cells highly stained, the system has a high specificity (94%) with moderate sensitivity (67.1%), i.e. some highly stained cells are classified as low or without staining, due mainly to artefacts, but very few low stained or without stain cells are classified as "highly stained" (0.7% and 0.08%, respectively).

		Predicted category				
		Highly	Low	Without	Se	Sp
True	Highly	**12.14**	3.60	2.36	67.07	93.96
Category	Low	0.70	**9.15**	11.64	42.54	62.69
	Without	0.08	1.80	**58.52**	96.89	80.69

Table 5.2: Confusion matrix (in %) for cell category prediction (highly stained, low stained and without stain) for the current version of OralInmunoAnalyser. The columns Se (Sensitivity) and Sp (Specificity) provide the system sensitivity and specificity for each class respectively.

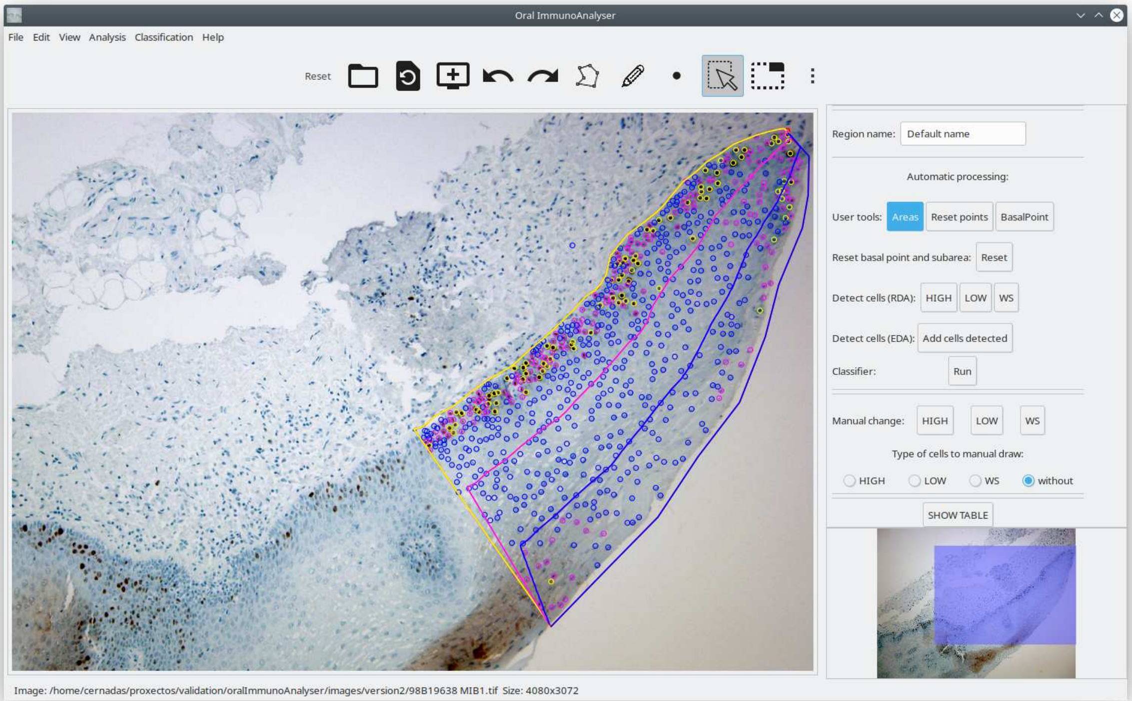
Following the recommendation of breast cancer studies [56], if, in our system, highly stained cells were considered positive and low stained and without stained cells were considered negative, the results in table 5.2 would achieve an accuracy of 93.6% and a kappa of 75.66%, maintaining the sensitivity and specificity for the positive cells and with a sensitivity and specificity of 98.92% and 93.66% respectively for the negative cells. The consideration of two levels of staining instead of three clearly improves the global performance of the cell classification.

5.5.2 Cell counting

In order to count the number of cells for each staining level and epithelial layer, the IHC image must be analyse, i.e. cells must be detected and classified within the ROI drawn by experts. Then, using the OIA editing tools, the expert must mark with the mouse a point on the image indicating to the software where is the basal part of the ROI in study. This step is necessary to do the count for the basal, medial and superior areas of the ROI, otherwise the count is only done globally. Given the basal point and the ROI under study, some geometrical computations can be done in order to divide the ROI in three areas: basal, medial and superior (see the algorithm described in section 4.3). The areas calculated can be visualized clicking the button *Areas* in the lateral panel, as it can be seen in figure 5.4, which shows overlapped on the IHC image an example of the resulting three epithelial layers calculated. Figure 5.5 shows other examples to illustrate the division of epithelium intro the three layers.

5.5.3 Analysis performance

OIA is compared with other procedures to quantify the IHC images of oral tissue. The comparison can be made from different points of view: the quality of statistical information provided by the analysis, the expert's analysis time and the expert's perception.



Region name: Default name

Type of cells to manual draw: ○ HIGH ○ LOW ○ WS ● without

Image: /home/cernadas/proxectos/validation/oralImmunoAnalyser/images/version2/98B19638 MIB1.tif Size: 4080x3072

Figure 5.4: Screenshot of OIA with an example visually illustrating how the cell counting algorithm works in the basal, medium and superior layers.

Once the region of interest is delineated on the image, OIA analyses the positivity for ki67 biomarker (the positive cells are those stained of brown color independently of its intensity) of that region, and exports to CSV files the following information:

– Percentage of positive cells.

– Percentage of cells for each staining intensity (highly stained, low stained and no stained cells).

– Distribution of the positivity in the basal, medium and superior layers of epithelium.

The clinical practice is implemented in OIA by analysing the nuclear expression of ki67 for the three epithelium layers, allowing to compare the cell proliferation among epithelium layers, similarly to the analysis of epithelial dysplasia.

The OIA software is reliable and precise because it allows to review the detection and classification of cells before counting. It is also easy-to-use and install. Another advantage that the program offers compared to manual quantification is the possibility of sharing the images and

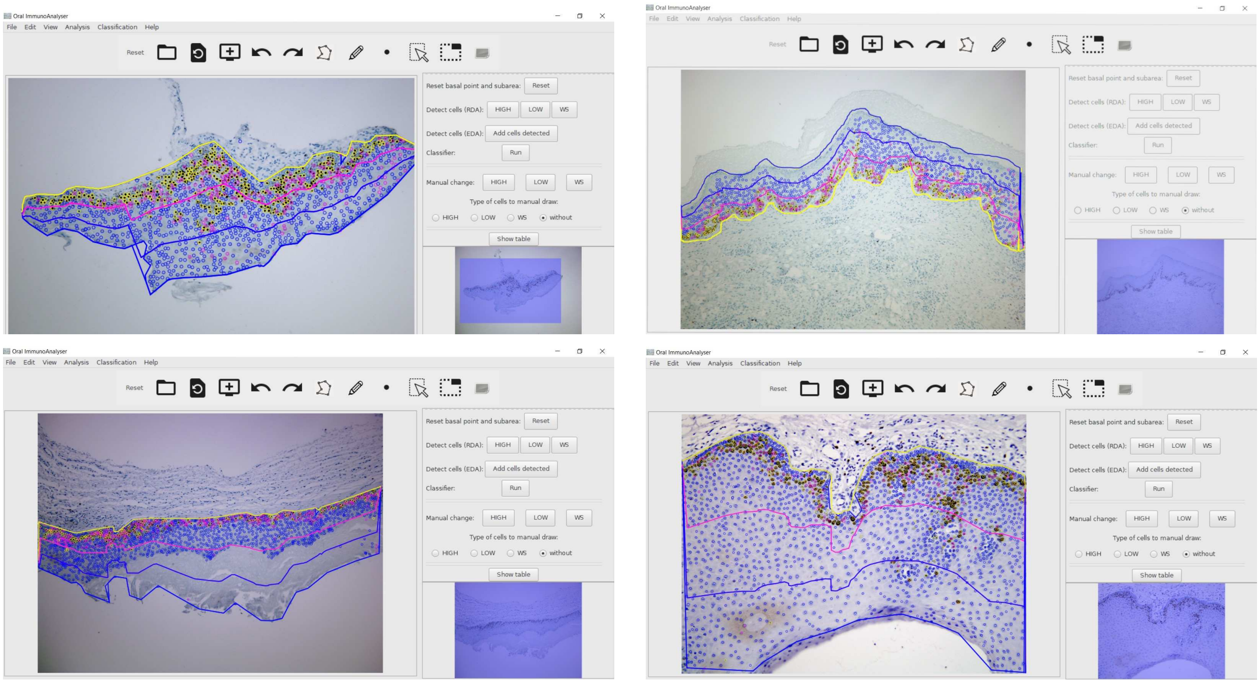


Figure 5.5: Other examples to illustrate the cell counting algorithm to divide the epithelium into the basal, medium and superior layers.

results among different experts to evaluate each case, reducing the variability among experts. To the best of our knowledge, there is no other application that counts automatically the cells of each staining level into the three epithelial stracta: basal, medial and superior (see the figures 5.4 and 5.5). This functionality of OIA allows a global analysis of the epithelium studied, which is of great importance for clinical diagnosis using IHC techniques, but it is currently not performed in the diary clinical practice due to its difficulty.

OralImmnunoAnalyser runs on a general purpose computer in a reasonable time. The delineation of the region of analysis and automatic processing takes less than 1 minute, similarly to the Aperio software[4]. The time required by the analysis is dominated by the expert's supervision, that depends on the review needs and the number of cells counted, being about 10 minutes for counting 1,000 cells (1 minute per 100 cells). This revision guarantees that the results are accurate and trustworthy. The traditional procedure to count manually 100 cells in the image takes approximately 2.3 minutes. So, OIA represents a saving of 1.3 minutes for 100 cells, being 2.3 times faster than manual counting and allowing to check the cells counted at any time, that reduces the chances of making human counting errors. Therefore, OIA reduces the analysis time by 56.5% with respect to the manual procedure.

---

[4] https://www.leicabiosystems.com/en-es/digital-pathology/manage/aperio-imagescope/

The expert's perception about OIA was evaluated using the system usability scale (SUS), a free questionnaire to measure the learning ability and subjectively perceived usability of computer systems [7, 10]. The questionnaire encloses the following 10 items with a five-point scale ranging from 1 (strongly disagree) to 5 (strongly agree):

1. I think that I would like to use this product frequently.
2. I found the product innecesarily complex.
3. I thought the product was easy to use.
4. I think that I would need the support of a technical person to be able to use this product.
5. I found the various functions in the product were well integrated.
6. I thought there was too much inconsistency in the product.
7. I imagine that most people would learn to use this product very quickly.
8. I found this product very awkward to use.
9. I felt very confident using the product.
10. I needed to learn a lot of things before I could get going with this product.

The score is calculated adding up the positively worded items (1, 3, 5, 7 and 9), subtracting one from the user responses, and the negatively worded items (2, 4, 6, 8 and 10), subtracting the user responses from five. Multiplying the SUS score by 2.5 re-scale the score from 0 to 100. An comprensive interpretation of SUS score [70] is: SUS < 25 is the worst imaginable system; from 25 to 39 is from the worst imaginable to poor; from 40 to 52 is from poor to OK; from 53 to 73 is OK to good; from 74 to 85 is good to excellent; and above 85 is excellent to the best imaginable system. A small sample between 8 to 12 users is enough to give a good assessment of how people see the software. The SUS questionnaire to evaluate OIA perception was filled out by eleven experts achieving a mean score of 80.9, which means that the system is from good to excellent from the experts point of view.

5.6 Summary and discussion

In conclusion, OralImmunoAnalyser (OIA) is a new reliable and easy-to-use software tool to estimate the oral leukoplakia from the IHC images of mouth tissues. This software combines the automatic detection and classification of cells in the image with a friendly GUI that allows the experts to review the cells recognized before the calculation of statistical results. OIA provides the number of positive (stained) and negative (without staining) cells in the region of analysis, the percentage of cells for each staining level (high, low and no staining) and the distribution of the cell positivity in different layers of epithelium (basal, medium and superior). From their daily use in the Odontology Faculty of the University of Santiago de Compostela, the automatic processing of the IHC images provided the following average performance: 1) the cell detection module achieved a sensitivity of 64.4%, specificity of 93% and precision of 60.7%; and 2) the cell classification in the three staining levels achieved an accuracy of 79.8%. The time required to analyse each image is dominated by the needs of supervision, being about 10 minutes to count 1,000 cells. So, OIA saves 56.5% of time spent by the traditional manual counting of cells, avoiding mistakes because the user can check at any time the cells counted. Despite OIA cannot operate fully automatically, it can accelerate considerably the analysis, that can be performed in the daily clinical practice, being a major advance over what is currently available. In fact, the expert's perception about OIA achieves a mean score of 80.9 in the SUS questionnaire, which means that the system is from good to excellent.

The OIA software is simple to use and install and has the following advantages:

- It works with a photograph taken under a microscope and not with a scan of the glass.
- It allows monitoring, i.e. to see the cells to be accounted for each category before image quantification.
- It allows to divide the epithelial strata into three thirds (basal / medium / superior), to compare the positivity of ki67 among layers and to study its relationship with the degree of epithelial dysplasia.

For these reasons, OIA is superior to other available tools and its use could be easily implemented in the daily practice of the biomedical labs. In addition, this possibility of supervision by the expert favors that OIA can be used as a tool in the teaching-learning process to instruct junior researchers in cell counting.

On other hand, the use of OIA in the biomedical lab allowed to annotate the OIADB dataset, which is publically available in the CiTIUS repository: https://gitlab.citius.gal/analyser/oiadb for their use in futher research.

Although OIA was validated with histological images of oral leukoplakia stained for ki67, our preliminary tests encourage its use with other molecular markers that also stain the cells with brown, such as p53 or p21. This possibility opens up new lines of research that we will address in the future: OIA will facillitate and optimize routinary immunohistochemical analysis and lead to a ever-increasing diagnostic accuracy.

CHAPTER 6

RESULTS

The content of the following paper is partially reproduced in this chapter. The publication details are:

Z.A. Al-Tarawneh, A.S. Tarawneh, A. Mbaidin, M. Fernández-Delgado, P. Gándara-Vila, A. Hassanat, E. Cernadas. **Cell Detection in Biomedical Immunohistochemical Images Using Unsupervised Segmentation and Deep Learning**. *Electronics*. Special issue *Applications of Computer Vision, 3rd Edition*, 2025, Vol. 14(18), p. 3705, ISSN 2079-9292. DOI: 10.3390/electronics14183705.

whose authors are:

1. Z.A. Al-Tarawneh, A.S. Tarawneh, A. Mbaidin and A. Hassanat from the Computer Science Department, Mutah University, Karak 61711, Jordan.

2. E. Cernadas and M. Fernández-Delgado from Centro Singular de Investigación en Tecnoloxías Intelixentes da USC (CiTIUS), Universidade de Santiago de Compostela, Spain.

3. P. Gándara-Vila from Oral Medicine, Oral Surgery and Implantology Unit, MedOralRes Group of University of Santiago, Santiago de Compostela, Spain.

The segmentation of IHC images to detect the cell's nuclei position is a challenging topic in automatic image analysis in pathology. The previous chapter (chapter 5) presented and evaluated the OralImmunoAnalyser (OIA) software, which was used to create the OIADB dataset with 41 IHC images of mouth tissue with annotations (see the section 3.2 for the details of the expert's annotations). OIA tool includes two image segmentation approaches to detect cell's nuclei in IHC images of oral tissues. Both approaches can be applied to the IHC images interactively through OIA graphical interface. Nevertheless, this use does not allow us to evaluate independently each approach in order to have a deeper understanding of the strengths and weaknesses of each technique. The two unsupervised segmentation algorithms included in OIA were called

EDA (an edge-based detection algorithm) and RDA (an region-based detection algorithm), both described in section 4.1 of chapter 4. In this chapter, they are called OIA-EDA and OIA-RDA in order to emphasize that they are included in the OIA software. This chapter aims to statistically compare the performance of these segmentation approaches with other state of the art segmentation techniques on this problem, including both unsupervised segmentation techniques (described in section 4.1) and supervised deep learning models (described in section 4.4).

The rest of chapter is organized as follows: section 6.1 describes the metrics used to compare the segmentation algorithms; section 6.2 addresses the experimental setup used in order to provide a fair comparison among unsupervised and supervised methods; sections 6.3 and 6.4 present respectively the results obtained using unsupervised and deep learning cell detection techniques; section 6.5 analyses the computationanl time spent by cell detection algorithms; and section 6.6 discusses the results presented and summarizes the main conclusions derived from this comparison of cell detection techniques on IHC images.

6.1 Metrics to measure cell detection

Statistical evaluation is crucial for assessing and comparing the detection methods used. The choice of metrics depends on both the segmentation output format and the nature of the ground-truth data. Segmentation techniques may produce boundary-based outputs (e.g., pixel-wise contours), region-based results (e.g., segmented objects), or keypoint-based representations (e.g., sparse feature points), each requiring different evaluation approaches. Ground-truth data further influences the selection of metrics, since well-posed problems with a single definitive solution allow for direct comparison, while poorly posed problems, where multiple valid interpretations exist, require more flexible evaluation strategies.

Regarding the cell detection on immunohistochemistry (IHC) images, ground-truth annotations are typically provided as scribbles, bounding boxes, or points marking the x and y coordinates of cells. Given the nature of this problem, we adopt an indirect evaluation framework tailored to cell detection by treating segmentation performance as a detection task and using the F_1-score for statistical assessment. Here, true positives (TP) represent correctly detected cells, false positives (FP) denote erroneously detected non-cells, and false negatives (FN) correspond to missed true cells. In this study, the evaluation metrics, recall (R), precision (P), and F_1-score (F_1) are defined as:

$$R = \frac{TP}{FN + TP} \qquad P = \frac{TP}{FP + TP} \qquad F_1 = \frac{2P \cdot R}{P + R} \tag{6.1}$$

To evaluate cell detection performance, we compare the positions of segmented nuclei against ground-truth cell positions. A segmented nucleus is counted as a true positive (hit) if its position lies within the minimum diameter of a true cell nucleus; otherwise, it is considered a false negative (miss). When multiple true cells satisfy this distance criterion for a single segmented nucleus, it is counted only once to avoid overestimation. Any segmented nucleus that does not meet this proximity requirement with any true cell is classified as a false positive. This rigorous matching approach ensures accurate quantification of detection performance in our analysis.

In order to study the performance of segmentation algorithms in relation to the level of nuclei staining in IHC images, we calculate the accuracy of hits for each type of nuclei according to their staining level for an IHC image as:

$$A_X = 100\frac{TP_X}{N_X} \tag{6.2}$$

where A_X is the accuracy (%) for detecting nuclei with staining level X, TP_X is the number of correctly detected nuclei with staining level X, and N_X is the total number of nuclei with staining level X in the image. In this research, we consider three staining levels: highly stained nuclei, low stained nuclei and nuclei without staining.

6.2 Experimental setup

The experiments are designed to do a fair comparison between unsupervised segmentation methods (described in section 4.1) and supervised deep learning cell detection approaches (described in section 4.4) to cell detection problem in IHC images. In the unsupervised segmentation methods, the algorithm is applied to the original IHC images, which return the position of the detected cells. As mentioned, our objective is only to estimate the position of detected cells. Then, when the output of the unsupervised segmentation methods is a labeled or binary image (case of FH, ChV and clustering methods), we applied the following post-processing step. Let a_{min} and a_{max} be the minimum and maximum area of a true cell calculated as the area of a circle with the minimum and maximun diameters of the cells established by pathologists (for the OIADB dataset, the minimum diameter for nuclei is $d_{min} = 20$ pixels). The regions in the segmented image whose area fall within the interval $[a_{min}, a_{max}]$ are considered candidates to represent a true cells, whose positions will be the mass center of those regions. In this case, all images in the OIADB dataset are processed and the metrics described in section 6.1 are collected.

In the case of supervised deep learning (DL) approaches, we employ a leave-one IHC image-out (LOO) cross-validation approach, which is particularly suitable given the limited size of the OIADB dataset. This methodology ensures a fair evaluation by iteratively selecting all patches of each IHC image in the dataset as the test sample, while the patches of the remaining images are used to train the model. This methodology ensures that all images in the dataset are used to test the method, as in the case of unsupervised algorithms. According to this methodology, it is not possible that patches of the same image are present both in the training and test sets. Therefore, there is no data leakage nor inflated performance estimates (i.e. optimistic bias). Since each image corresponds to a different patient, i.e. each patient has only one image, it is not possible that information of the same patient are used both in the training and test datasets.

Both DL models used, U-Net and YOLO, are trained for 50 epochs, with early stopping (patience = 5), to strike an optimal balance between underfitting and overfitting, allowing the model sufficient time to learn the underlying patterns without an excessive risk of overfitting. The training batch size is set to 16, balancing computational efficiency with graphical processing unit (GPU) memory constraints to enable smoother training. The models are trained using a learning rate schedule, which starts at 0.01 and follows a cosine annealing pattern that decays to 0.001 during the course of training. This scheduling strategy automatically adapts well to our patch-based training approach. Stochastic gradient descent is used as an optimizer, with a momentum of 0.937 and weight decay set to 0.0005.

Finally, for all the methods, an overlapping test was applied in order to avoid two detections of the same true cell. If the distance between the position of two cells is smaller than the minimum diameter of a cell, we assume that these two positions correspond to the same true cell.

Table 6.1: Precision, recall and F_1-score and their standard deviation (SD) to detect cell's nuclei in IHC images of OIADB dataset for the unsupervised segmentation methods.

Method	**Config.**	**Recall**		**Precision**		**F_1-score**	
		%	**SD**	**%**	**SD**	**%**	**SD**
OIA-RDA	HIGH	8.4	5.3	87.6	26.5	15.0	8.8
	LOW	16.7	12.6	90.3	13.4	26.3	16.2
	WS	24.4	14.1	87.6	10.4	35.9	15.2
OIA-EDA		44.1	11.2	46.7	18.9	41.0	8.2
FH		34.6	11.4	40.9	13.1	34.5	3.4
ChV	MorphACWE	12.6	7.6	89.4	9.1	21.4	10.8
	MorphGAC	13	8.2	88.3	9.4	21.6	11.5
Clustering	kmeans-RGB	12.2	16.5	40.8	39.5	17	20
	kmeans-LAB	12.1	16.5	40.1	39.5	16.7	19.6
	SLIC	0.5	0.6	9.7	12.3	1	1.2
OIA	RDA-EDA	49.2	18	50.9	11.7	46.4	7.2

6.3 Cell detection using unsupervised segmentation

Table 6.1 shows the average precision, recall and F_1-score, with their standard deviation, for all the unsupervised segmentation techniques used in this study. OIA-RDA provides a high precision, about 87.6%. Recall and F_1-score are increasing from option HIGH (R=8.4% and F_1=15%) to option WS (R=24.4% and F_1=35.9%), but recall is low for all configurations, although the performance depends on the specific image. A behavior similar to OIA-RDA, although with lower global performance, is observed by active contours approaches (ChV) with an average F_1=21% (P=89.4%, R=12.6%). On the other hand, OIA-EDA shows a performance more balanced (i.e. R and P more similar between them) than OIA-RDA, achieving R=44.1% and P=46.7% with F_1=41.0%, that is the highest performance over all unsupervised methods. The FH graph cut method shows similar behavior than OIA-EDA, but with lower overall performance, F_1 =34.5% (P=40.9% and R=34.6%). Clustering-based approaches generally underperform, with K-means implementations in both RGB and LAB color spaces yielding recall and precision values around 12-13% and 40%, respectively, and correspondingly low F_1 scores (16.7-17%). The SLIC clustering method performs particularly poorly, with near-negligible recall (0.3%) and precision (4.2%). Quantitative results of MeanShift algorithm are not provided due to its very poor performance.

Table 6.2: F_1-score to detect cell's nuclei in IHC images of OIADB dataset for all unsupervised segmentation methods. The last three columns (A_{High}, A_{Low} and $A_{Without}$) show the accuracy (%) detecting high, low and without staining level of nuclei.

Method	**Config.**	F_1**-score**	A_{High}	A_{Low}	$A_{Without}$
OIA-RDA	HIGH	15.0	49.7	3.4	0.6
	LOW	26.3	59.0	17.7	7.6
	WS	35.9	63.9	24.0	15.8
OIA-EDA		41.0	46.3	36.7	45.3
FH		34.5	39.8	27.9	34.6
ChV	MorphACWE	21.4	58.2	17.2	2.5
	MorphGAC	21.6	57.7	17.7	3
Clustering	kmeans-RGB	17	22.4	10.9	10.2
	kmeans-LAB	16.7	22.8	10.3	10.3
	SLIC	1	0.2	0.2	0.3
OIA	RDA-EDA	46.4	69.8	42.4	48.7

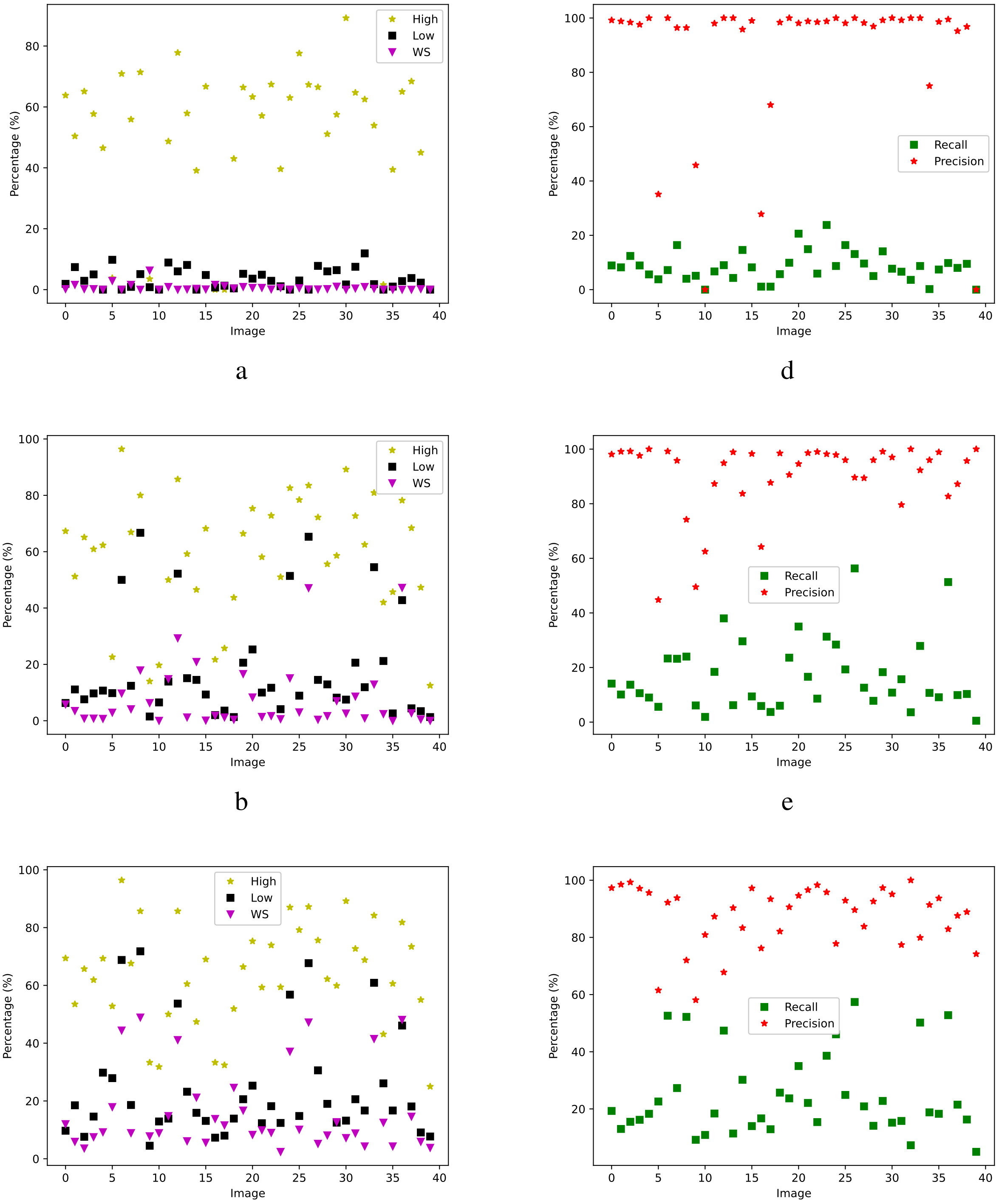
Table 6.2 shows the percentage of cells detected for each staining level and allows analysis of the behavior of the cell detection algorithms according to their staining level. In general, the region-based segmentation methods (see OIA-RDA and ChV methods in table 6.2) achieve good performance for the detection of highly stained cells, but its reliability is poor detecting the low stained or nonstained cells. In contrast, edge-based segmentation techniques exhibits a uniform behavior with respect to the level of cell staining (see OIA-EDA and FH methods in table 6.2), but their main drawback is the high number of false positives due to its moderate precision. Specifically, OIA-RDA detected, in average, 63.9% of the highly stained cells, where it detects less than 24% and 15.8% of low stained and without staining cells, respectively.

Figures 6.1, 6.2 and 6.3 show the behavior (accuracy, precision and recall) of all unsupervised segmentation algorithms for each image in the OIADB dataset. Figure 6.1 shows the performance for OIA-RDA approach for each IHC image in OIADB dataset. The rate of correct cell detection for each staining level of nuclei is shown in the graphs of the first column for the three options. The right-side graphs show the precision and recall for the three options. Normally, precision is high (red asterisks) for all images and recall (green squared) is rather low, almost for the HIGH option. This algorithm detect quite well the highly stained cells (yellow asterisks), but its performance is rather bad for cells without staining (purple diamonds).

Figure 6.2 shows the performance of OIA-EDA, MorphACWE and FH methods to detect cells in IHC images of OIADB dataset. Left-side graphs show the rate of correct cell detection for each staining level: highly stained cells as yellow asterisks, low stained cells as black squared and cell without staining as purple diamonds. The behavior of MorphACWE algorithm (panel b) is similar to OIA-RDA (see figure 6.1), showing much higher rates to detect highly stained cells than the remaining types of cells. The performance of OIA-EDA algorithm seem to be independent of the cell staining level for the majority of images (see panel a of figure 6.2). FH algorithm (panel c) detects better highly stained cells in some cases and its behavior is uniform with the staining level of cells in other cases. In relation to precision and recall (the corresponding right-side graphs in the figure 6.2), the precision and recall values are moderate for both OIA-EDA and FH methods, while MorphACWE method achieves high precision and low recall for almost all images, similary to OIA-RDA method.

Figure 6.3 shows the performance of clustering methods to detect cells in IHC images of OIADB dataset. Specifically, the kmeansRGB approach (first row of the figure), kmeansLAB approach (second row) and SLIC algorithm (third row). The behavior of all approaches are similar to OIA-RDA method, but these algorithm did not converge to a solution for some images. So, their global performance would be lower than the remaining approaches.

Figures 6.4 , 6.5, 6.6 and 6.7 illustrate the visual performance of unsupervised segmentation algorithms for images 4, 12, 17 and 34 of OIADB dataset respectively. The upper left image of



a

d

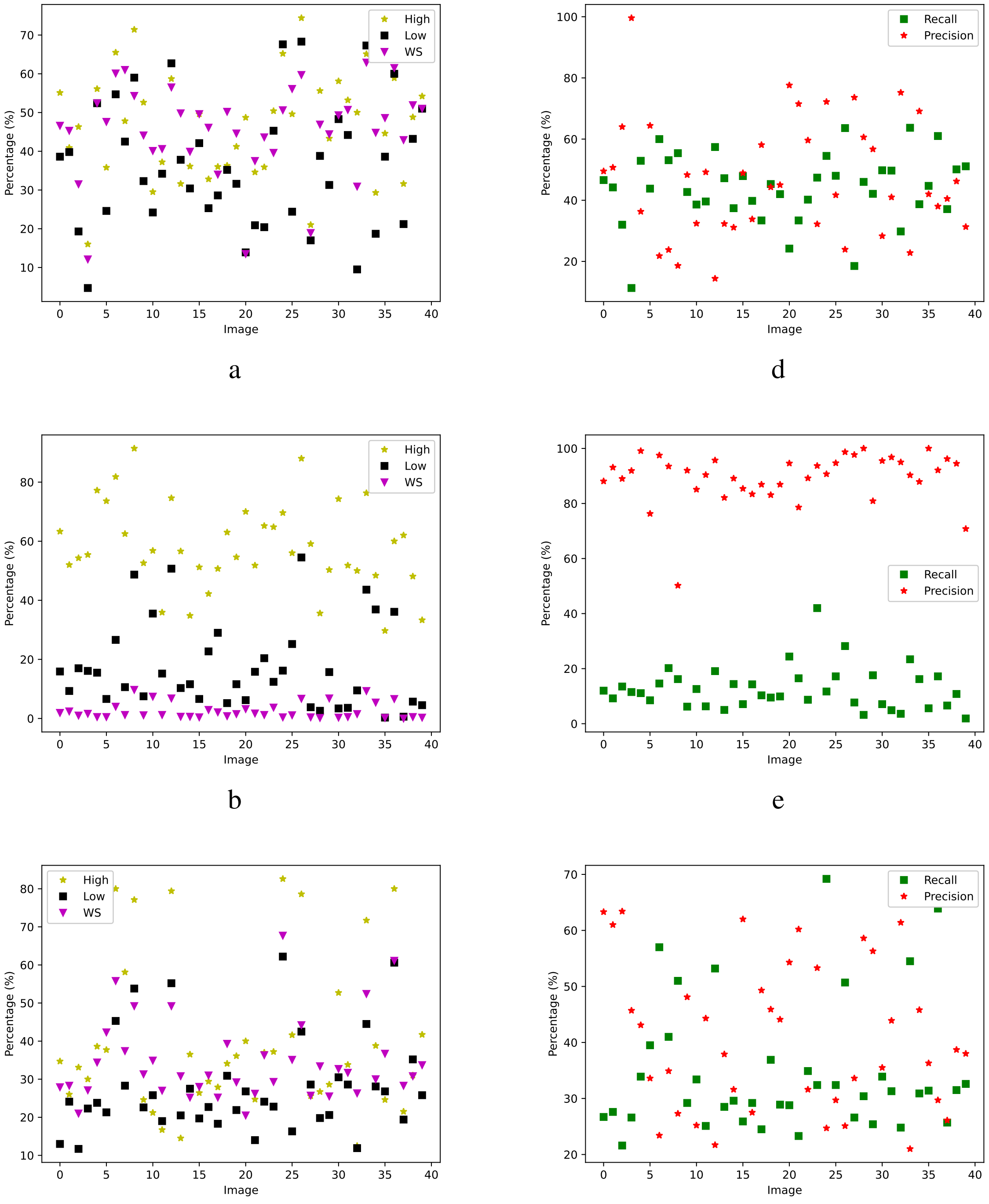
b

e

c

f

Figure 6.1: The left column shows the accuracy (in %) for detecting nuclei belonging to each staining level and the right column shows the precision and recall (in %) for each image in OIADB datasets using OIA-RDA algorithm with option **HIGH** (a) and (d); **LOW** (b) and (e); and **WS** (c) and (f).



a d

b e

cc

c f

Figure 6.2: The left column shows the accuracy (in %) for detecting nuclei belonging to each staining level and the right column shows the precision and recall (in %) for each image in OIADB datasets: OIA-EDA (a) and (d); Morphological Active Contours without Edges (MorphACWE) (b) and (e); and the graph cut method FH (c) and (f).

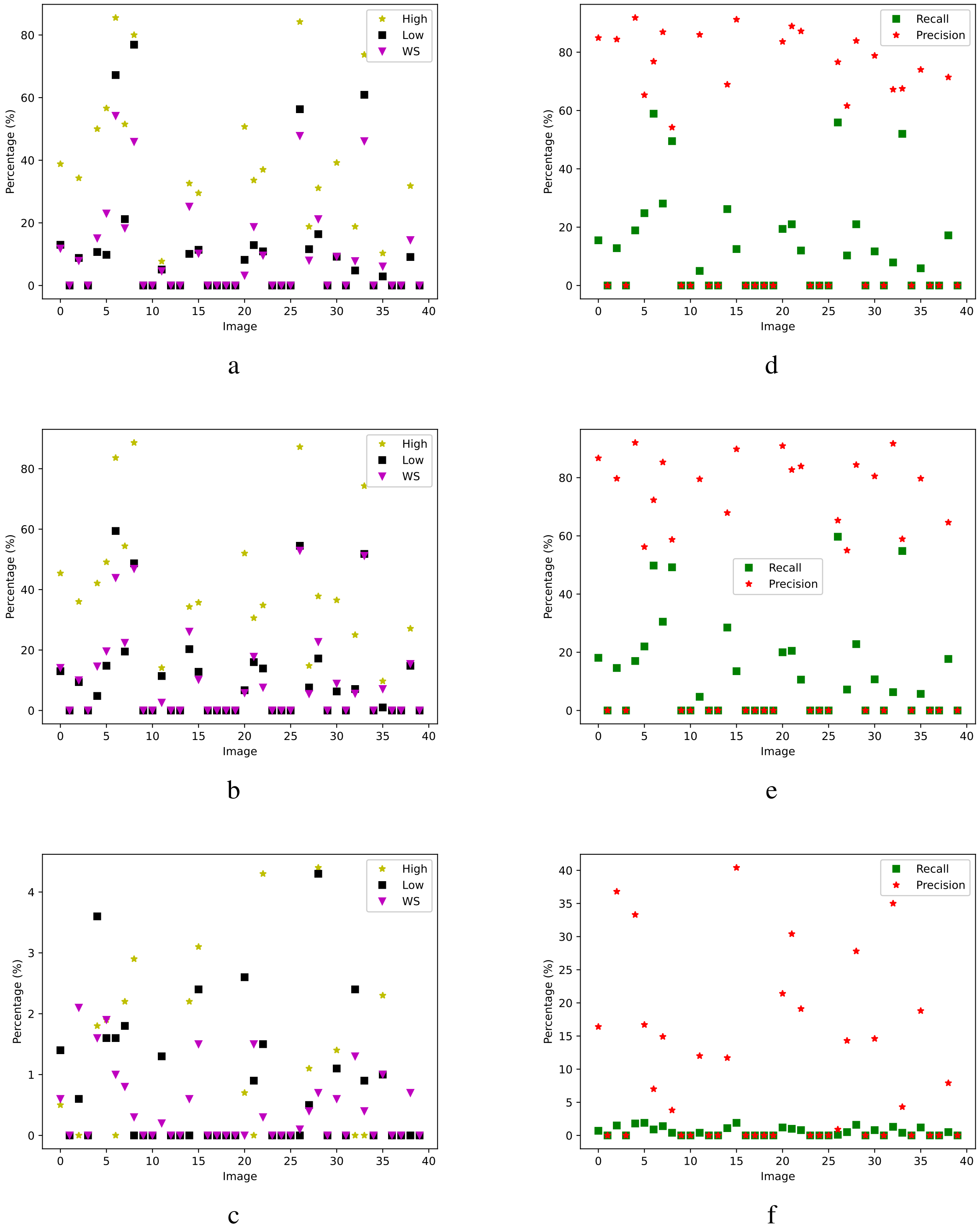
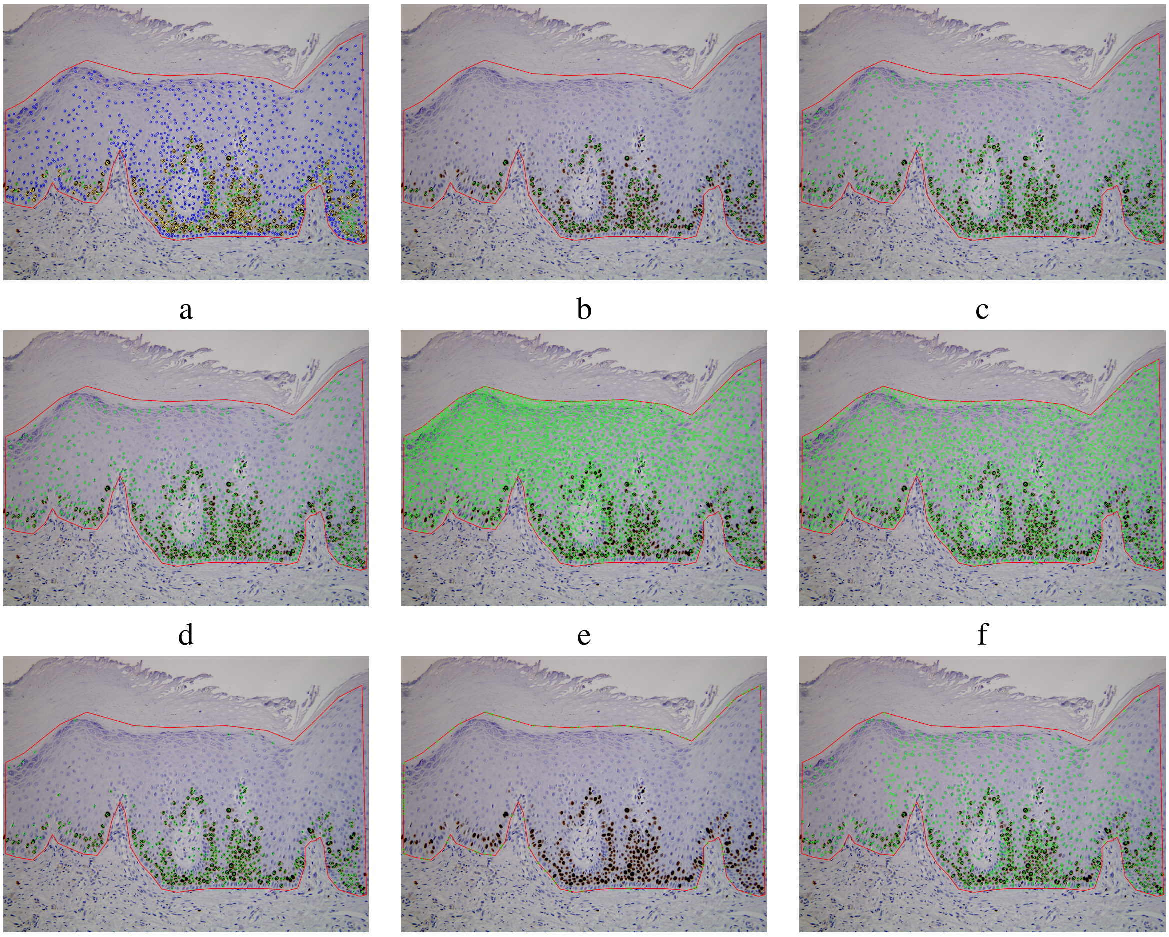


Figure 6.3: The left column shows the accuracy (in %) for detecting nuclei belonging to each staining level and the right column shows the precision and recall (in %) for each image in OIADB datasets: kmeansRGB (a) and (d); kmeansLAB (b) and (e); and SLIC (c) and (f).


a b c

d e f

g h i

Figure 6.4: Example (`image4`) of unsupervised cell detection approaches (cells are overlapped to the IHC image): (b) OIA-RDA (option HIGH); (c) OIA-RDA (option LOW); (d) OIA-RDA (option WS); (e) OIA-EDA; (f) FH; (g) ChV (option MorphACWE); (h) SLIC ; and (i) kmeans-RGB. The groundtruth is shown in (a), where the color means the staining level of nuclei: yellow (highly stained), green (low stained) and blue (without staining).

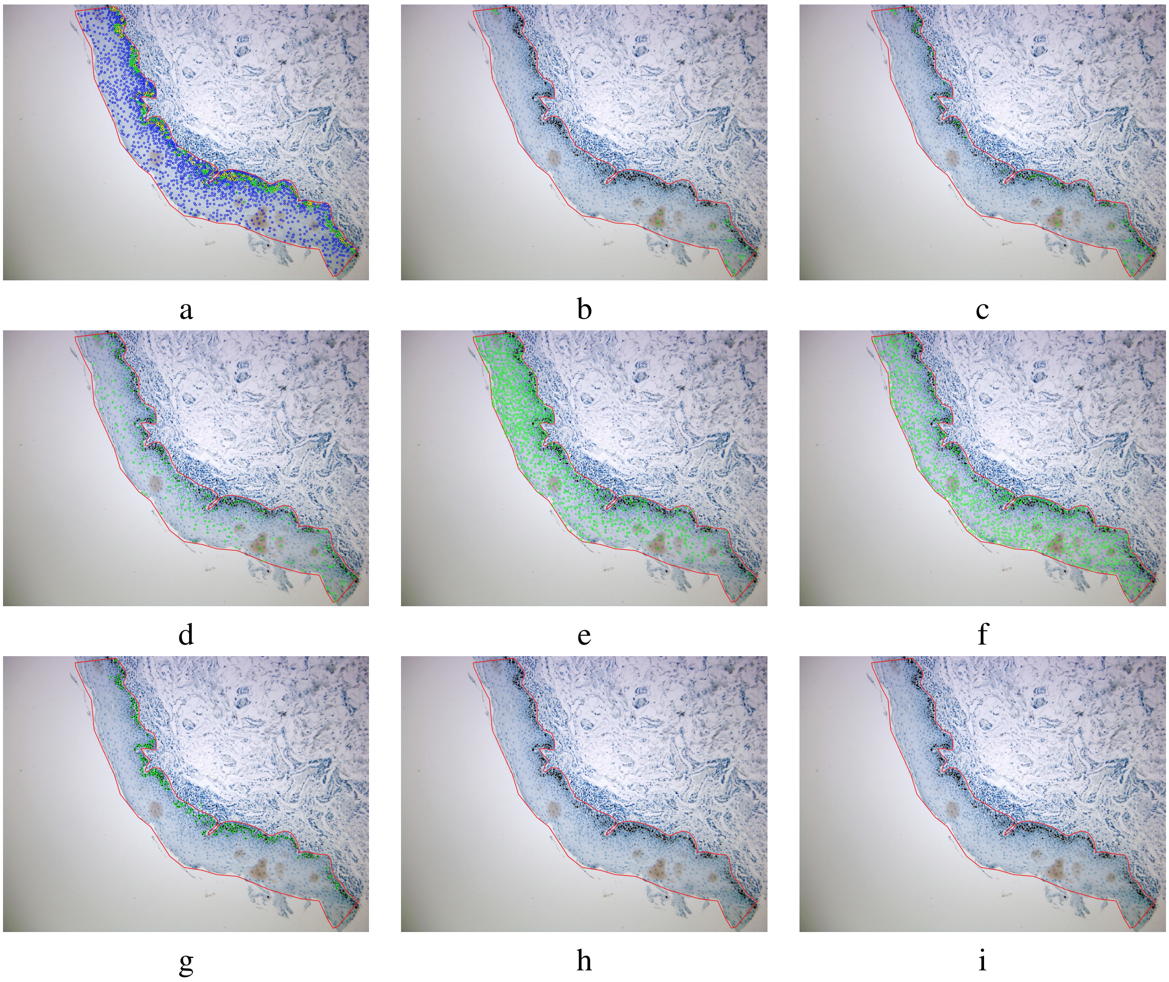
each figure is the original IHC image, where the ground truth cells annotated by experts are overlapped. The color means the staining level of cells nuclei: yellow for the highly stained nuclei, green for low stained ones and blue for the nuclei without staining. As mention in chapter 3, the composition of IHC images according to the staining levels of their cell's nuclei varies among them (see the details for these images in table 3.1). The behavior of cell detection algorithms is statisically different, and depends on percentage of cells for each staining level, as it can be see in the table 6.3, 6.4, 6.5 and 6.6 for images 4, 12, 17 and 34 of OIADB dataset respectively. For example, the image of figure 6.4 (`image4`) contains 1378 cells with 19.3% highly stained, 12.1% low stained and 68.6% without staining (see table 3.1). Table 6.3 shows the precision, recall, F_1 and percentage of correct detected cells for each staining level for the `image4` of figure 6.4. The cells detected by OIA-RDA approaches (see panels b, c and d of figure 6.4) are nearly all true (high precision) but some cells are missed, mainly for HIGH option, decreasing recall. Therefore, the F_1-score, that represents a weighted average of precision and recall, is still high ($F_1 = 70\%$) for WS option. On other hand, OIE-EDA and FH approaches achieved similar performance (see panels e and f of figure 6.4), lower values of precision, which are close to 25% and higher values of recall, between 50% to 65%, which means that both methods detect many of true cells, but providing many false positive (low precision), achieving lower global performance (F_1 =33-35%) than OIA-RDA (see panels g and h of figure 6.4). MorphACWE and kmeans-RGB approaches (see panels g and i of figure 6.4) have a similar behaviour than OIA-RDA methods (high precision and moderate recall) but with lower global performance, achieving F_1=43.8% and 64.6% respectively. SLIC works very bad (see panel h in figure 6.4).

Table 6.3: Precision, recall and F_1-score (columns three, four and five) to detect nuclei cells in IHC image `image4.tif`. The last three columns (SH, SL and SW) show the sensitivity for the high, low and without staining level of nuclei.

Method	**Config.**	**Recall**	**Precision**	F_1**-score**	A_{High}	A_{Low}	$A_{Without}$
OIA-RDA	HIGH	13.1	100	23.1	67.3	0	0.1
	LOW	56.3	89.6	69.2	83.3	65.3	47.1
	WS	57.4	89.6	70	87.2	67.7	47.2
OIA-EDA		63.6	23.9	34.8	74.4	68.3	59.7
FH		50.7	25.1	33.5	78.6	42.5	44.2
ChV	MorphACWE	28.2	98.7	43.8	88	54.5	6.7
Clustering	kmeans-RGB	55.9	76.6	64.6	84.2	56.3	47.8
	SLIC						

A_{High}: accuracy(%) for high stained nuclei, A_{Low}: sensitivity for low stained nuclei and $A_{Without}$: sensitivity for nuclei without staining.

Figure 6.5 and table 6.4 show the visual and quantitative results of the unsupervised segmentation methods on IHC image `image12.tif`, where the percentage of highly stained nuclei is lower and the percentage of cells without staining is higher than `image4.tif` (see table 3.1). In table 6.4, similar tend as with `image4` is observed (compare to 6.3), but with lower performance (the highest F_1 is only 42.4% using OIA-EDA algorithm). OIA-RDA approaches achieved high precision, from 68% to 93%, but with very low recall (only 13%), achieving only a F_1 =22.6%. This may be due to the fact that this IHC image is of poorer quality and that it presents a lower percentage of highly stained cells and higher percentage of cells without staining. In this case, OIA-EDA and FH algorithms present a similar behavior than with the previous



a b c d e f g h i

Figure 6.5: Example (`image12`) of unsupervised cell detection approaches (cells are overlapped to the IHC image): (b) OIA-RDA (option HIGH); (c) OIA-RDA (option LOW); (d) OIA-RDA (option WS); (e) OIA-EDA; (f) FH; (g) ChV (option MorphACWE); (h) SLIC ; and (i) kmeans-RGB. The groundtruth is shown in (a), where the color means the staining level of nuclei: yellow (highly stained), green (low stained) and blue (without staining).

image or even better, achieving higher global performance (F_1 =33-42%) than with `image4`. Kmeans-RGB did not converge to a solution and SLIC algorithm is also very bad.

Figure 6.6 and table 6.5 show the visual and quantitative results of the unsupervised segmentation methods on IHC image `image17.tif` (see table 3.1), which has a very high percentage of highly stained cells (59.1%). As shown in table 6.5, the highest F_1-score is provided by MorphACWE algorithm (F_1 =58%). OIA-RDA approaches achieved very high precision, higher than 95.8% for all configurations, with moderate recall (38.6%), archieving a F_1 =55%. OIA-EDA and FH algorithms present a moderate global performance (F_1 =38.3% and 40.3% respectively) resulting from a moderate precision and recall. Kmeans-RGB did not converge to a solution and SLIC algorithm is also very bad.

Table 6.4: Precision, recall and F_1-score (columns three, four and five) to detect nuclei cells in IHC image `image12.tif`. The last three columns (SH, SL and SW) show the sensitivity for the high, low and without staining level of nuclei.

Method	**Config.**	**Recall**	**Precision**	**F_1-score**	A_{High}	A_{Low}	$A_{Without}$
OIA-RDA	HIGH	1.1	68	2.2	0	1.3	1.2
	LOW	3.7	87	7.1	25.7	3.6	1.2
	WS	12.9	93.4	22.6	32.4	8	11.6
OIA-EDA		33.4	58.1	42.4	36	28.6	34
FH		24.5	49.3	32.7	27.9	18.3	25.2
ChV	MorphACWE	10.3	86.9	18.4	50.7	29	2.1
Clustering	kmeans-RGB	0	0	0	0	0	0
	SLIC						

A_{High}: accuracy(%) for high stained nuclei, A_{Low}: sensitivity for low stained nuclei and $A_{Without}$: sensitivity for nuclei without staining.

Table 6.5: Precision, recall and F_1-score (columns three, four and five) to detect nuclei cells in IHC image `image17.tif`. The last three columns (SH, SL and SW) show the sensitivity for the high, low and without staining level of nuclei.

Method	**Config.**	**Recall**	**Precision**	**F_1-score**	A_{High}	A_{Low}	$A_{Without}$
OIA-RDA	HIGH	8.2	99	15.1	66.7	4.8	0.1
	LOW	31.3	98.2	47.5	51	4.1	0.6
	WS	38.6	95.8	55	59.4	12.4	2.4
OIA-EDA		47.4	32.2	38.3	50.4	45.3	39.6
FH		32.4	53.3	40.3	37.2	22.8	29.3
ChV	MorphACWE	42	93.7	58	64.8	12.4	3.7
Clustering	kmeans-RGB	0	0	0	0	0	0
	SLIC						

A_{High}: accuracy(%) for high stained nuclei, A_{Low}: sensitivity for low stained nuclei and $A_{Without}$: sensitivity for nuclei without staining.

Finally, figure 6.7 and table 6.6 show the visual and quantitative results of the unsupervised segmentation methods on IHC image `image34.tif`. As shown in table 6.6, the highest F_1-score is provided by OIA-RDA algorithm with WS option (F_1 =64.5%). OIA-EDA and FH algorithms present a moderate global performance (F_1 =46.8% and 40.6% respectively) resulting from a moderate precision and recall. In this case, MorphACWE algorithm provides low

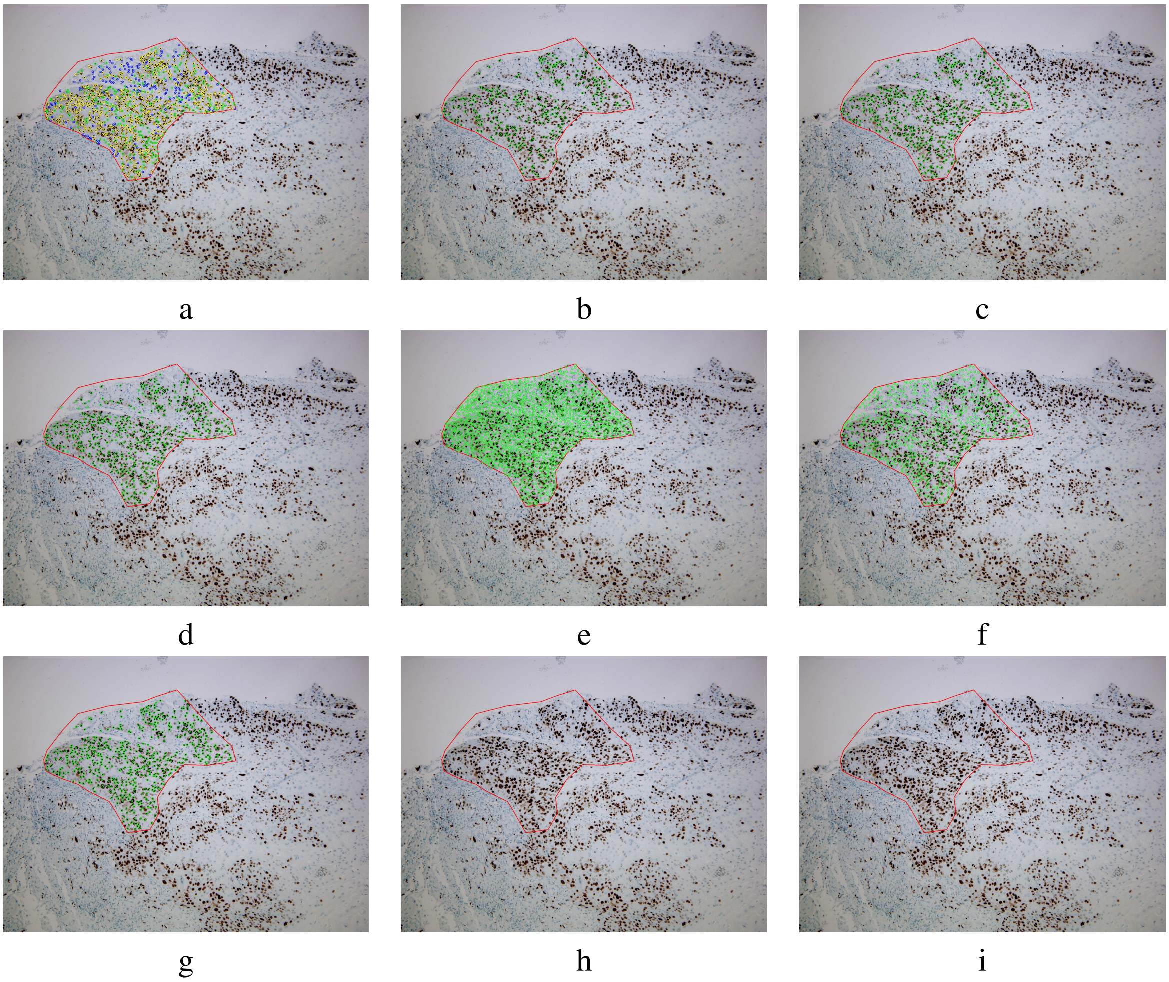


Figure 6.6: Example (`image17`) of unsupervised cell detection approaches (cells are overlapped to the IHC image): (b) OIA-RDA (option HIGH); (c) OIA-RDA (option LOW); (d) OIA-RDA (option WS); (e) OIA-EDA; (f) FH; (g) ChV (option MorphACWE); (h) SLIC ; and (i) kmeans-RGB. The groundtruth is shown in (a), where the color means the staining level of nuclei: yellow (highly stained), green (low stained) and blue (without staining).

performance , only F_1 =29%. Kmeans-RGB did not converge to a solution and SLIC algorithm is also very bad.

From the common behavior of OIA-RDA and OIA-EDA, we can assert that OIA-RDA method provides better results when the IHC images contains higher rates of stained cells (as for example the `image17` in figure 6.6 and table 6.5), for which it frequently provides a low number of false positives and all the detected cells are really true cells, giving high precision values. On other hand, OIA-EDA detects more cells (recall higher than OIA-RDA) and works better detecting low and no stained cells due to they are normally isolated, as in `images4` or `image12` (see figures 6.4 and 6.5, and tables 6.3 and 6.4). Since both algorithms are available in the OralImmunoAnalyser (OIA) software for experts to run separately or in combination, we

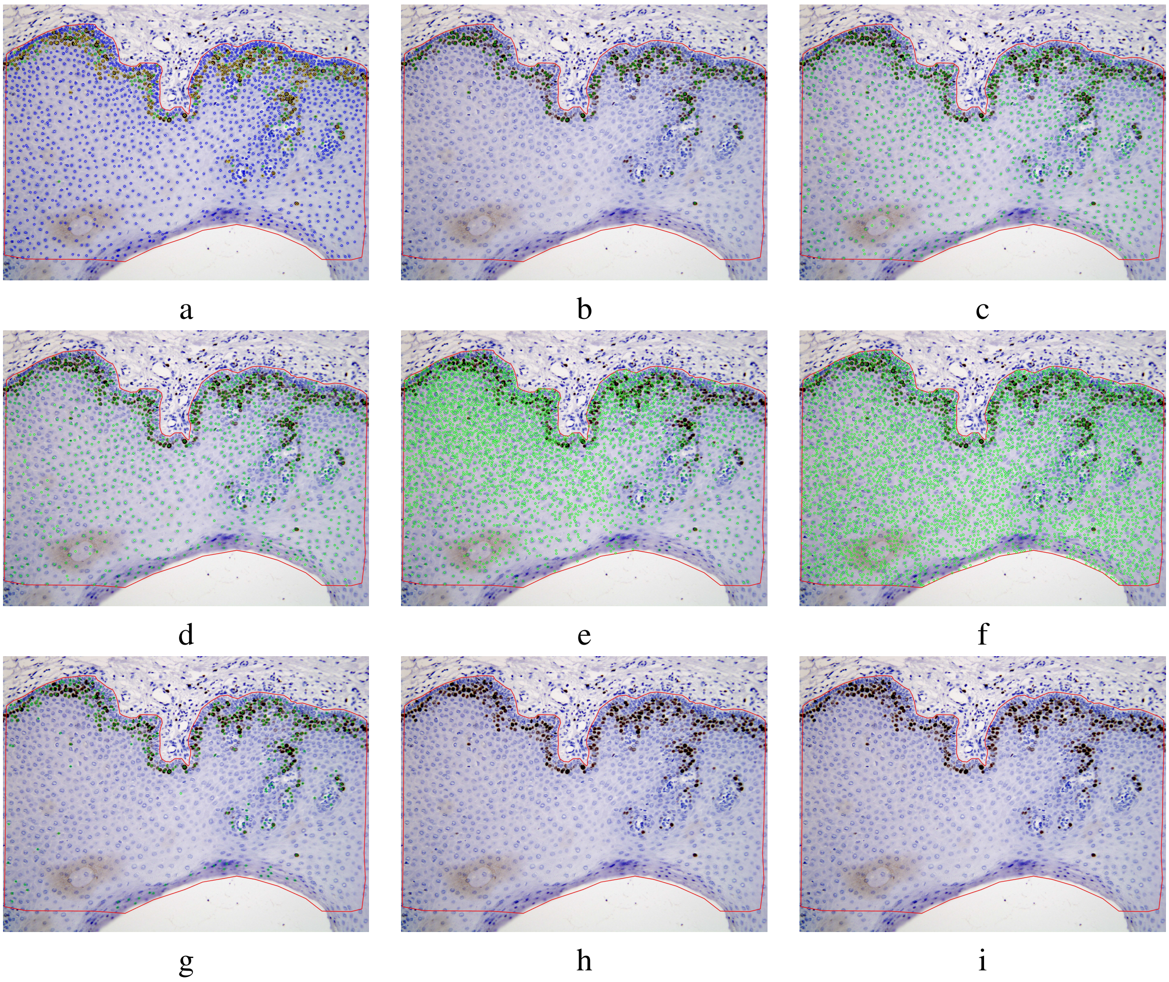


Figure 6.7: Example (`image34`) of unsupervised cell detection approaches (cells are overlapped to the IHC image): (b) OIA-RDA (option HIGH); (c) OIA-RDA (option LOW); (d) OIA-RDA (option WS); (e) OIA-EDA; (f) FH; (g) ChV (option MorphACWE); (h) SLIC ; and (i) kmeans-RGB. The groundtruth is shown in (a), where the color means the staining level of nuclei: yellow (highly stained), green (low stained) and blue (without staining).

experimented by adding the cells detected by both approaches (see the row with Method=OIA and Config=RDA-EDA in table 6.1). The performance metrics increased slightly to R=49.2%, P=50.9%, and F_1=46.4% and OIA-RDA-EDA emerges as the most effective unsupervised segmentation method.

6.4 Cell detection using deep learning

Deep learning networks U-Net and YOLO described in section 4.4 were applied to the segmentation of IHC images of the OIADB dataset employing the leave-one-image-out cross-validation configuration described in section 6.2 [3]. Average precision, recall, and F_1-score values are

Table 6.6: Precision, recall and F_1-score (columns three, four and five) to detect nuclei cells in IHC image `image17.tif`. The last three columns (SH, SL and SW) show the sensitivity for the high, low and without staining level of nuclei.

Method	Config.	Recall	Precision	F_1-score	A_{High}	A_{Low}	$A_{Without}$
OIA-RDA	HIGH	9.8	99.5	17.4	65	2.8	0
	LOW	51.3	82.7	63.3	78.2	42.8	47.6
	WS	52.8	82.9	64.5	81.8	46.1	48.1
OIA-EDA		61	38	46.8	58.9	60	61.5
FH		63.9	29.7	40.6	80	60.6	61.1
ChV	MorphACWE	17.2	92.1	29	60	36.1	6.6
Clustering	kmeans-RGB	0	0	0	0	0	0
	SLIC						

A_{High}: accuracy(%) for high stained nuclei, A_{Low}: sensitivity for low stained nuclei and $A_{Without}$: sensitivity for nuclei without staining.

computed for a collection of values for the parameters threshold and diameter. The threshold (called confidence in YOLO model) controls the heatmap intensity level at which centroids are captured, and the diameter defines the allowable distance between a predicted cells position and a ground-truth point for the prediction to be considered correct. Threshold values ranging from 0.1 to 0.9 with step 0.1 are considered. The diameter values are established taking into account the minimum diameter of cells in this dataset, which is 20 pixels. So, we use diameter values close to 20, ranging from 10 to 30 with step 1.

Table 6.7: Precision, recall and F_1-score and their standard deviation (SD) to detect cell's nuclei in IHC images of OIADB dataset using deep learning models.

Method	Config.	Recall		Precision		F_1-score	
		%	SD	%	SD	%	SD
OIA	RDA-EDA	49.2	18	50.9	11.7	46.4	7.2
DL	U-Net	74.5	8.6	76.8	7.1	75.3	6.2
	YOLO	68.9	8.7	80.9	6.7	74.0	5.9

Figure 6.8 shows the heatmaps with the average precision, recall, and F_1-score for all IHC images under different threshold and diameter values for the U-Net model (left column) and YOLO model (right column). The performance value is overlapped on the heatmap as a number (in percentage) and also coded as a different color. U-Net slightly outperformed YOLO, achieving the

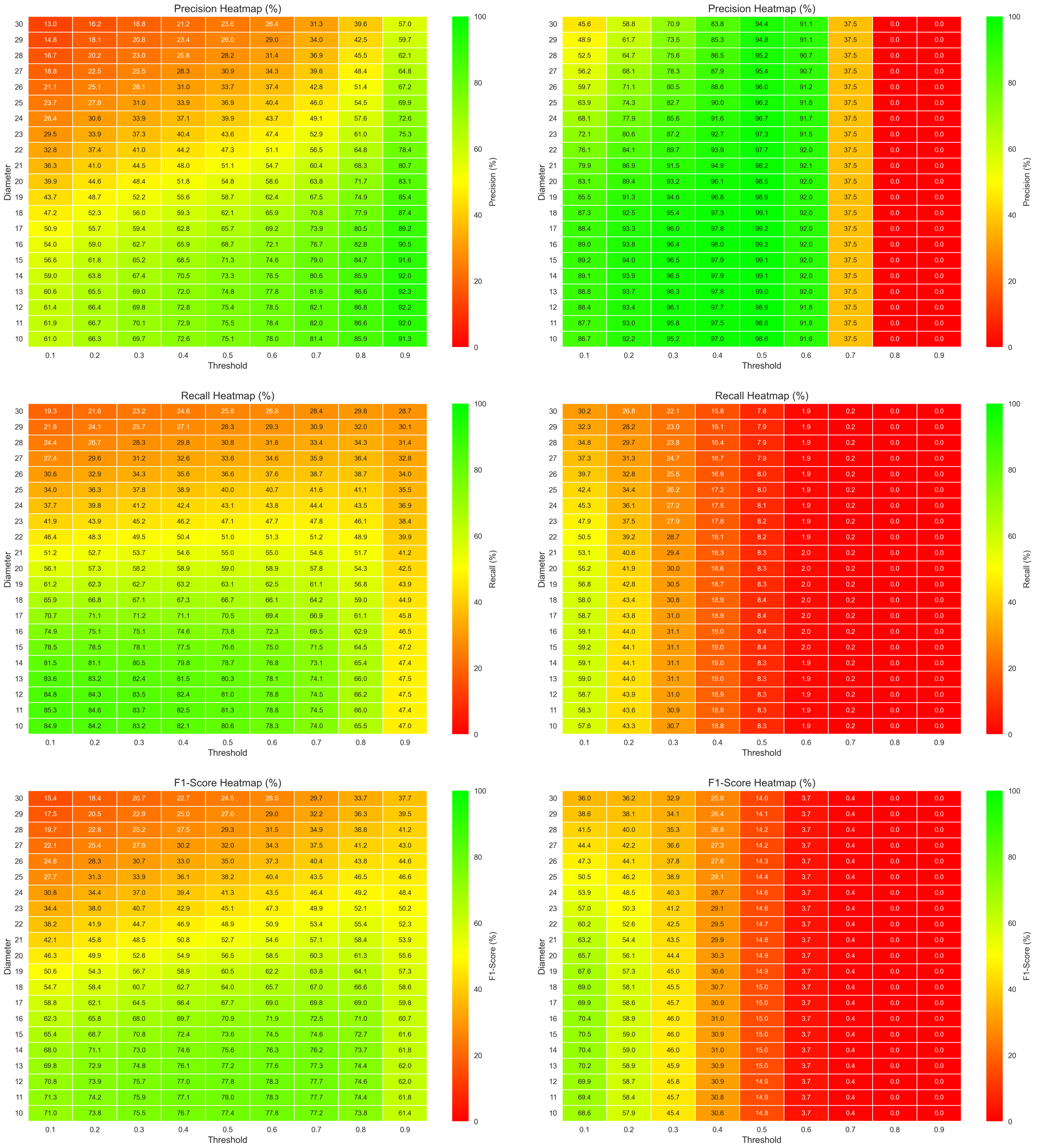


Figure 6.8: Average precision (upper row) for U-Net (left panel) and for YOLO (right panel); recall (middle row) for U-Net (left panel) and for YOLO (right panel); and F_1-score (lower row) for U-Net (left panel) and for YOLO (right panel), across the images in the dataset, using different threshold (x-axis) and diameter (y-axis) values.

highest F_1-score, approximately 77%, for thresholds in the range of 0.4 to 0.7 and diameters between 10 and 13 pixels. Precision reaches its maximum values at higher thresholds (0.8–0.9) within the same diameter range, indicating a reduction in false positives. In contrast, recall is the highest at lower thresholds (0.1–0.2), reflecting increased sensitivity to true positives under

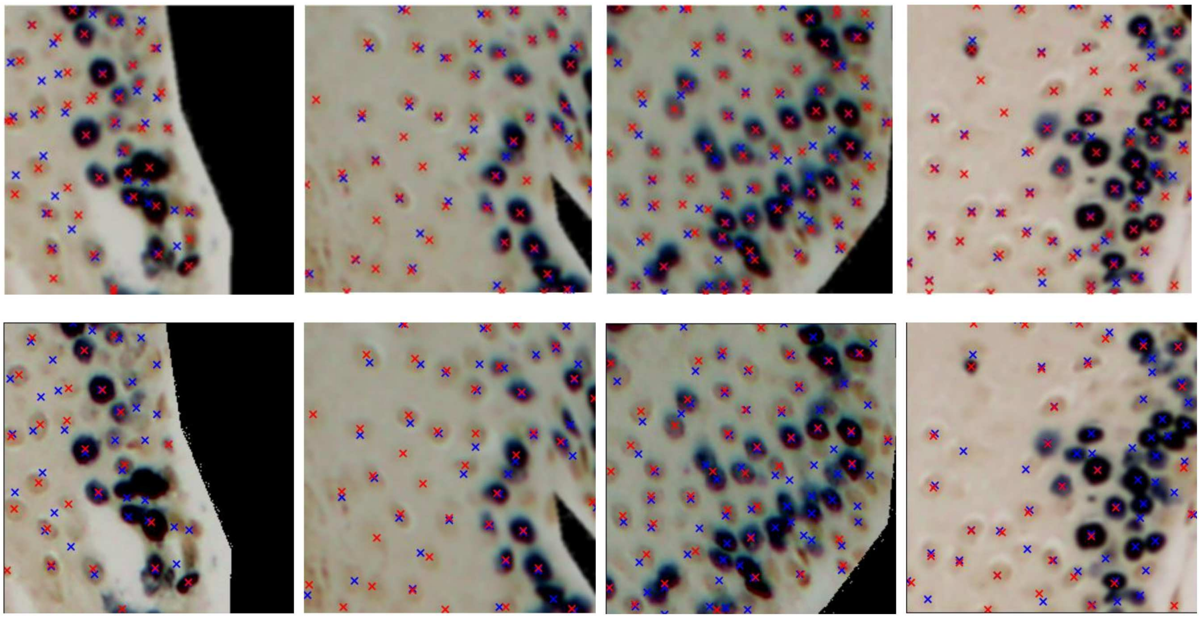


Figure 6.9: Samples from the dataset with the ground-truth and the predicted centroids, using U-Net (top panel) and YOLO (bottom panel), highlighted as blue and red crosses, respectively.

Table 6.8: F_1-score to detect cell's nuclei in IHC images of OIADB dataset. The last three columns (A_{High}, A_{Low} and $A_{Without}$) show the accuracy (%) detecting high, low and without staining level of nuclei.

Method	**Config.**	F_1**-score**	A_{High}	A_{Low}	$A_{Without}$
OIA	RDA-EDA	46.4	69.8	42.4	48.7
DL	U-Net	75.3	73.0	64.4	77
	YOLO	74.0	73.1	63.1	69.2

more permissive detection criteria. In this study, we focus on the parameter combinations that maximize the F_1-score as it provides a balanced trade-off between precision and recall. Specifically, we selected the operation point on a threshold of 0.6 and a diameter of 11 pixels. For the YOLO model, we selected the operation point in threshold = 0.1 and diameter = 14 pixels.

Table 6.7 shows the performance for deep learning models, including also the best unsupervised segmentation approach for comparative purposes (indicated by the algorithm OIA and configuration RDA-EDA). U-Net achieves the highest performance with F_1=75.3%, R=74.4% and P=76.8%. A low precision indicates a high rate of false positives, which may lead to incorrect detections, while a low recall suggests that many true positives are missed, reducing the sensitivity of the model. YOLO is slightly lower, in terms of F_1, achieving F_1=74.0%, R=69% and P=81%. Figure 6.9 shows some patches, extracted from the test set, with their true and pre-

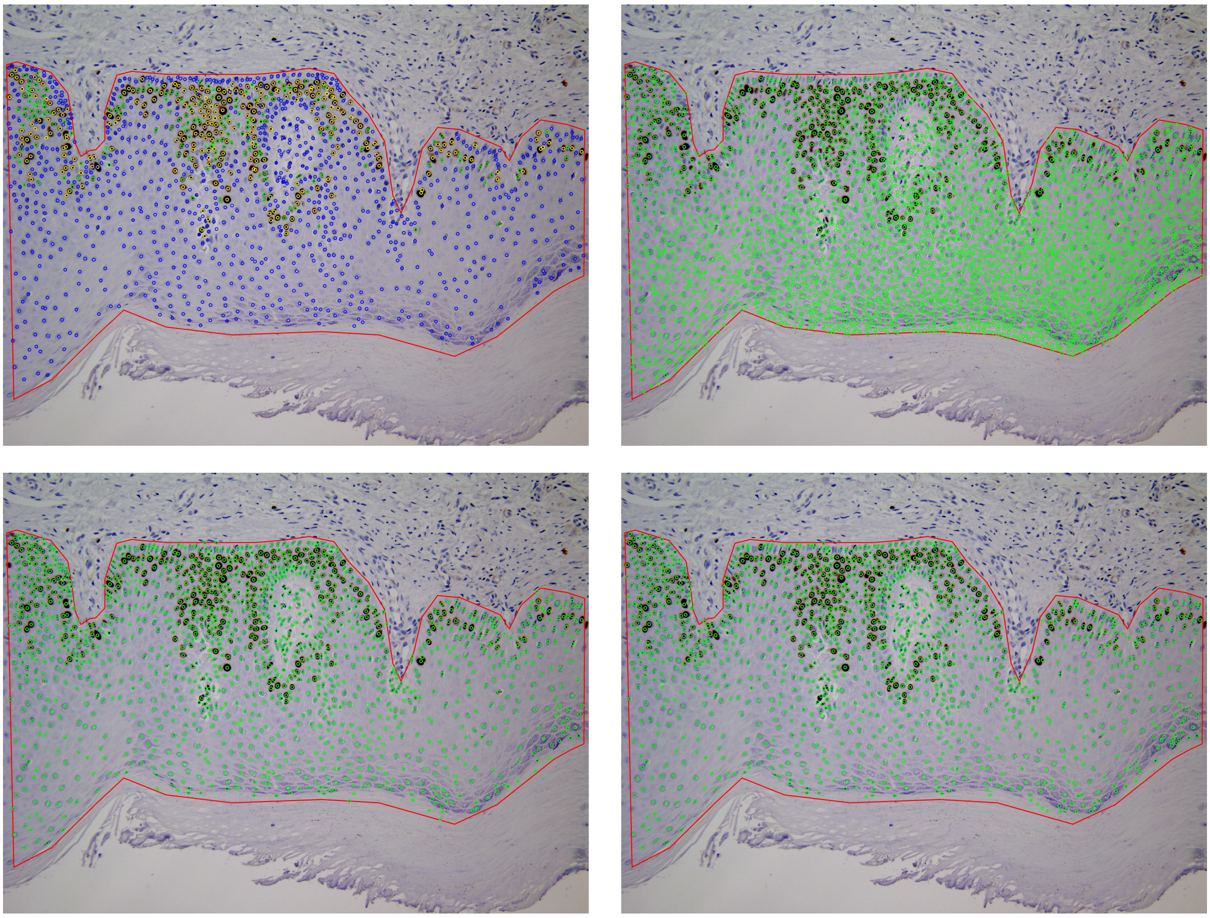


Figure 6.10: Example (`image 4`) of cell detection (cells are overlapped to the IHC image) for methods: OIA-RDA-EDA (upper right panel); U-Net (lower left panel) ; and YOLO (lower right panel). The ground truth of image 34 is the upper left panel, where the color means the staining level of nuclei: yellow (highly stained), green (low stained) and blue (without staining).

dicted centroids for the two DL models. Both deep learning networks show consistent behavior depending on the level of cell staining, as seen in the last three columns of table 6.8 (accuracy ranging from 63% to 77% for both DL models and all staining levels of cell's nuclei), and clearly outperform the unsupervised segmentation methods.

After collecting the predicted centroids from the individual patches of a test image, we map these centroids to the original image. Figures 6.10, 6.11, 6.12 and 6.13 illustrate the final prediction result on `image4`, `image12`, `image17` and `image34`, respectively, of the OIADB dataset using the best performing methods in this study: the unsupervised segmentation algorithm OIA-RDA-EDA and the deep learning approaches U-Net and YOLO. Table 6.9 shows the precision, recall and F_1-score for the images in figures from 6.10 to 6.13. For example, in `image34`, the F_1-score is 64.5% for OIA-RDA, 46.8% for OIA-EDA, 52.2% for OIA-RDA-

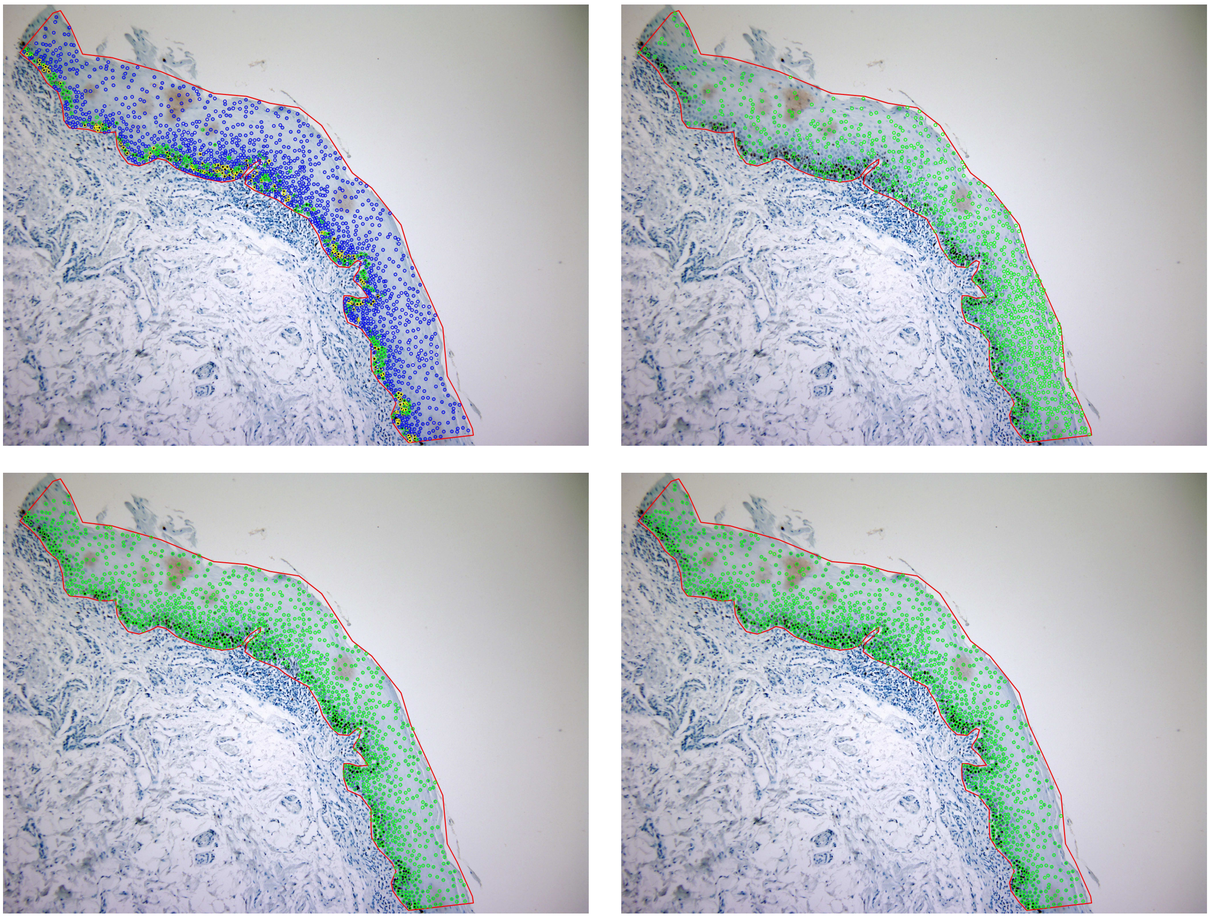


Figure 6.11: Example (`image 12`) of cell detection (cells are overlapped to the IHC image) for methods: OIA-RDA-EDA (upper right panel); U-Net (lower left panel) ; and YOLO (lower right panel). The ground truth of image 34 is the upper left panel, where the color means the staining level of nuclei: yellow (highly stained), green (low stained) and blue (without staining).

EDA, 87.5% for U-Net and 86.6% for YOLO; and for `image12`, the F_1-score is 22.6% for OIA-RDA, 42.4% for OIA-EDA, 46.5% for OIA-RDA-EDA, 72.1% for U-Net and 70.9% for YOLO.

6.5 Computation time

The unsupervised segmentation experiments were performed on a desktop computer with Intel® CoreTM i7-9700 processor at 3.6 GHz and 64 GB of RAM under Ubuntu 24.04. The OIA algorithms are done in the C/C++ programming language using the OpenCV 4.6 computer vision library, and the remaining algorithms are implemented in Python 3.12.3 and scikit-image 0.22 libraries.

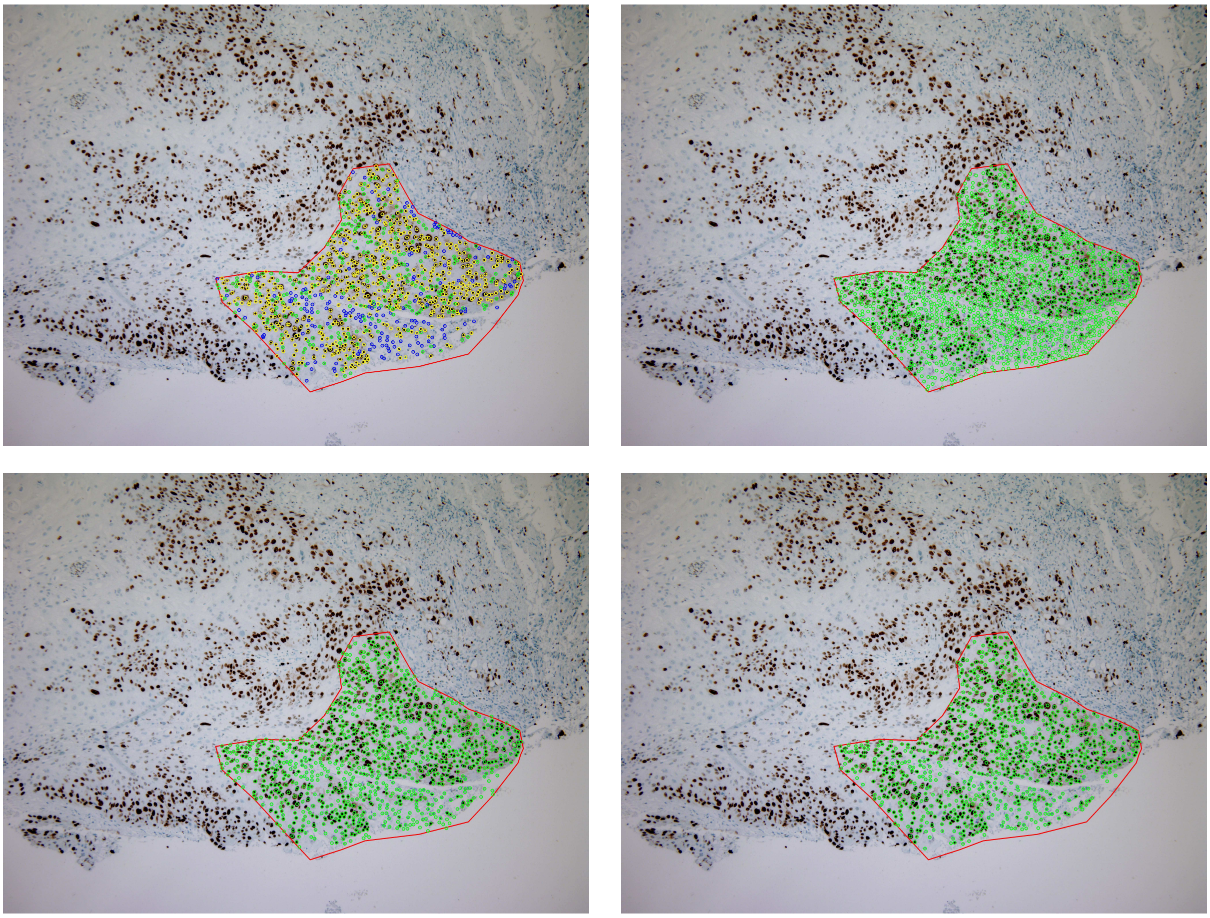


Figure 6.12: Example (`image 17`) of cell detection (cells are overlapped to the IHC image) for methods: OIA-RDA-EDA (upper right panel); U-Net (lower left panel) ; and YOLO (lower right panel). The ground truth of image 34 is the upper left panel, where the color means the staining level of nuclei: yellow (highly stained), green (low stained) and blue (without staining).

The deep learning experiments were conducted on a Windows Server 2022 Standard Evaluation 64-bit (10.0, Build 20348) with Intel(R) Xeon(R) Gold 6426Y (64 CPUs), ∼2.5GHz with 128 GB of RAM memory, equipped by a GPU NVIDIA RTX A4000 with 8 GB RAM memory. All deep learning experiments are implemented using Python programming language. U-Net is imported from the segmentation-models-pytorch (SMP) module[1]. The preparation of datasets and data load was performed using PyTorch (2.7.0.dev20250122+cu126). For YOLO, we used the implementation provided by Ultralytics[2].

The average computational time per image for each method is shown in table 6.10. For unsupervised segmentation methods, the computational time also includes the pre- and post-

---

[1] https://pypi.org/project/segmentation-models-pytorch/0.0.3/

[2] https://github.com/ultralytics

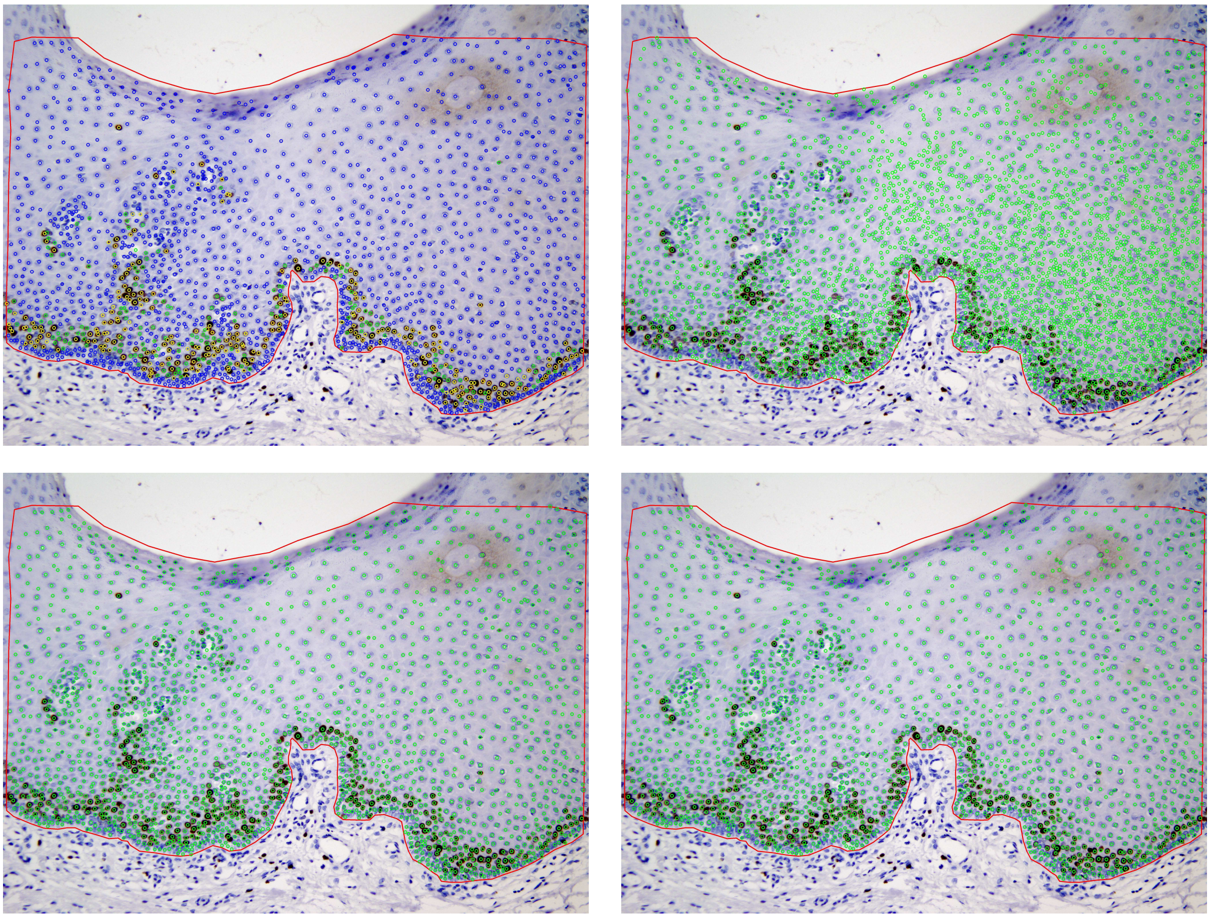


Figure 6.13: Example (`image 34`) of cell detection (cells are overlapped to the IHC image) for methods: OIA-RDA-EDA (upper right panel); U-Net (lower left panel) ; and YOLO (lower right panel). The ground truth of image 34 is the upper left panel, where the color means the staining level of nuclei: yellow (highly stained), green (low stained) and blue (without staining).

processing steps. The algorithms included in OralImmunoAnalyser are the fastest, requiring approximately only one second to process an image, making them optimal for interactive applications. The slowest algorithms are the active contours, which take nearly 10 minutes to process a single image. The methods based on clustering or graph cuts fall in between, with processing times ranging from a few seconds for K-means approaches to up to two minutes for the FH graph cut method. The last four rows of table 6.10, named U-Net and YOLO, report the average times spent by these networks to process a image, i.e. the average time spent to test the deep network on one image. The rows labeled with “using CPU” means the computational time needed by the DL technique to proccess a test image without employing specialized hardware, i.e a GPU. These test times, that do not include the time required to train the DL network, are similar to the fastest unsupervised segmentation methods. Training deep neural networks is very intensive,

Table 6.9: Precision, recall and F_1-score (columns three, four and five) to detect nuclei cells in differents IHC images. The last three columns (A_{High}, A_{Low} and $A_{Without}$) show the sensitivity for the high, low and without staining level of nuclei.

Method	**Config.**	**Recall**	**Precision**	F_1**-score**	A_{High}	A_{Low}	$A_{Without}$
Image 4							
OIA	RDA-EDA	28	75.9	40.9	89.5	78.4	71.6
DL	U-Net	83.2	91.8	87.3	95.5	93.4	90.5
	YOLO	84.4	88	86.2	95.5	92.8	85.1
Image 12							
OIA	RDA-EDA	61.1	37.5	46.5	44.1	27.7	38.6
DL	U-Net	82.1	64.3	72.1	54.4	54.5	67.3
	YOLO	81.6	62.7	70.9	58.1	49.1	65.7
Image 17							
OIA	RDA-EDA	42.6	59	49.5	68.1	49.1	40.9
DL	U-Net	80.5	76.8	78.6	81.5	69.3	70.7
	YOLO	85	73.5	78.8	81.7	62.9	59.8
Image 34							
OIA	RDA-EDA	41.2	71.1	52.2	86.1	62.8	69.2
DL	U-Net	85.5	89.7	87.5	90.0	85.6	90.1
	YOLO	85.6	87.7	86.6	90.4	84.4	87.6

even using dedicated hardware like GPUs. In our case, it takes approximately between 4 to 6 hours.

6.6 Discussion

Regarding the cell detection performance in tables 6.1 and 6.7, the best F_1=75.3% is provided by a supervised technique, the U-Net deep learning model. The best unsupervised segmentation technique, OIA-RDA-EDA, included in the OralImmunoAnalyser software [2], achieved F_1=46.4%. Figure 6.14 shows the average precision (blue bars), recall (green bars) and F_1-score (red bars) for the best supervised and unsupervised approaches tested on OIADB dataset. The vertical error bars show the standard deviation of each method. As expected, supervised techniques provided the highest performance results, because they use knowledge of the problem (location of cell's nuclei) through training before process a test image. Figure 6.15 shows the

Table 6.10: Computational time of segmentation approaches.

Method	Config.	Time (s)	Standard deviation
OIA-RDA	HIGH	0.91	0.29
	LOW	0.90	0.28
	WS	0.91	0.29
OIA-EDA		1.22	0.51
FH		118.74	52.1
ChV	MorphACWE	497.25	248.97
	MorphGAC	553.66	0.62
Clustering	kmeans-RGB	18.98	0.52
	kmeans-LAB	18.73	0.71
	SLIC	60.32	60.14
U-NET	using GPU	0.92	0.20
	using CPU	2.90	1.31
YOLO	using GPU	0.87	0.23
	using CPU	2.02	0.95

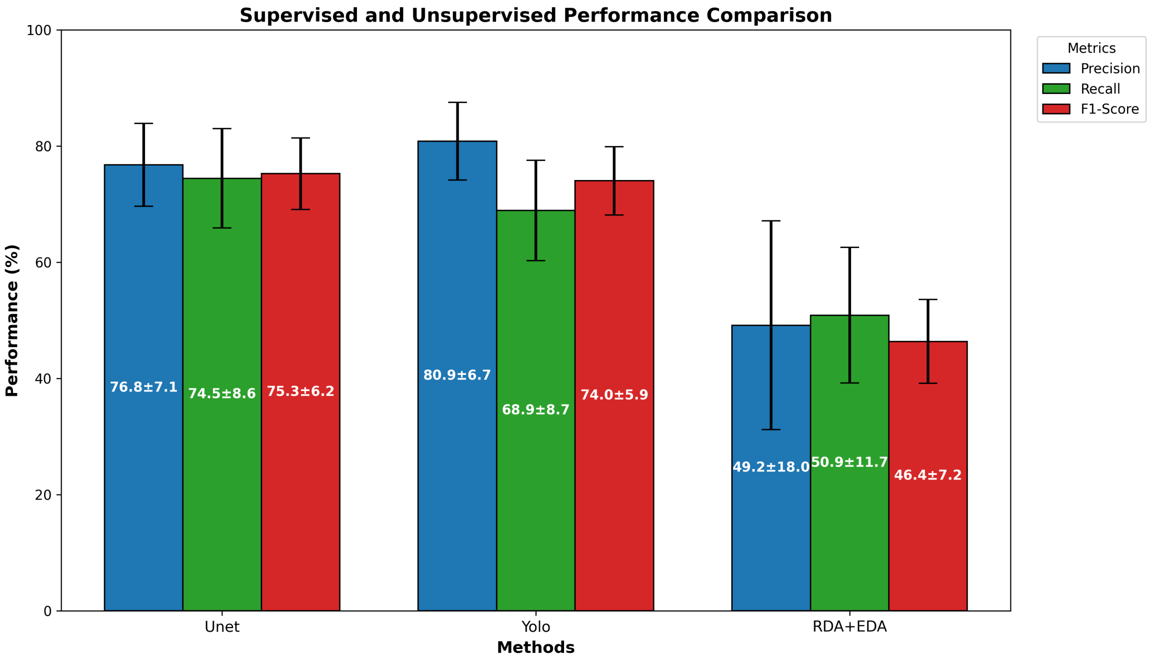


Figure 6.14: Average of precision, recall, and F_1-score for the best supervised and unsupervised segmentation algorithms tested on OIADB dataset.

precision, recall, and F_1-score achieved by the best methods for both approaches to detect cells in IHC images for each image in the OIADB dataset. A Wilcoxon ranksum test is used to compare the F_1-score of the best cell detection method (U-Net deep learning model) and of the other methods. This difference (p-value) is not statistically significant only for YOLO (p-value=0.358122), being significant for the others with $p \sim 10^{-14}$, which satisfy that $p < 0.05$. This means that both Deep Learning methods U-Net and YOLO clearly outperform the unsupervised methods as expected.

There are a number of difficulties when comparing the computational resources needed for supervised versus unsupervised segmentation approaches. Traditional segmentation methods such as edge detection, thresholding, and clustering are often made to function well on common hardware. Generally speaking, these techniques can be used on a general-purpose computer and often finish processing in a respectable amount of time, approximately one or two seconds for average images. Because of this, users with little processing power can utilize them.

However, supervised segmentation methods, specially those that rely on deep learning, require much more processing power. Due to the large number of calculations required for training and inference, these methods often require specialized hardware, such as graphics processing units (GPUs) or tensor processing units (TPUs). Many layers and parameters are commonly included in deep learning models used for segmentation/detection, which can result in increased memory needs and processing times. Consequently, the application of these techniques can be limited to settings with high-performance computer capability. One important point to mention here is the fact that DL inference (or test) time is very low compared to its training time (see Table 6.10). Another drawback for the practical use of deep learning is that the training stage needs a collection of annotated images. This requirement does not exist for the unsupervised methods, that process each image independently and do not need any collection, that may be difficult or impossible to gather or to annotate.

Furthermore, the deep learning models' training phase may be especially resource-intensive, frequently needing huge datasets and a significant amount of time to reach peak performance. This stands in stark contrast to conventional approaches, which often may be used immediately without requiring large datasets or intensive training. Overall, even though deep learning methods can produce better detection outcomes, especially in complex situations, their increased resource consumption and dependence on specialized hardware may make them impractical for some applications, particularly in settings with limited computational resources. Since deep learning models can be run on general-purpose (CPU-only) computers in a reasonable amount of time, the technological limitations for small biomedical units could be overcome if they have facilities for computing-intensive structures to train deep learning algorithms.

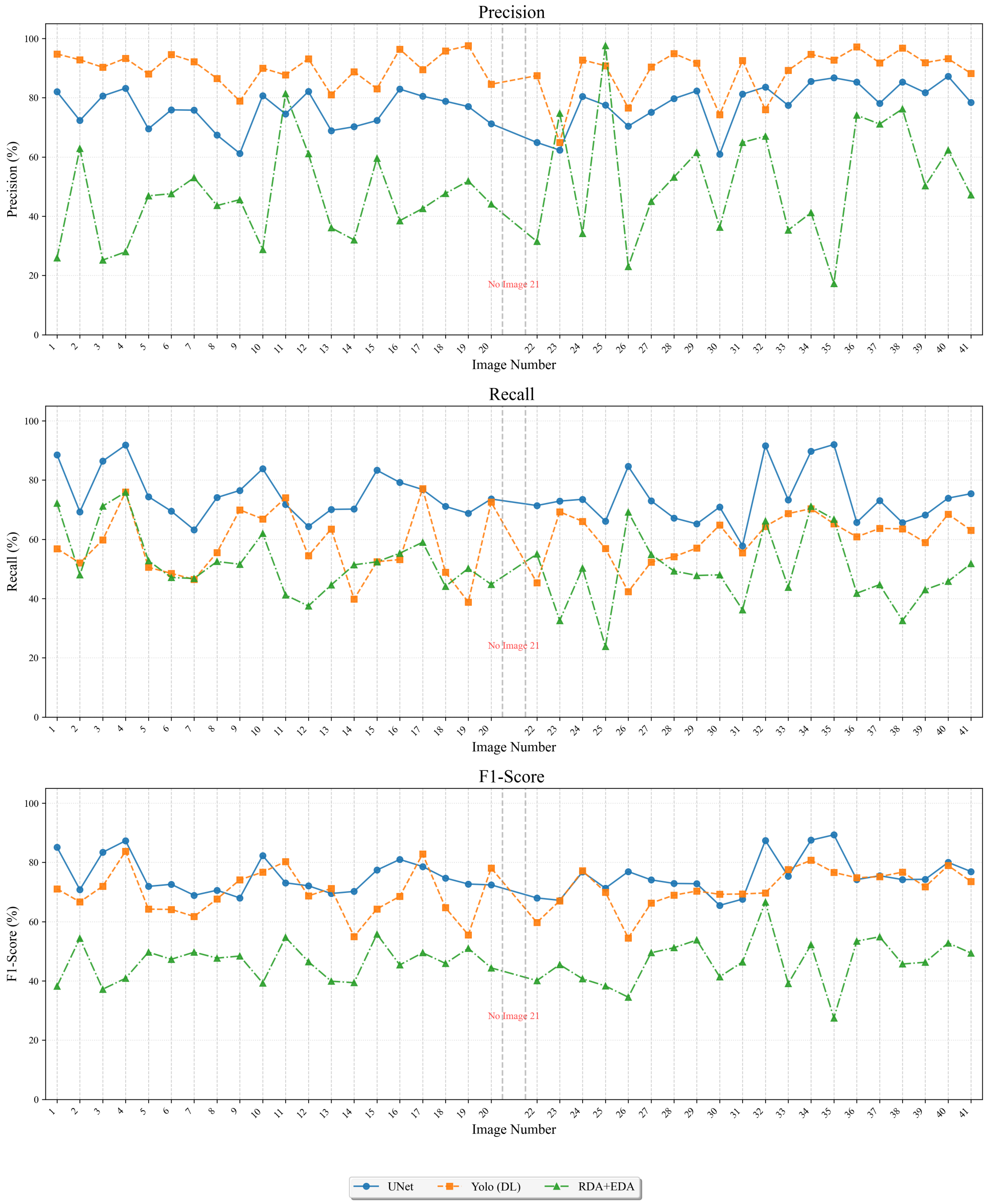


Figure 6.15: Precision (upper panel), recall (middle panel), and F_1-score (lower panel) for each image in the dataset for the U-Net model (blue line), OIA-EDA (orange line), and RDA+EDA (green line).

Note that the use of ground truth by Deep Learning methods (U-Net and YOLO) does not mean that comparison to unsupervised method is unfair. Rather, this is a fundamental difference between both methodological approaches. Their comparison is of practical relevance because highlights researchers between investing time in data annotation and supervised training or applying readily available unsupervised methods, reporting the trade-off between these two legitimate strategies for cell detection. The conclusions of our study are that supervised methods achieve superior performance, but at the cost of requiring training data and computational resources.

Recently, the study [4] tested the performance of deep learning-based digital image analysis systems relative to manual counting by expert pathologists in stomach cancer tissues using IHC images stained with ki-67 biomarker. The study compared three systems (AetherAI, 3D Histech, Roche) analyzing 239 microscopic specimens and found a considerable variability among the results. No deep learning-based application exceeded the average performance among human experts, and variability also occurred among the experts themselves according to the method of counting (whole TMA, box selection, or selection by hand-free selection). These results indicate the potencial, but also the current shortcomings, of deep learning systems in this field, evidencing the need of further research.

In summary, we presented a comprehensive evaluation of cell detection methods applied to immunohistochemical (IHC) images of oral tissue, comparing traditional unsupervised segmentation techniques with modern deep learning-based approaches. The results demonstrated that classical methods such as K-means clustering, edge detection, and active contours can provide acceptable performance under specific conditions. For example, highly stained cells are well detected by the region-based algorithm OIA-RDA. Although they are very fast (need only one or two seconds per image) and do not require specific hardware resources, they are generally limited in handling the complexity of some IHC images, particularly in the presence of artefacts or defects in sample preparation. Globally, the highest performance achieved is only F_1=46.4% using OIA-RDA-EDA. Regarding their performance in detecting cells with different levels of staining, region-based techniques such as OIA-RDA were highly sensitive in detecting highly stained cells (63.9% of highly stained cells) and achieved low sensitivity with the rest of the staining levels. However, recent research suggests that intensely stained cells are associated with specific cellular phases, being present more often in cases of invasive carcinoma or carcinoma in situ. This makes their detection particularly important. On the other hand, unstained or lightly stained cells identify cellular phases that may be consistent with normality or the presence of hyperplasia. Edge-based algorithms, such as OIA-EDA, perform uniformly with all levels of cell staining.

As expected, the supervised cell detection using deep learning models achieved better performance than unsupervised segmentation, because they use the problem knowledge during the training step before their use for cell detection. The U-Net network, employing heatmap regression to detect cell centroids, achieved the highest performance, F_1=75.3%. YOLOv8, adapted with synthetic bounding boxes, also showed strong results, F_1=74%. As well, the behavior of both methods is uniform with all staining levels of cells. This highlights the effectiveness of deep learning for robust and scalable cell detection in histopathological analysis.

Despite the promising results, the study acknowledges the limitations posed by the relatively small number of IHC images in the OIADB dataset and the reliance on centroid-only annotations, but cell drawing labeling requires a great deal of workload for professionals. However, we want to highlight that each image has on average more than 1000 cells, so the dataset has more than 40,000 labeled cells, which is a significant number of objects to detect. Despite all this, the generalizability of the results could be further improved by including a comprehensive comparative study evaluating our methods against other published cell detection approaches across larger standardized datasets to provide broader performance benchmarking. Additionally, we plan to extend our evaluation to include cross-dataset generalization studies to assess method robustness across different tissue types and imaging protocols.

CHAPTER 7

CONCLUSION

The current thesis includes two objectives: 1) the development of the OralImmunoAnalyser (OIA) software to perform quantitative analysis of immunohistochemical (IHC) images of oral tissue for the study of oral cancer; and 2) a comprehensive comparative evaluation of algorithms to detect cells in IHC images.

OIA is avilable from CiTIUS repository[1] for research purposes, and it was installed in the laboratory of the Faculty of Odontology of the University of Santiago de Compostela (Spain), where experts have evaluated its operation and usability. This software combines the automatic detection and classification of cells in the image with a friendly GUI that allows the experts to review the recognition before the calculation of statistical results. OIA provides the number of positive (stained) and negative (without staining) cells in the region of analysis, the percentage of cells for each staining level (high, low and no staining) and the distribution of the cell positivity in different layers of epithelium (basal, medium and superior).

The main conclusions from the OIA evaluation are:

1. From the automatic processing to cell detection and classification: 1) the cell detection module achieved a sensitivity of 64.4%, specificity of 93% and precision of 60.7%; and 2) cell classification in the three staining levels achieved an accuracy of 79.8%.

2. The computational time spent to process an images is aproximatly one second, and the time required to analyse each image is dominated by the needs of supervision. Despite OIA cannot operate fully automatically, it can accelerate considerably the analysis, saving 56.5% of time spent by the traditional manual counting of cells. OIA can be performed in the daily clinical practice, being a major advance over what is currently available.

3. The expert's perception about OIA achieves a mean score of 80.9 in the SUS (system usability scale) questionnaire, which means that the system is from good to excellent.

---

[1] https://citius.usc.es/transferencia/software/oralimmunoanalyser

4. OIA software is simple to use and install and has the following advantages:
    a) It works with a photograph taken under a microscope and not with a scan of the glass.
    b) It allows monitoring, i.e. to see the cells to be accounted for each category before image quantification.
    c) It can be used to annotated IHC images of oral tissue. Hence, OIADB dataset, containing 41 IHC images of oral leukoplakia annotated by experts, was created using OIA and it was published in CiTIUS repository[2].
    d) It allows to divide the epithelial strata into three thirds (basal / medium / superior), to compare the positivity of ki67 among layers and to study its relationship with the degree of epithelial dysplasia.

For these reasons, OIA is superior to other available tools and its use could be easily implemented in the daily practice of the biomedical labs. In addition, this possibility of supervision by the expert favors that OIA can be used as a tool in the teaching-learning process to instruct junior researchers in cell counting.

Chapter 6 focussed on a comprehensive evaluation of cell detection methods applied to immunohistochemical (IHC) images of oral tissue, comparing traditional unsupervised segmentation techniques with modern deep learning-based approaches. As expected, the main conclusion is that supervised methods outperformed unsupervised segmentation methods, mainly due to they use the problem knowledge during the training step before their use for cell detection. The highest performance was achieved by U-Net network, employing heatmap regression to detect cell centroids, with F_1=75.3%, follow by the another deep learning model, YOLOv8, adapted with synthetic bounding boxes, with F_1=74%. The best unsupervised segmentation technique, OIA-RDA-EDA, included in the OralImmunoAnalyser software, achieved only F_1=46.4%. Hence, a Wilcoxon ranksum test comparing F_1-score of the best cell detection method (U-Net deep learning model) and of the other methods concluded that U-Net is not statistically significant only for YOLO (p-value=0.358122), being significant for the others methods with $p \sim 10^{-14}$, which satisfy that $p < 0.05$. This means that both Deep Learning methods U-Net and YOLO clearly outperform the unsupervised methods as expected.

Other conclusions derived from the results presented in chapter 6 are:

1. In relation with unsupervised segmentation algorithms, such as K-means clustering, edge detection, and active contours can provide acceptable performance under specific conditions. For example, highly stained cells are well detected by the region-based algorithm

---

[2] https://gitlab.citius.gal/analyser/oiadb

OIA-RDA. But, they are generally limited in handling the complexity of some IHC images, particularly in the presence of artefacts or defects in sample preparation. Globally, the highest performance achieved is only F_1=46.4% using OIA-RDA-EDA.

2. Regarding their performance in detecting cells with different levels of staining, region-based techniques such as OIA-RDA were highly sensitive in detecting highly stained cells, while edge-based algorithms, such as OIA-EDA, perform uniformly with all levels of cell staining. The behavior of both deep learning methods is also uniform with all staining levels of cells.
3. Deep learning approaches need a collection of annotated images to train the models, which may be difficult or impossible to gather, while unsupervised methods are applied directly to IHC images.
4. Comparing processing power, unsupervised segmentation techniques can be run is a general purpose computer, without requiring specialized hardware, such as graphics processing units (GPUs).
5. Comparing computational time, unsupervised methods are faster than deep learning techniques, even when only testing time is only considered.

However, recent research suggests that intensely stained cells are associated with specific cellular phases, being present more often in cases of invasive carcinoma or carcinoma in situ. This makes their detection particularly important. On the other hand, unstained or lightly stained cells identify cellular phases that may be consistent with normality or the presence of hyperplasia.

Future work will aim to extend the deep learning models YOLO and U-Net to also make classifications of the cell staining levels of the detected objects. We are also planning to include a trained deep learning U-Net model in the OralImmunoAnalyser software. Additionally, we plan to extend our evaluation to include cross-dataset generalization studies to assess method robustness across different tissue types and imaging protocols.

APPENDIX A

ORALIMMUNOANALYSER USER GUIDE

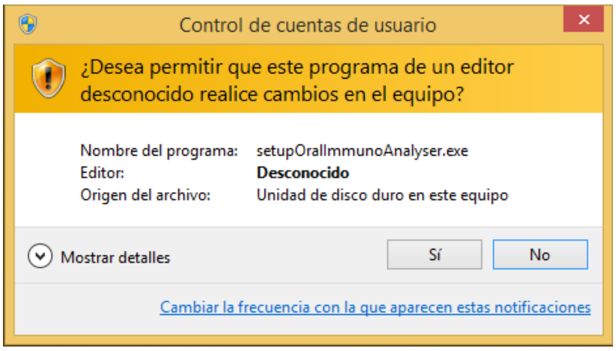
A.1 What is OralImmunoAnalyser?

OralImmunoAnalyser is a free software tool to quantitative analyse oral immunohistochemistry images of patients. It is a multi-platform software written in C/C++ programming language. It was developed by the Research center of Intelligent Tecnhologies (CiTIUS) [1] in colaboration with the Stomatology Department of Medicine and Odontology Faculty, both belonging to the University of Santiago de Compostela (Spain). OralImmunoAnalyser (OIA) examines automatically the immunohistochemistry images using advanced computer vision and machine learning techniques. But, sometime the automatic processing is not satisfactory to the expert. So, OralImmunoAnalyser includes a friendly GUI (Graphical User Interface) to review, draw, interact and visualize data.

OralImmunoAnalyser work with three types of files: image files and the text files XML (eXtensible Markup Language) and CSV (Comma-Separated Values). The image formats supported are the most frequently used formats like GIF, TIF, PNG, BMP, PPM, JPG, etc. There is one XML file per image to save the quantitative analysis, which includes the region of analysis and the staining level of the cells in the region. This information is overlapped on the image and it is frequently called overlays. The CVS file contains the statistical information of the quantitive analysis of the image. So, OralImmunoAnalyser manages in different files the image, the overlays with the analysis and the results. This organization allows that the experts can load, share or review an analysis every time later. The software counts the number and percentage of cells for each staining level (highly stained, low stained and without staning) in the region of analysis, and provides also these measures for different subareas in the region of analysis (basal, medial and superior areas).

---

[1] http://citius.usc.es/


Control de cuentas de usuario

¿Desea permitir que este programa de un editor desconocido realice cambios en el equipo?

Nombre del programa: setupOralImmunoAnalyser.exe
Editor: **Desconocido**
Origen del archivo: Unidad de disco duro en este equipo

Cambiar la frecuencia con la que aparecen estas notificaciones

Figure A.1: Window to install OralImmunoAnalyser.

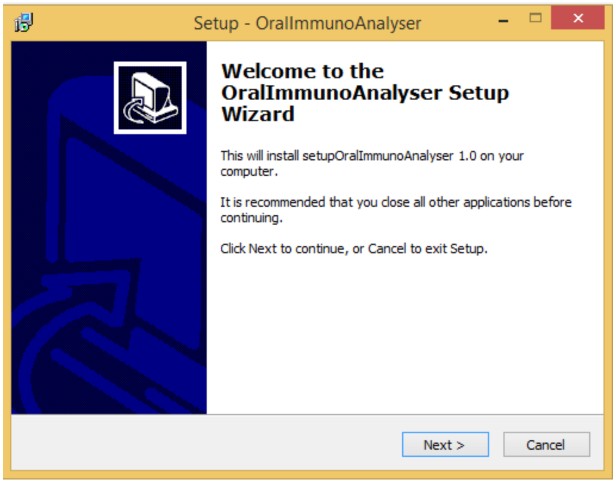
The current user guide is organized as follows: sections A.2 and A.3 describe the installation steps for Windows and Linux operating systems respectively; section A.4 describes the editing tools of the Graphical User Interface (GUI) of OralImmunoAnalyser. Section A.5 describes the configuracion of the working preferences. Sections A.6, A.7 and A.8 describe the functionality of File, Edit and View menus, respectively. The section A.9 describes how to do the quantitative analysis of one image. Section A.10 describe the statistical results provided by the software. Section A.11 describes the issues related to the classification of the stained level of cells. Finally, Section A.12 refers to Menu Help.

A.2 Windows installation

The installation process is the common process to install programs in Windows. First, do double click on the file setupOralImmunoAnalyser.exe. As OralImmunoAnalyser (OIA) is a foreing program, you need to set permisions to allow the installation of foreing programs. Then, a window opens asking us about to allow to install a foreing program setupOralImmunoAnalyser.exe on your computer. You must click "Yes" in figure A.1 to allow the installation of the program. After clicking "Yes", the window of Figure A.2 is open. If you click "Next", the program go on installing.

The next installation window is shown in Figure A.3, which asks if you want to create a direct access icon in the desktop. We recommend you activate the check box because it will be faster the access to OralImmunoAnalyser. Then, you click on the button "Next" to go to the next window (Figure A.4). It is a window to confirm if you want to install OralImmunoAnalyser (clicking with the mouse on button "Install") or cancel the installation (clicking the button "Cancel").

After starting the installation process, it will pop up a dialogue showing the installation process, as it can be seen in the Figure A.5. This process can take a few seconds. If you click on the button "Cancel", the installation process will be cancelled. When the installation process



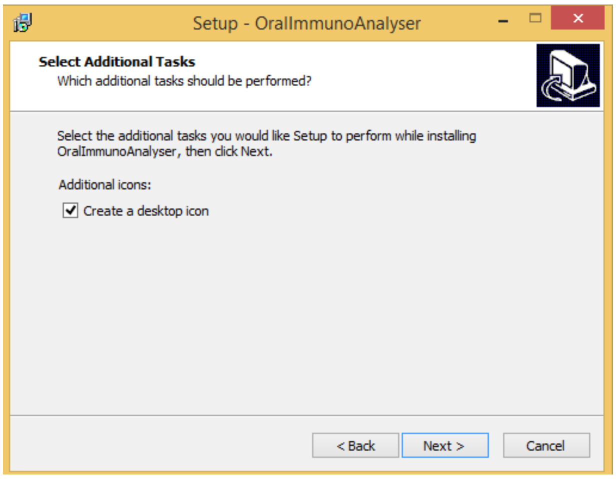
Welcome to the OralImmunoAnalyser Setup Wizard

This will install setupOralImmunoAnalyser 1.0 on your computer.

It is recommended that you close all other applications before continuing.

Click Next to continue, or Cancel to exit Setup.

Figure A.2: Window to install OralImmunoAnalyser.



Select Additional Tasks

Which additional tasks should be performed?

Select the additional tasks you would like Setup to perform while installing OralImmunoAnalyser, then click Next.

Additional icons:

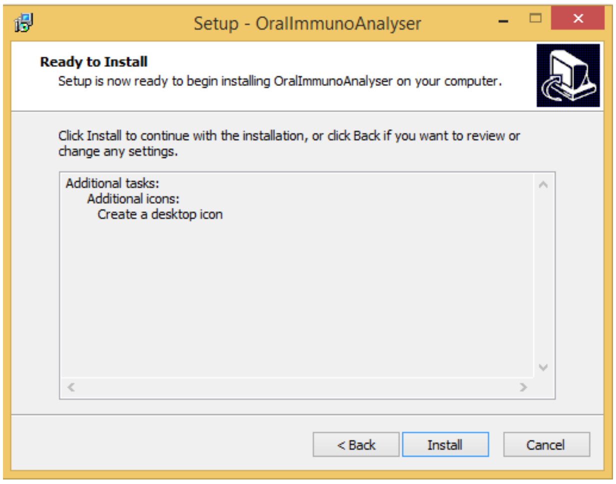
☑ Create a desktop icon

Figure A.3: Window to install OralImmunoAnalyser. Click on the check box to create a direct access in the desktop.

finished, it will be shown the window of Figure A.6. In this window, you click the "Finish" button to finish the installation process. If the check box *Lunch OralImmunoAnalyser* is marked, OralImmunoAnalyser will be run. Otherwise, OralImmunoAnalyser will not run now, but you can run every time later double clicking the desktop icon of OralImmunoAnalyser, which was created in the desktop during the installation process.

A.3 Linux installation

In Linux, there are several types of packages, and every distribution has its own preferred package format. Ubuntu distributions used the Debian packages (format DEB). It is in construction.



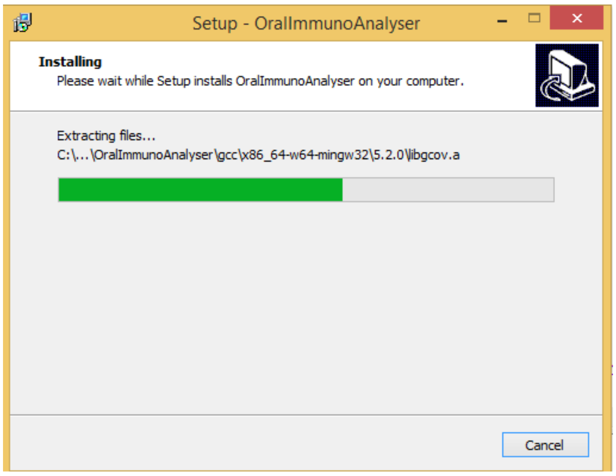
Ready to Install

Setup is now ready to begin installing OralImmunoAnalyser on your computer.

Click Install to continue with the installation, or click Back if you want to review or change any settings.

Additional tasks:
  Additional icons:
    Create a desktop icon

Figure A.4: Window to install OralImmunoAnalyser. Ready to install OralImmunoAnalyser.



Installing

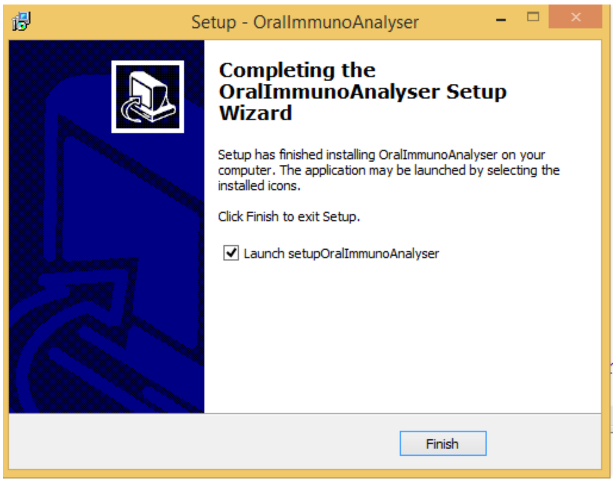
Please wait while Setup installs OralImmunoAnalyser on your computer.

Extracting files...
C:\...\OralImmunoAnalyser\gcc\x86_64-w64-mingw32\5.2.0\libgcov.a

Figure A.5: Window to install OralImmunoAnalyser. Installing OralImmunoAnalyser.

A.4 Run OralImmunoAnalyser

After installing OralImmunoAnalyser (OIA) in your computer using Windows, double click on the desktop icon of OralImmunoAnalyser to run it. OIA checks if the classifier was loaded. If there is a classifier trained in the computer, the pop-up message of upper panel of figure A.7 will be shown, otherwise, it is shown the pop-up message of lower panel of figure A.7, which informs that the classifier must be trained before classifing cells (see the section A.11 for the details of classifier). After clicking **Ok** in the window, the main window of Figure A.8 will be open. It encloses a menu bar (at the top of screen) containing all menu commands, a toolbar (under the menu bar) containing tools to access to the main functionality of OralImmunoAnalyser, and an


Setup - OralImmunoAnalyser

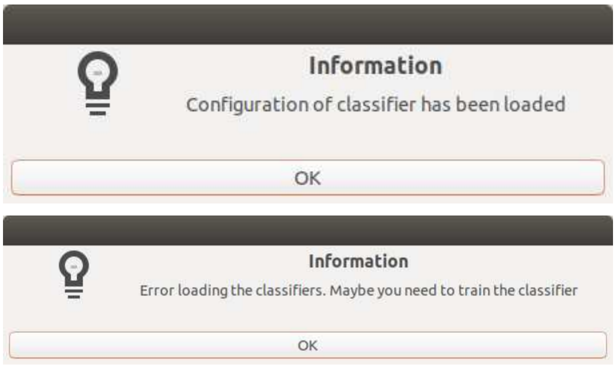
Completing the OralImmunoAnalyser Setup Wizard

Setup has finished installing OralImmunoAnalyser on your computer. The application may be launched by selecting the installed icons.

Click Finish to exit Setup.

☑ Launch setupOralImmunoAnalyser

Figure A.6: Window to install OralImmunoAnalyser. Completing the OralImmunoAnalyser Setup Wizard.


Information

Configuration of classifier has been loaded

Information

Error loading the classifiers. Maybe you need to train the classifier

Figure A.7: Pop-up windows show when OIA is run to inform if the classifier is loaded or not.

image window (at the bottom of screen) in which the immunohistochemistry images are open. The menu bar lists all **OralImmunoAnalyser** commands. It is organized in six menus:

1. **File**: Basic file operations (opening images and XML files, saving the overlays on the image in files and statistical results); set and load the working preferences; and exit of OralImmunoAnalyser.
2. **Edit**: redo and undo operations, fit the image to the window size and set the image to the original size.
3. **View**: show or hide the processing panel, also called lateral panel.
4. **Analysis**: provides functionalities to calculate the results of many images jointly.
5. **Classification**: provides functionalities to classify cells into their stained level (high stained, low stained, without staining) and to train the classifier.

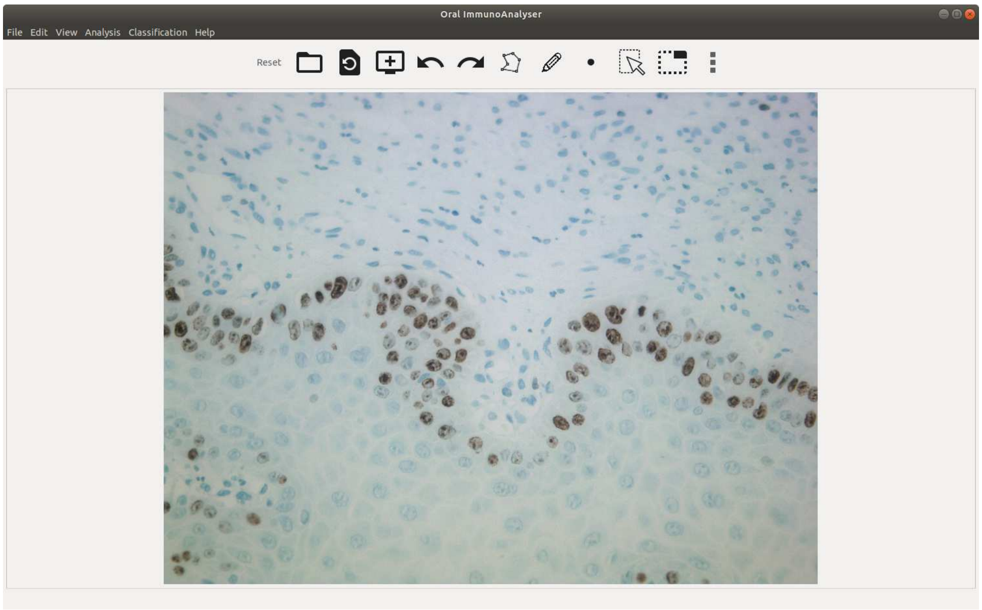


Figure A.8: Main window of OralImmunoAnalyser.

6. **Help**: the help functionality.

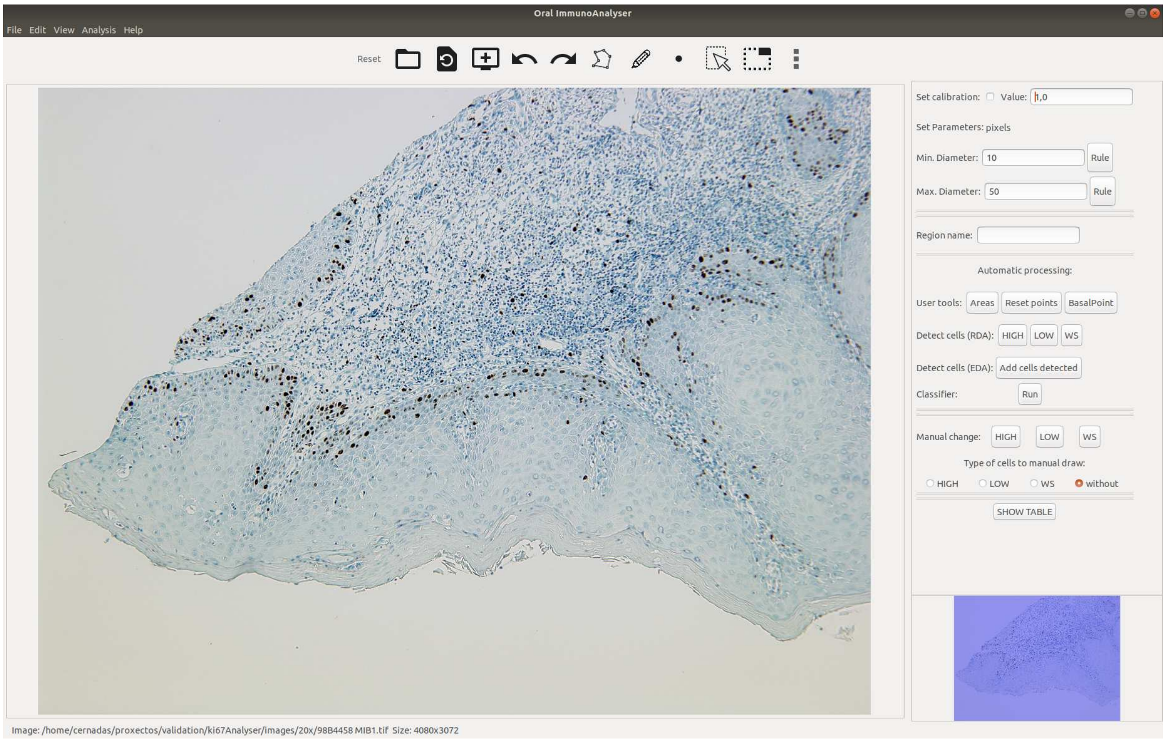
The toolbar is a fast access to the main functionality of OralImmunoAnalyser, which contains the following icons (if the mouse is put on the icon, a pup up message showing the functionality of the icon is open):

1. **Reset**: (first icon) clear all the objects drawn or overlayed on the image.
2. **Open** : (second icon) open a dialog to select the immunohistochemistry image to be open (see section A.6).
3. **Zoom Fit**: (third icon) fit the image zoom to the image window.
4. **Original Zoom**: (fourth icon) set the original image zoom.
5. **Undo**: (fifth icon) when you click this button, undo the last overlay deleted on the image (operate only with the overlays manually deleted).
6. **Redo**: (sixth icon) when you click this button, redo the overlays deleted (operate only with the overlays manually deleted).
7. **Draw with points**: (seventh icon) activate the draw of regions tool. While this button is pressed, the user draws a region marking points with the left button of the mouse and finishes the outline of the region when the user click the middle button of the mouse.

8. **Draw freehand**: (eighth icon) activate the draw freehand regions tool. When this button is pressed, the user can draw a freehand region pressing the left button of mouse and keeping it pressed while you are drawing the region. When you release the mouse button, the region is finished.
9. **Draw points**: (ninth icon) when this button is activated, you can mark points, which can represent points of a region, basal point or cells (as the software OIA considers the cells as points), on the image clicking the left button of the mouse.
10. **Select**: (tenth icon) when this button is activate, you can select an object drawn on the image. The object is selected clicking into the region (for select regions) or near the point (to select points) with the left button of the mouse. To select more than one object, keep the key **Ctrl** or key "Control" pressed while selecting objects.
11. **Select with rectangle**: (eleven icon) when this button is activate, you can select many objects drawn on the image drawing a rectangle. For that, click the left button of the mouse and keeping the button pressed, drag the mouse to draw a rectangle and release the left button. The objects inside the rectangle are selected.
12. **Lateral panel**: (twelfth icon) open the lateral panel, which will be used to process and analyse the open image (see figure A.9).

In the following sections, the functionality of the Graphical User Interface (GUI) of OralImmunoAnalyser will be described. The submenus of the File menu from the menu bar are:

1. **Open image**: open a dialogue window to load a new image.
2. **Open image and XML**: open a dialogue window to select the image to load. But, if the XML file for that image exists, its overlays are also loaded and drawn on the image. To do this process, the XML file must be named with the same name as the image and it is in the XML folder specified in the working preferences.
3. **Open XML**: open a dialogue window to select the XML file and put its content on the image loaded. The user controls that this XML file is suitable to the image loaded in OralImmunoAnalyser.
4. **Save XML**: open a dialogue window to select the name of XML file in which you want to save the overlays drawn on the image. By default, OralImmunoAnalyser sets the name of the image file, but with extension XML, and it is stored in the XML folder specified in the working preferences.



Set calibration: ☐ Value: 1,0

Set Parameters: pixels

Min. Diameter: 10

Max. Diameter: 50

Region name:

Type of cells to manual draw:
○ HIGH ○ LOW ○ WS ● without

Image: /home/cernadas/proxectos/validation/ki67Analyser/images/20x/98B4458 MIB1.tif Size: 4080x3072

Figure A.9: Main window of OralImmunoAnalyser with a typical immunohistochemistry image loaded and the lateral panel open.

5. **Export CSV**: open a dialogue window to select the name of CSV file in which you want to save the statistical analysis of the image. By default, OralImmunoAnalyser sets the name of the image file, but with extension CSV, and it is stored in the CSV folder specified in the working preferences.

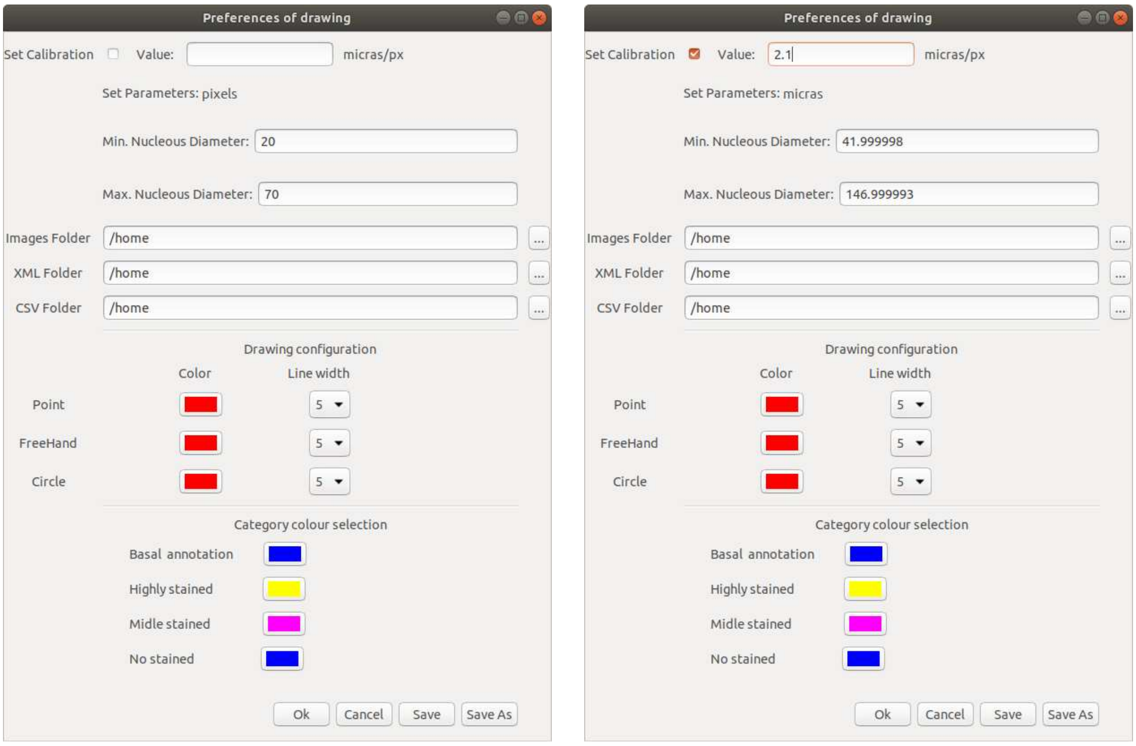
6. **Preferences**: open a dialogue window to set our working preferences in OralImmunoAnalyser (working directories, colour and width of lines, calibration and diameters). The working preferences can be saved in the computer to be available in next sessions.

7. **Reset preferences**: reset all the OralImmunoAnalyser preferences set previously in the computer.

8. **Load preferences**: internally, the preferences are saved in a XML file. This option allows to load the working preferences from an external XML file.

9. **Exit**: exit of the program.

Section A.5 describes the configuracion of the working preferences in OralImmunoAnalyser.



Preferences of drawing (left panel)

Set Calibration ☐ Value: micras/px

Set Parameters: pixels

Min. Nucleous Diameter: 20

Max. Nucleous Diameter: 70

Images Folder /home

XML Folder /home

CSV Folder /home

	Color	Line width
Point	🟥	5
FreeHand	🟥	5
Circle	🟥	5

Category colour selection	
Basal annotation	🟦
Highly stained	🟨
Midle stained	🟪
No stained	🟦

Preferences of drawing (right panel)

Set Calibration ☑ Value: 2.1 micras/px

Set Parameters: micras

Min. Nucleous Diameter: 41.999998

Max. Nucleous Diameter: 146.999993

Images Folder /home

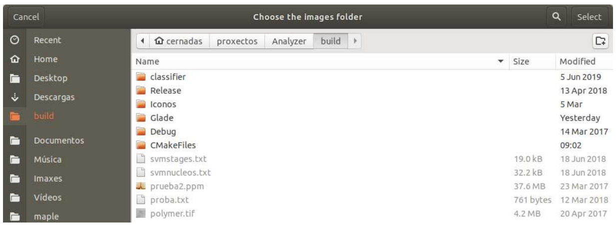
XML Folder /home

CSV Folder /home

	Color	Line width
Point	🟥	5
FreeHand	🟥	5
Circle	🟥	5

Category colour selection	
Basal annotation	🟦
Highly stained	🟨
Midle stained	🟪
No stained	🟦

Figure A.10: Window to configure the preferences of OralImmunoAnalyser: default configuration (left panel) and after setting the calibration (right panel).



Name	Size	Modified
classifier		5 Jun 2019
Release		13 Apr 2018
Iconos		5 Mar
Glade		Yesterday
Debug		14 Mar 2017
CMakeFiles		09:02
svmstages.txt	19.0 kB	18 Jun 2018
svmnucleos.txt	32.2 kB	18 Jun 2018
prueba2.ppm	37.6 MB	23 Mar 2017
proba.txt	761 bytes	12 Mar 2018
polymer.tif	4.2 MB	20 Apr 2017

Figure A.11: Window to choose a working directory.

A.5 Set preferences

Selecting the submenu “Preferences” of menu File, the window shown in Figure A.10 (left panel) will be open. The items to configure are:

1. **Calibration**: it is the number of micrometers per pixel in the image, which depends on the image acquisition system. It allows to set the units to measure the objects (region of analysis in our case).

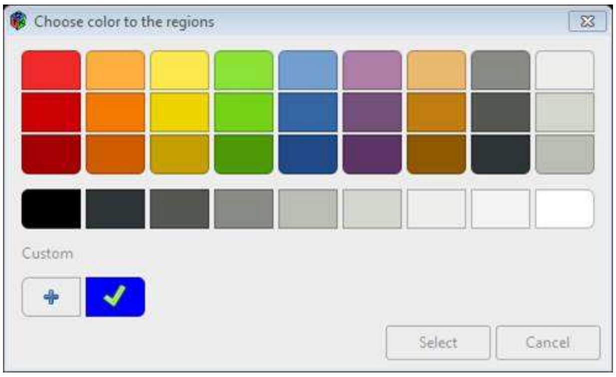


Figure A.12: Window to choose the colour to draw cells on the image.

2. **Diameters**: it allows to put the minimum and maximum size of the cells in the images. These parameters can be set in pixels (if the calibration is not set) or micras (if the calibration is set).

3. **Working directories**: it allows to set the default directories of the images, XML files and CSV files to use OralImmunoAnalyser.

4. **Drawing configuration**: it allows to set the colour and line width preferences to draw the objects overlapped to the images.

5. **Category colour selection**: it allows to set the colour to visualize different categories in OralImmunoAnalyser. The categories used by the software are: 1) the stained level of the cells, which we consider "high stained", "low stained" and "without staining"; and 2) the colour to visualize the basal point on the image, which is set using the button **BasalPoint** in the lateral panel (see section A.9.2 for more details).

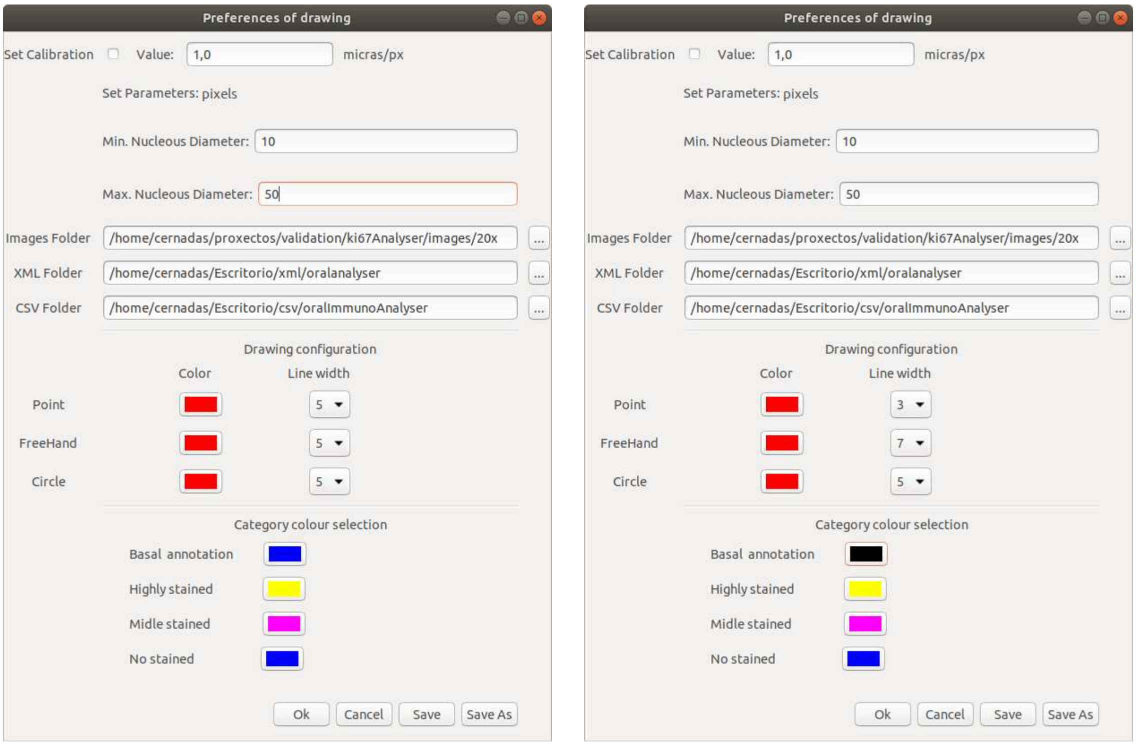
6. **Set or change the configuracion**: the buttons **Ok**, **Cancel**, **Save** and **Save As** at the bottom of Figure A.10 set, cancel or save the configuration.

A.5.1 Calibration

The top of Figure A.10 shown a check box after the label *Set Calibration*. This check mark assures that the Calibration real value has been added, if it is non-checked, then the calibration is considered in pixels. After the label *Value*, and if the check box of calibration is active, you can set the number of micras or micrometers per pixel in the image (press the "Enter" key after put the value). The calibration value are fixed by the digitalization process, depending on the magnification used in the microscope and the spatial resolution of the digital camera connected to it. The right panel of Figure A.10 shows the left panel after activating the calibration and set its value to 2.1 micras per pixel. As it can be seen comparing both panels in Figure A.10, when the check box of calibration is modified, the units and values of diameters change.

A.5.2 Diameters

The automatic detection of cells in OralImmunoAnalyser improves if the user provides the minimum and maximum diameter of the cells that the user want to detect. These parameters can be set after the labels *Min. Diameter* and *Max. Diameter* respectively, puting the values of minimum and maximum diameter (in micrometers if the calibration is active and in pixels if the check box of calibration is not marked). Obviously, if the minimum diameter set by the user is higher than the maximum diameter or the maximum diameter set is lower than the minimum diameter, the value of diameter is not updated. After set each diameters, you must press the “Intro” or “Enter” key to update the value in OralImmunoAnalyser. The minimum and maximum diameters of the cells to recognize can also be set on the lateral panel as text or graphically (see section A.8).



Left panel:

Set Calibration ☐ Value: 1,0 micras/px
Set Parameters: pixels
Min. Nucleous Diameter: 10
Max. Nucleous Diameter: 50
Images Folder /home/cernadas/proxectos/validation/ki67Analyser/images/20x
XML Folder /home/cernadas/Escritorio/xml/oralanalyser
CSV Folder /home/cernadas/Escritorio/csv/oralImmunoAnalyser

Drawing configuration

	Color	Line width
Point	🟥	5
FreeHand	🟥	5
Circle	🟥	5

Category colour selection

Basal annotation	🟦
Highly stained	🟨
Midle stained	🟪
No stained	🟦

Right panel:

Set Calibration ☐ Value: 1,0 micras/px
Set Parameters: pixels
Min. Nucleous Diameter: 10
Max. Nucleous Diameter: 50
Images Folder /home/cernadas/proxectos/validation/ki67Analyser/images/20x
XML Folder /home/cernadas/Escritorio/xml/oralanalyser
CSV Folder /home/cernadas/Escritorio/csv/oralImmunoAnalyser

Drawing configuration

	Color	Line width
Point	🟥	3
FreeHand	🟥	7
Circle	🟥	5

Category colour selection

Basal annotation	⬛
Highly stained	🟨
Midle stained	🟪
No stained	🟦

Figure A.13: Window of preferences of figure A.10 after: setting the working directories (left panel) and setting the drawing configuration and categories colour (right panel).

A.5.3 Working directories

After the labels *Images Folder*, *XML Folder* and *CSV Folder* in Figure A.10, there are their correspondent entry widgets to visualize the default directory to store the images, XML files and CSV files respectively. After each entry widget there is a button with three points. Clicking this button open the folder chooser dialogue of Figure A.11 to choose the work directory for images and the left panel of Figure A.13 shows the preferences window after modifying the working directories and the diameters.

A.5.4 Drawing configuration

You can change the colour and line width to draw the regions and points (cells) overlapped to the image. Clicking with the mouse the red button after the label *FreeHand* or *Point* opens the colour chooser dialogue shown in Figure A.12. Selecting a colour in this window and pressing the button “Select”, you change the drawing colour to freehand or point objects. Clicking in the button after the colour button, a drop-down list is open to select the line width to the freehand or point object. Points are shown as a hole circle in order to visualise the cells.

A.5.5 Category colour selection

As in the section A.5.4, clicking the colour button after the labels of the fundamental categories: basal point, hight stained, low stained and no-stained open the colour chooser dialogue, in which you can change the colour to represent that categories in OralImmunoAnalyser. Figure A.13 shows the figure A.10 in which the width and colour are set for some categories.

A.5.6 Set or change the configuration

The four buttons **Ok**, **Cancel**, **Save** and **Save As** at the bottom of Figure A.10 have the following functionalities:

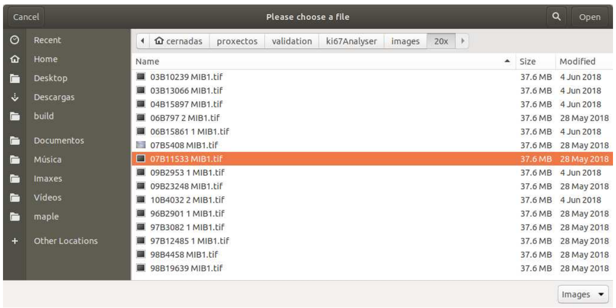
1. Clicking the button **Ok**, you set these preferences to the present working session.
2. Clicking the button **Cancel**, you cancel the operation of setting the preferences and the preferences will not be updated.
3. Clicking the button **Save**, you set these preferences to the present working session and to the following working sessions in the future.
4. Clicking the button **Save As**, you set these preferences to the present working session and OralImmunoAnalyser saves these preferences in a XML file selected by the user, which

can be loaded every time using the submenu “Load preferences” of menu File in this computer or other one. This option allows to share the working preferences with other expert (it may be possible that the working directories do not work correctly).

A.6 File menu

6 FILE MENU

The items available within **File** menu are: **Open Image**, **Open Image and XML**, **Open XML**, **Save XML**, **Export CSV**, **Preferences**, **Load Preferences** and **Exit**. The **Exit** menu or the **X** button in the top-right side of window quit OralImmunoAnalyser. As mentioned, OralImmunoAnalyser works with three types of files: image files, XML (eXtensible Markup Language) and CSV (Comma-Separated Values) files. There is one XML file per image, which saves the contour of region of analysis with the cells (visualize as points in the image window) detected, in order to allow that this analysis can be loaded into the software in other instant of time. The CVS file exports the statistical information of the quantitave analysis of the image.

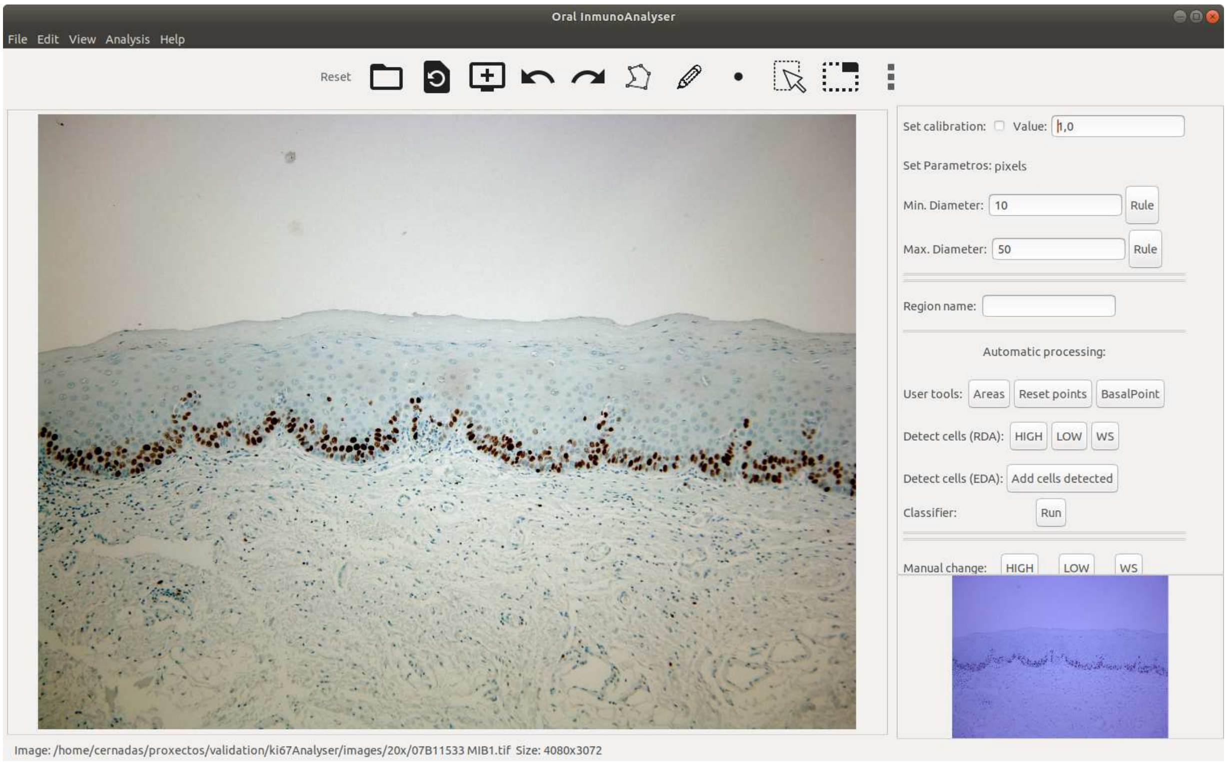


Name	Size	Modified
03B10239 MIB1.tif	37.6 MB	4 Jun 2018
03B13066 MIB1.tif	37.6 MB	4 Jun 2018
04B15897 MIB1.tif	37.6 MB	4 Jun 2018
06B797 2 MIB1.tif	37.6 MB	28 May 2018
06B15861 1 MIB1.tif	37.6 MB	4 Jun 2018
07B5408 MIB1.tif	37.6 MB	28 May 2018
07B11533 MIB1.tif	37.6 MB	28 May 2018
09B2953 1 MIB1.tif	37.6 MB	4 Jun 2018
09B23248 MIB1.tif	37.6 MB	28 May 2018
10B4032 2 MIB1.tif	37.6 MB	4 Jun 2018
96B2901 1 MIB1.tif	37.6 MB	28 May 2018
97B3082 1 MIB1.tif	37.6 MB	28 May 2018
97B12485 1 MIB1.tif	37.6 MB	28 May 2018
98B4458 MIB1.tif	37.6 MB	28 May 2018
98B19639 MIB1.tif	37.6 MB	28 May 2018

Figure A.14: File chooser dialogue to choose the image to be loaded in OralImmunoAnalyser.

The **Open Image** submenu or the second button of the toolbar opens the file chooser dialogue of Figure A.14. The directory open is the one set in the configuration of the preferences (see the section A.5). Choose the path to the image and click the button **Open** in the bottom of the window to load the image in OralImmunoAnalyser. Take care that the image path is not biggest than 256 characters or the image path contains rare symbols, because OralImmunoAnalyser could not work correctly. Figure A.15 shows an image loaded in OralImmunoAnalyser. When an image is loaded, the lateral panel is open and in the bottom of the main window appear a state bar with the name of image and its size.

If the image was analysed in other instant and the XML file was stored, you can load this XML file and overlap to the image selecting the submenu **Open XML**. This operation opens a file chooser dialogue to choose the XML file (see the Figure A.16) and overlap the region outline and cells to the image, as can be seen in Figure A.17. The user must select the right XML file



Value: 1,0

Set Parametros: pixels

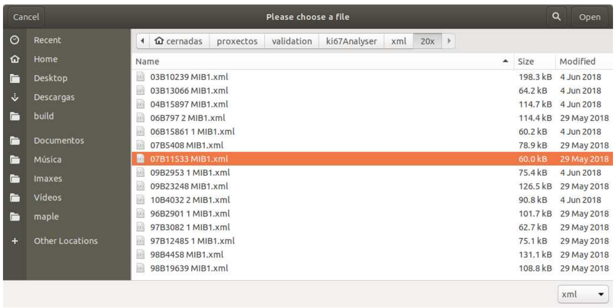
Min. Diameter: 10

Max. Diameter: 50

Region name:

Image: /home/cernadas/proxectos/validation/ki67Analyser/images/20x/07B11533 MIB1.tif Size: 4080x3072

Figure A.15: The image selected in Figure A.14 is loaded in OralImmunoAnalyser.



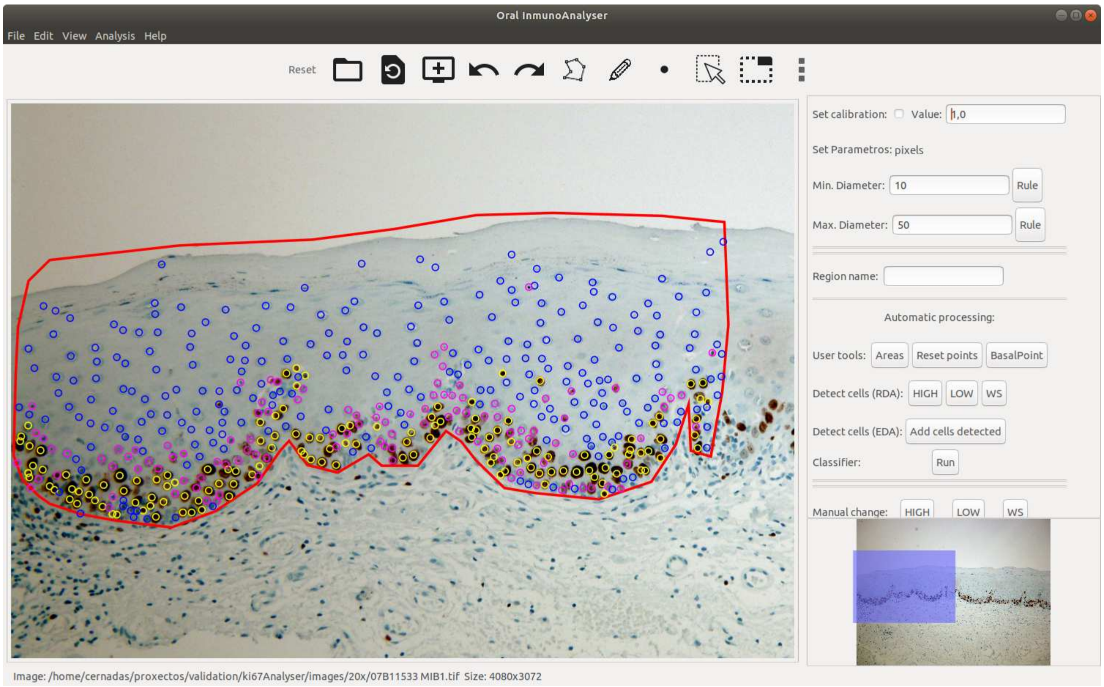
Name	Size	Modified
03B10239 MIB1.xml	198.3 kB	4 Jun 2018
03B13066 MIB1.xml	64.2 kB	4 Jun 2018
04B15897 MIB1.xml	114.7 kB	4 Jun 2018
06B797 2 MIB1.xml	114.4 kB	29 May 2018
06B15861 1 MIB1.xml	60.2 kB	4 Jun 2018
07B5408 MIB1.xml	78.9 kB	29 May 2018
07B11533 MIB1.xml	60.0 kB	29 May 2018
09B2953 1 MIB1.xml	75.4 kB	4 Jun 2018
09B23248 MIB1.xml	126.5 kB	29 May 2018
10B4032 2 MIB1.xml	90.8 kB	4 Jun 2018
96B2901 1 MIB1.xml	101.7 kB	29 May 2018
97B3082 1 MIB1.xml	62.7 kB	29 May 2018
97B12485 1 MIB1.xml	75.1 kB	29 May 2018
98B4458 MIB1.xml	131.1 kB	29 May 2018
98B19639 MIB1.xml	108.8 kB	29 May 2018

Figure A.16: File chooser dialog to choose the XML file to be loaded in OralImmunoAnalyser.

for each image, OralImmunoAnalyser does not check if this XML file corresponds to a specific image file.

The above process can be done in one step with the submenu **Open Image and XML**, which open a file chooser dialogue, as in Figure A.14, to choose the image to load and OralImmunoAnalyser checks in the XML path set in the preferences if there is a XML file with the same image name and extension .xml. In this case, OralImmunoAnalyser opens this XML file and overlaps its content over the image loaded (go directly to figure A.17).

Once the image was analysed (automatically or manually), the statistical analysis can be saved in the CSV file. Figure A.18 shows the basal, medial and superior regions (see section



Set calibration: ☐ Value: 1,0

Set Parametros: pixels

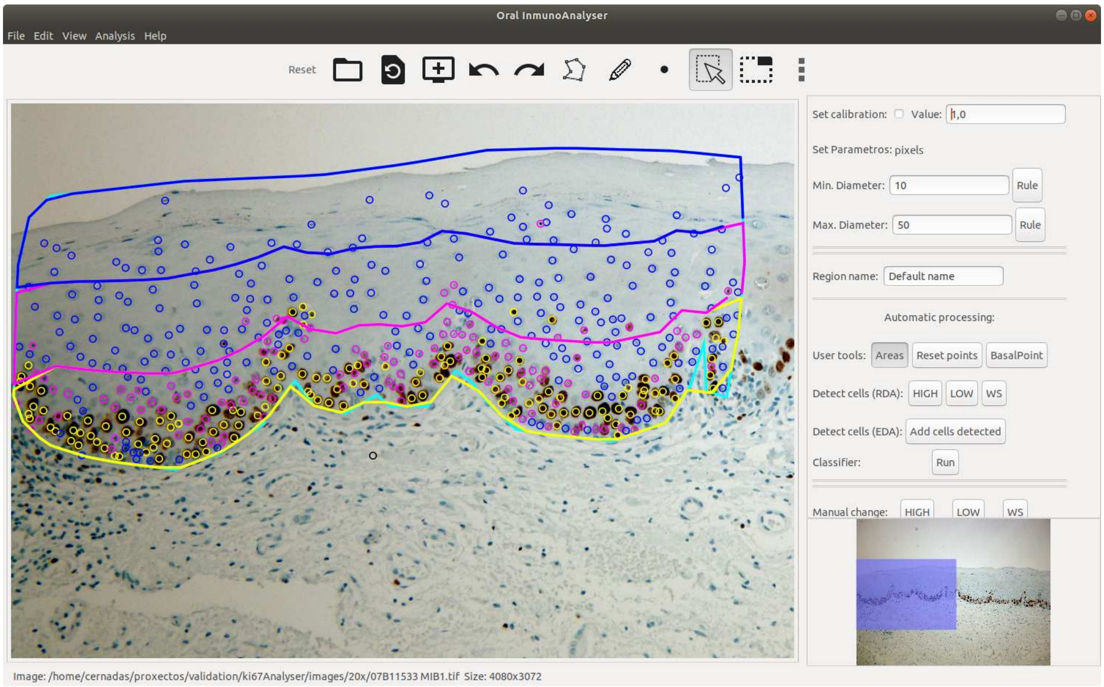
Min. Diameter: 10

Max. Diameter: 50

Region name:

Image: /home/cernadas/proyectos/validation/ki67Analyser/images/20x/07B11533 MIB1.tif Size: 4080x3072

Figure A.17: OralImmunoAnalyser shows the image of figure A.15 with the date of XML file selected in figure A.16 overlapped.



Set calibration: ☐ Value: 1,0

Set Parametros: pixels

Min. Diameter: 10

Max. Diameter: 50

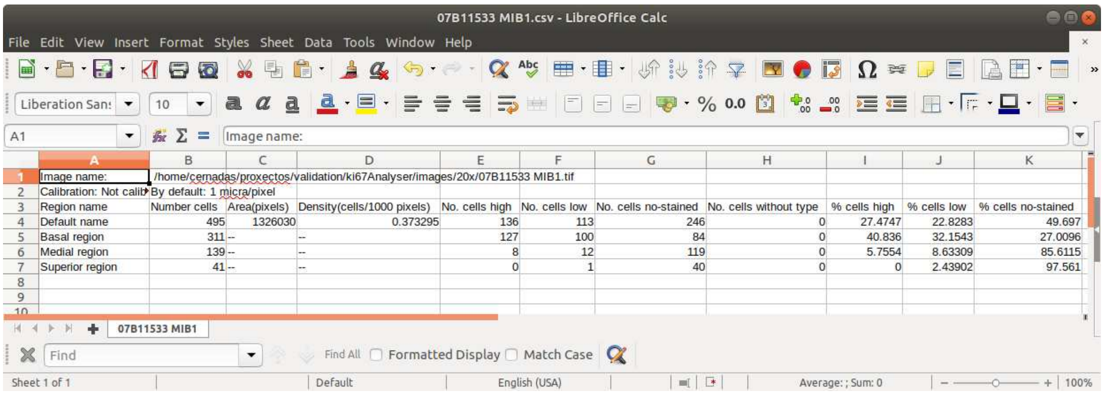
Region name: Default name

Image: /home/cernadas/proyectos/validation/ki67Analyser/images/20x/07B11533 MIB1.tif Size: 4080x3072

Figure A.18: OralImmunoAnalyser shows the image of figure A.17, in which the calculation of basal, medial and superior areas are visualized.

Name: 07B11533 MIB1.csv

Figure A.19: File chooser dialog to choose the CSV file to store the statistical analysis of the image.



07B11533 MIB1.csv - LibreOffice Calc

A1: Image name:

	A	B	C	D	E	F	G	H	I	J	K
1	Image name:	/home/cernadas/proxectos/validation/ki67Analyser/images/20x/07B11533 MIB1.tif									
2	Calibration: Not calib	By default: 1 micra/pixel									
3	Region name	Number cells	Area(pixels)	Density(cells/1000 pixels)	No. cells high	No. cells low	No. cells no-stained	No. cells without type	% cells high	% cells low	% cells no-stained
4	Default name	495	1326030	0.373295	136	113	246	0	27.4747	22.8283	49.697
5	Basal region	311	--	--	127	100	84	0	40.836	32.1543	27.0096
6	Medial region	139	--	--	8	12	119	0	5.7554	8.63309	85.6115
7	Superior region	41	--	--	0	1	40	0	0	2.43902	97.561
8											
9											

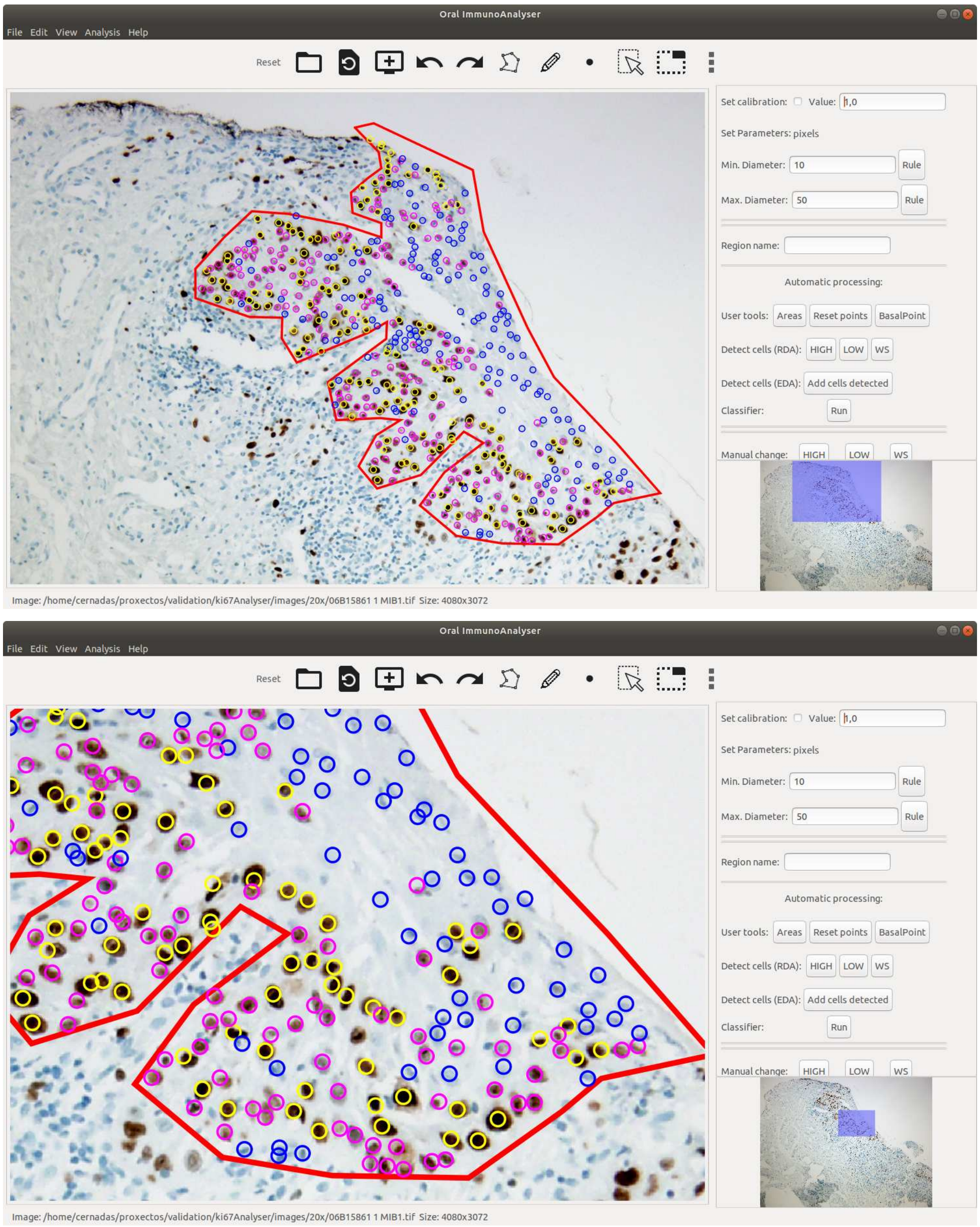
Figure A.20: An example of CSV file of the image of Figure A.18 imported in LibreOffice Calc.

A.9.2 to a further description), which are visualized if the toggle button **Areas** in the lateral panel was activated and the basal point was marked by the user. The submenu **Export CSV** opens the file chooser dialogue in the CSV directory selected in the preferences (see figure A.19) to select the name of CSV file (by default it is used the name of image with the extension CSV). OralImmunoAnalyser stores the statistical analysis of the image in that file. The information stored is: image path, calibration used, total number of cells, area of the region of analysis (in pixels if calibration is not set and in micrometers otherwise), density of cells in the region of analysis, number of cells high stained, number of cells low stained, number of cells without staining, number of cells that the user do not assign any category, percentage of high stained cells, percentage of low stained cells and percentage of cells without staining. If the basal point was marked by the user, the number of cells of each category and their percentages for the basal, medial and superior area are provided. The information stored in the CSV file can be loaded in a spreadsheet as LibreOffice Calc[2], as it can be seen in Figure A.20 (take care that the decimal numbers are represented by a point in the CSV file and not by comma).

The **Load preferences** submenu opens a file chooser dialogue to select the XML file preferences (preferences.xml file) to load in the computer.

---

[2] https://www.libreoffice.org/discover/calc/



Set calibration: ☐ Value: 1,0
Set Parameters: pixels
Min. Diameter: 10
Max. Diameter: 50
Region name:
Image: /home/cernadas/proxectos/validation/ki67Analyser/images/20x/06B15861 1 MIB1.tif Size: 4080x3072

Set calibration: ☐ Value: 1,0
Set Parameters: pixels
Min. Diameter: 10
Max. Diameter: 50
Region name:
Image: /home/cernadas/proxectos/validation/ki67Analyser/images/20x/06B15861 1 MIB1.tif Size: 4080x3072

Figure A.21: A zoom of the image in OralImmunoAnalyser.

A.7 Edit menu

The items available within this menu are also available in the toolbar (see section A.4). They are: **Undo** (fifth icon in the toolbar), **Redo** (sixth icon in the toolbar), **Fit Image** (third icon in the toolbar) and **Original Size** (fourth icon in the toolbar).

The **Fit Image** item fits the image zoom to the image window (for example see the figure A.9) and the **Original Size** sets the original image zoom. Another image zoom can be achieved rolling up the mouse wheel to increase the zoom and rolling down the mouse wheel to decrease the zoom. If the lateral panel is open, the position of the visible image area in the image window can be seen in the icon image located in the bottom of lateral panel (see Figure A.21). Keeping this zoom, you can move to another part of the image by two methods: 1) pressing simultaneously the left and right buttons of the mouse and move the mouse to displace the visible area; and 2) press the left mouse button on the blue square in the icon image of the lateral panel and move it. Both the visible area in the window image and the icon image of lateral panel are synchronized. As can be seen, the overlays of the image are zoomed with the image.

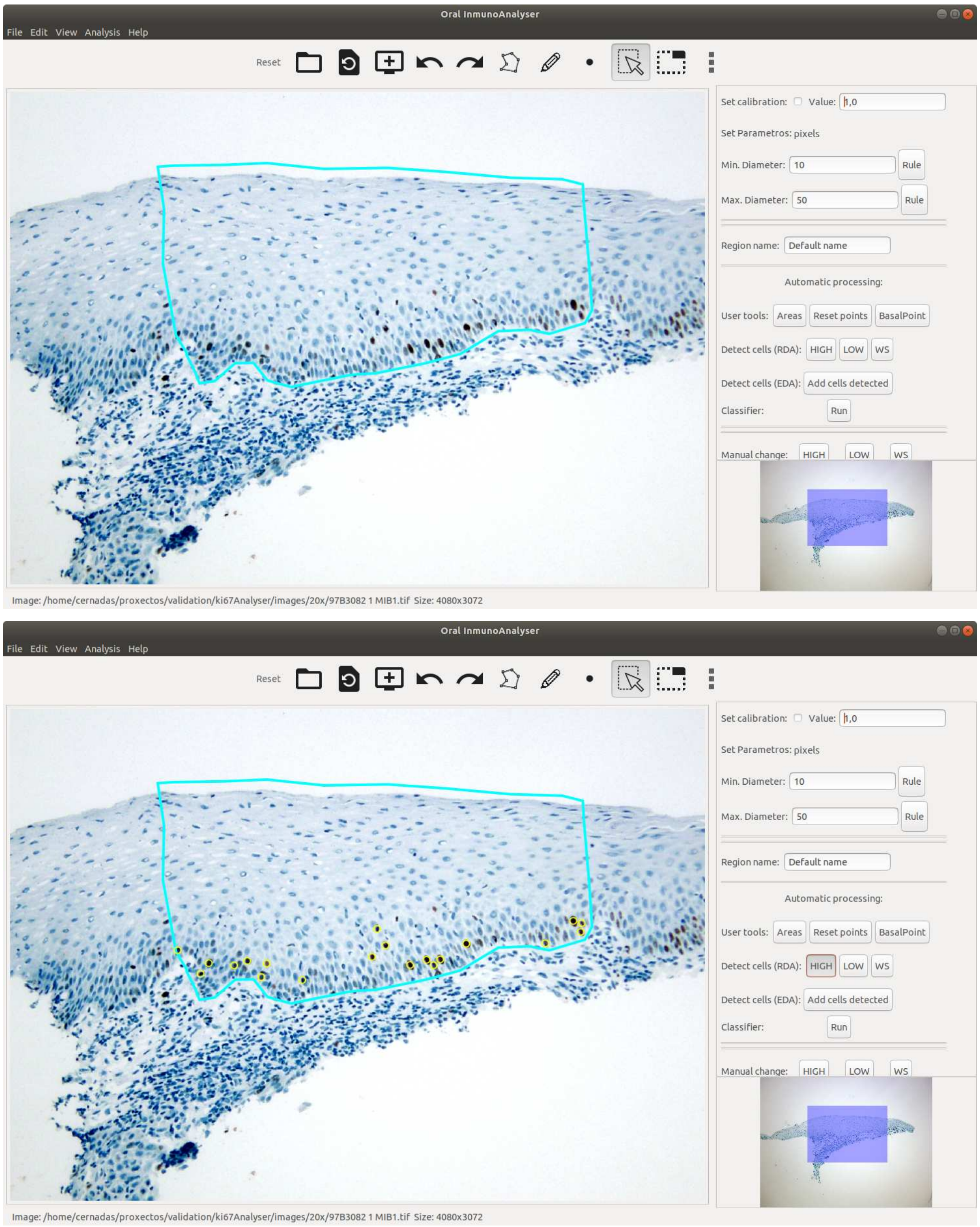
A.8 View menu or lateral panel

The only item available is **Processing Panel** (also included in the last icon of the toolbar), which closes or opens the lateral panel of figure A.9. The lateral panel contains the following functionality from top to bottom:

1. **Calibration** (first line): shows the calibration preferences set in the software or allows to change the calibration options as in the configuration preferences window (see section A.5).

2. **Diameters** (third and fourth line): shows the minimum and maximum diameters set in the preferences of the program or allows to set these values. The values of the minimum or maximum diameter can be set numerically, using the entry widgets (as it was described in section A.5), or graphically, using the button **Rule**. When the calibration or diameters are modified, all the overlays of the image are deleted. Using the entry widgets, the value of the diameter would be provided in pixels, if calibration is not activated, or in micrometers otherwise. To choose the diameter graphically, do the following steps: 1) click on the **Rule** button (the button remains activated); 2) draw a line on the image window pressing the left mouse button and when you move the mouse, with the left mouse button pressed, the line is drawing; and 3) when you release the left mouse button, the **Rule** button will become deactivated and the length of the line is put in the entry widget (this value does not appear if the minimum diameter is superior than maximum diameter or the maximum diameter is lower than the minimum one). The changes done in the lateral panel for calibration and diameters are only applied for the current working session. But, if you want to save permanently these changes you must go the preferences submenu (see section A.5).

3. **Region name** (fifth line): after this label widget, there is a entry widget in which the user can assign a name (text) to the selected region. The software assigns the name “Default name” by default.



Set calibration: ☐ Value: 1,0

Set Parametros: pixels

Min. Diameter: 10

Max. Diameter: 50

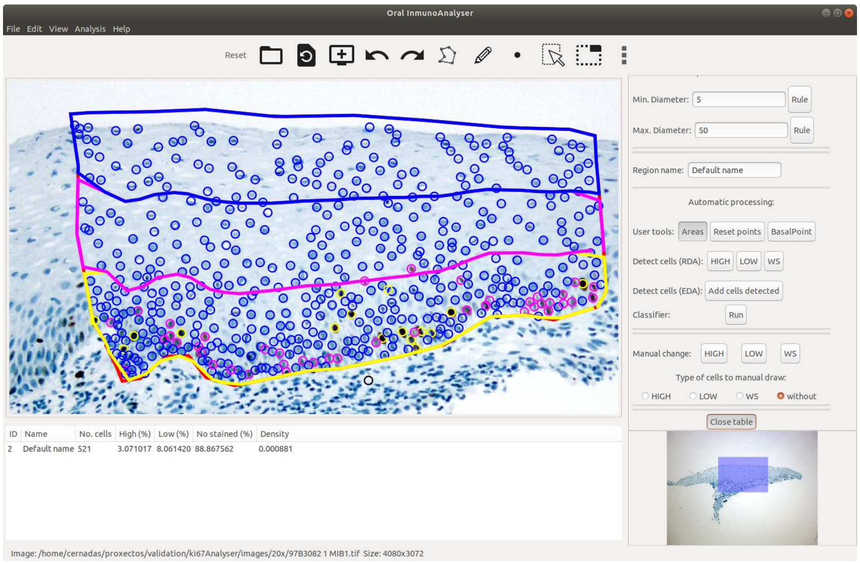
Region name: Default name

Image: /home/cernadas/proxectos/validation/ki67Analyser/images/20x/97B3082 1 MIB1.tif Size: 4080x3072

Figure A.22: After a region of analysis was drawn and selected (upper panel) and the upper region after the use the button **HIGH** of RDA method (lower panel).

4. **Automatic processing**: (from seven to ten line) after this label, there is a set of four lines to automatically process the image, which will be described further in section A.9.

5. **Category asignment**: (from eleven to thirteen line): there are a set of functionalities to manage the issues related to the assigment of categories to the cells detected (see also the A.9 section).



Min. Diameter: 5

Max. Diameter: 50

Region name: Default name

○ HIGH ○ LOW ○ WS ● without

ID	Name	No. cells	High (%)	Low (%)	No stained (%)	Density
2	Default name	521	3.071017	8.061420	88.867562	0.000881

Image: /home/cernadas/proxectos/validation/ki67Analyser/images/20x/97B3082 1 MIB1.tif Size: 4080x3072

Figure A.23: Visualization of information of the image analysis in the bottom table (after clicking the button **Show Table** in the lateral panel).

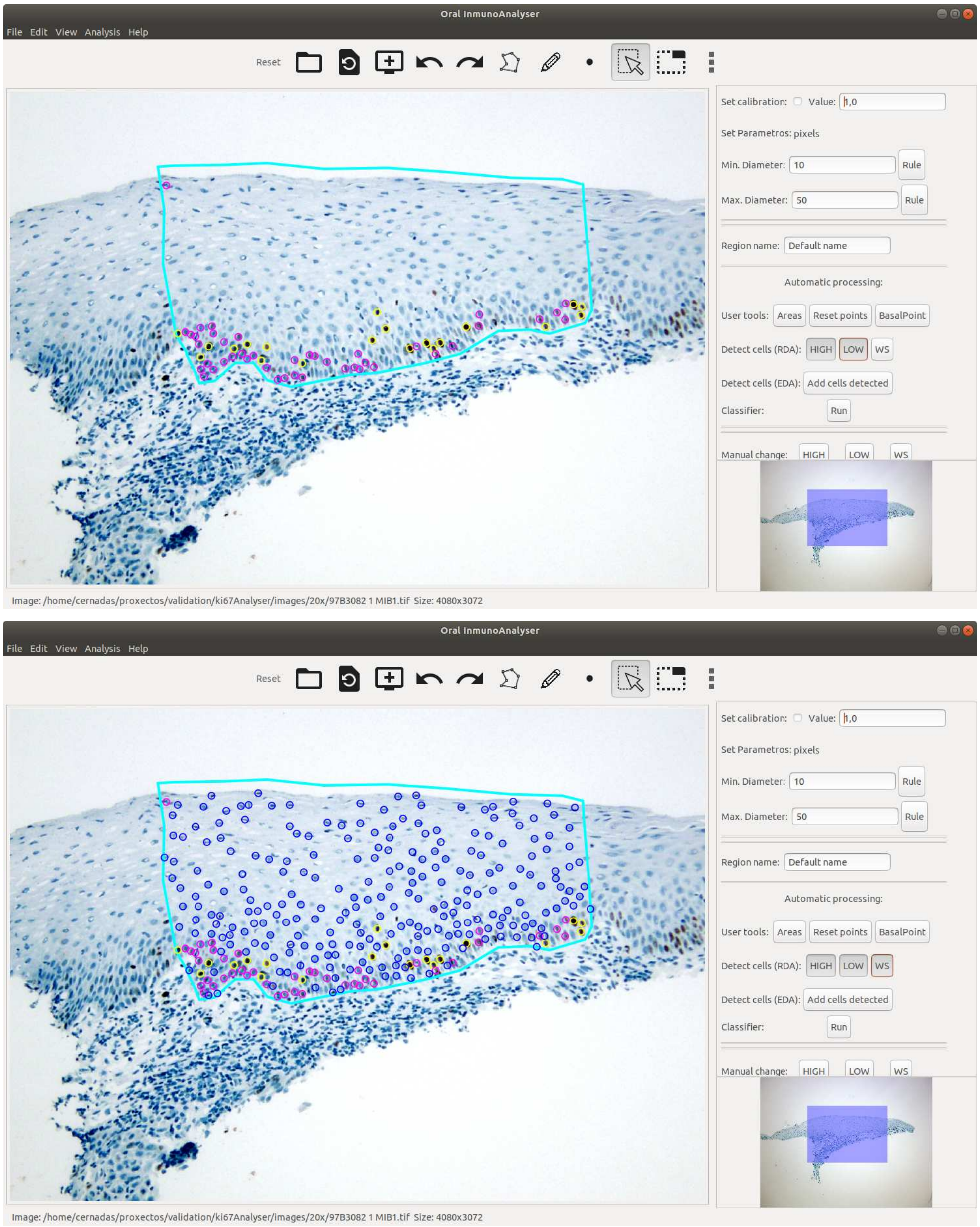
6. **Visualization of results**: when pressing the toggle button **Show Table**, a table containing analysis information is open in the bottom (more details in section A.9). If the table is open, you can click that button (now with the message **Close table**) to close the table.

7. **Visualization position**: at the bottom of the lateral panel, there is a miniature or icon image of the original image loaded in OralImmunoAnalyser. Over this miniature, there is overlapped a blue shadown square showing the part of the original image, which is shown in the image window. The position and size of this blue square depends on the zoom used in that moment, as can be see in figure A.21.

A.9 Automatic processing

To process automatically an immunohistochemistry image, the user must do the following steps, pointing out if that step is required or optional:

1. **Draw the region of analysis** (required): the user must manually draw the region of analysis. So, activate the freehand button in the toolbar (seven or eight button) and draw a region in the image window using the mouse. Finally, activate the button **Select** (tenth button in toolbar) and click inside the region with the left button of the mouse to select the region (see upper panel of figure A.22).



Set calibration: ☐ Value: 1,0
Set Parametros: pixels
Min. Diameter: 10
Max. Diameter: 50
Region name: Default name

Image: /home/cernadas/proxectos/validation/ki67Analyser/images/20x/97B3082 1 MIB1.tif Size: 4080x3072

Set calibration: ☐ Value: 1,0
Set Parametros: pixels
Min. Diameter: 10
Max. Diameter: 50
Region name: Default name

Image: /home/cernadas/proxectos/validation/ki67Analyser/images/20x/97B3082 1 MIB1.tif Size: 4080x3072

Figure A.24: Image of the upper panel of figure A.22 after the use the button **LOW** of RDA method (upper panel) and after the use of the three buttons of the RDA method (lower panel).

2. **Change the name to the region** (optional): by default the name of the region is “Default name”. The user can change this name writing a new text in the entry widget after the label *Region name* in the lateral panel.

3. **Detect cells inside the region** (required): detect automatically the cells inside the selected region using the buttons of lines eight and nine of lateral panel. A further description is in section A.9.1.

4. **Classify cells** (required): once the cells are automatically detected, they must be classified to assign a category to each cell. The categories included in OralImmunoAnalyser are: 1) cells high stained; 2) cells with low stained; and 3) cells without staining (further details in section A.9.1).

5. **Manual supervision** (optional): as the immunohistochemitry images are very complex and exist a high variance among images, the automatic processing is frequently incomplete or not perfect. Every time the user can use the edition tools provided by the GUI to delete and add cells. The user can also change the cell category using the buttons **HIGH**, **LOW** and **WS** after the label *Manual change* (line eleven in the lateral panel). To change the category of a cell or set of cells click in the suitable button taking in mind that the cells must be selected previously.

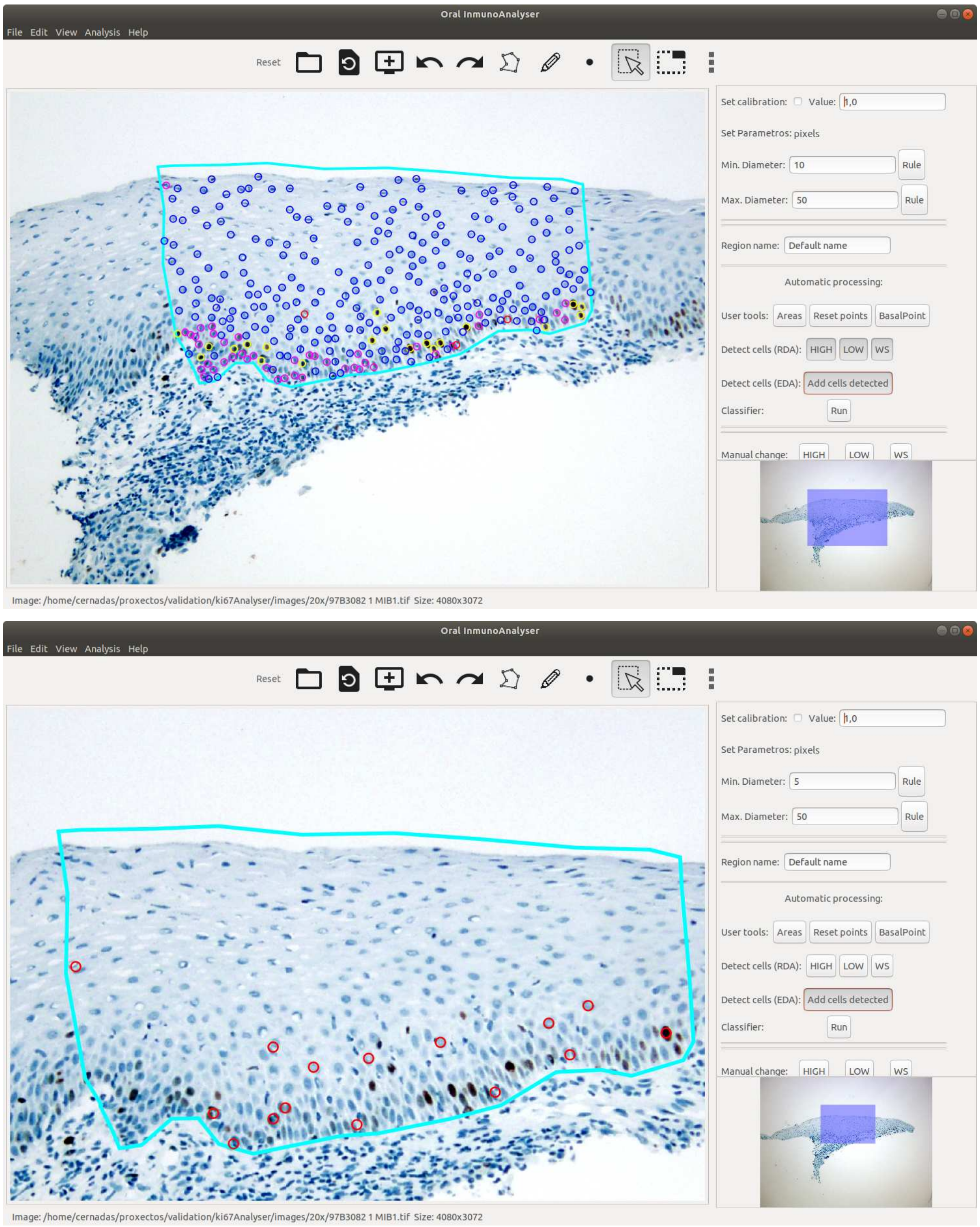
6. **Counts by subregions** (optional): OralImmunoAnalyser allows to count cells into the region of analysis by the following three subareas: basal, middle and superior (further details in section A.9.2).

7. **Visualization of results** (optional): pressing the button **Show Table** of the lateral panel open a table at the bottom of window (see Figure A.23) to show the following information: region name, number of cells in the region, percentage of cells high stained, percentage of cells low stained, percentage of cells without staining and the density of cells. The information shown in this table can be exported to a CSV file, as it has been described in the section A.6.

A.9.1 Detect and classify cells

OralImmunoAnalyser encloses two differents algorithms to dectect the cells in the image, based on two different paradigms: 1) the method called RDA in the lateral panel (line eight); and 2) the method called EDA (line nine in the lateral panel). The RDA method searchs differences in some property among pixels in the image, while the EDA method tries to find the discontinuities between an object (in our case cell) and its neighborhood in some property. So, as both algorithms are based on different principles, it is possible that one algorithm works better on one type of images and the another on another type of images. It is important to emphasize that both methods are synchronized, which implies that the software checks if there is some overlapping among the detected cells by both methods (see below for examples).

Once the region of analysis was drawn and selected (as in the upper panel of figure A.22), you can use the automatic detection of cells. Normally, the high stained cells are better detected by RDA algorithm (especially when the image is well-contrasted). The EDA methods detect better



Set calibration: ☐ Value: 1,0
Set Parametros: pixels
Min. Diameter: 10
Max. Diameter: 50
Region name: Default name
Image: /home/cernadas/proxectos/validation/ki67Analyser/images/20x/97B3082 1 MIB1.tif Size: 4080x3072

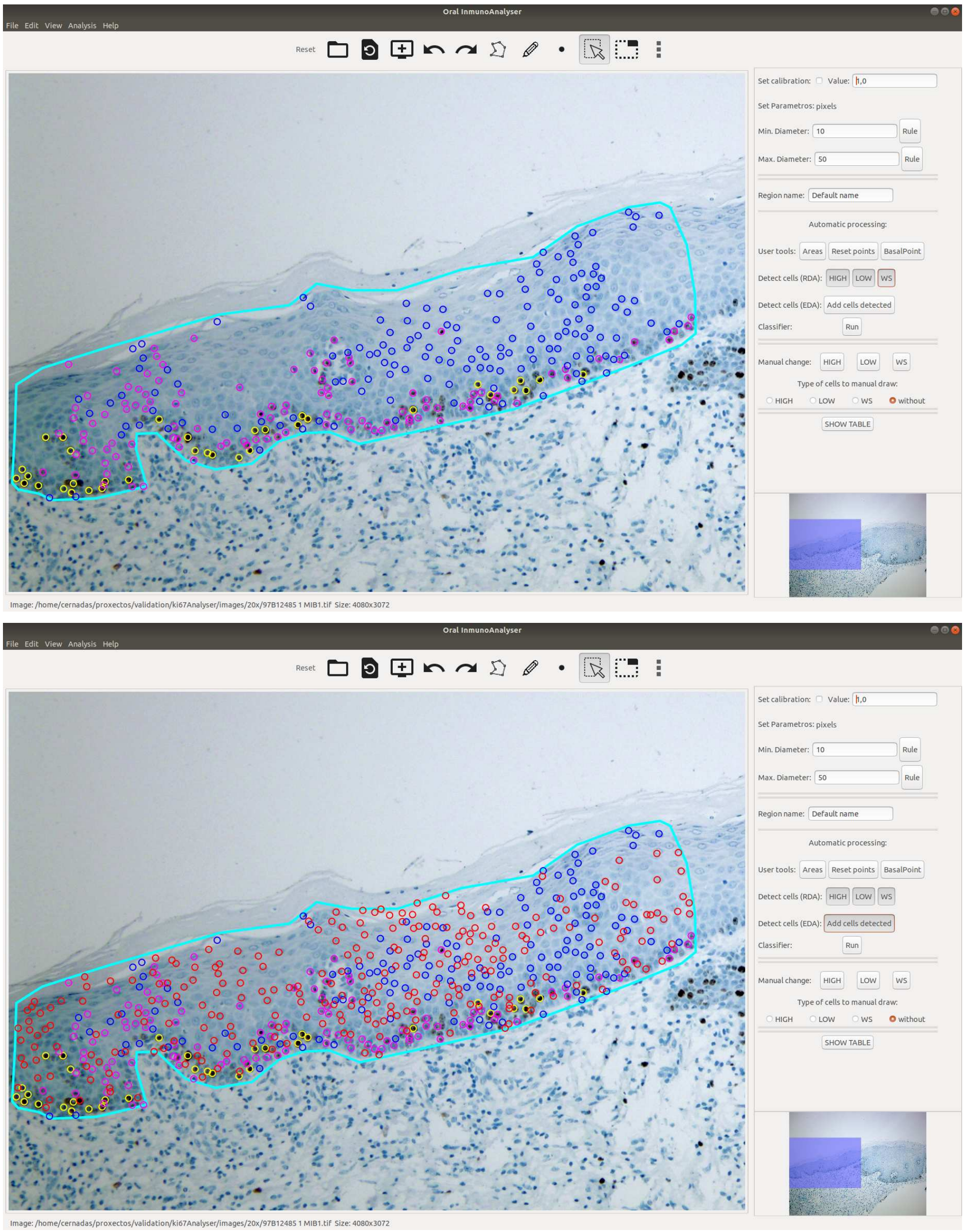
Set calibration: ☐ Value: 1,0
Set Parametros: pixels
Min. Diameter: 5
Max. Diameter: 50
Region name: Default name
Image: /home/cernadas/proxectos/validation/ki67Analyser/images/20x/97B3082 1 MIB1.tif Size: 4080x3072

Figure A.25: Image of the lower panel of figure A.24 after adding the cells of EDA method (upper panel) and cells detected if you choose to apply the EDA method first (lower panel).

the non-stained cells (specially when image contrast is low), but it is rather sensitive to the image noise and, sometimes, provides an unwanted number of false positive. The more important is that both algorithms are quite fast and you can undo every automatic processing. So, the use of the four buttons of lines eight and nine in the lateral panel (buttons **HIGH**, **LOW**, **WS** and **Add detected cells**) can be used in any combination. All these buttons are toggle buttons, which mean that if the button is active, the detected cells are visualize on the image window and otherwise the detected cells are hidden (and you must active to visualize the cells again). The order in which the buttons are activated is also important due to the overlapping test, which is applied every time you modified the state of some button. The buttons **HIGH**, **LOW** and **WS** do not mean that the detected cells after clicking these buttons are the high stained, low stained and non-stained cells respectively (although the button **HIGH** detect the most stained cell, the **LOW** button a less stained cells and the **WS** the much less stained cells, but there are not absolute staining and they are only relative staining which depends on the image characteristics. Our purpose was to simulate the idea of the scale and to use different colours to draw the cells detected by each button in order to facilitate or visualize explicitly the cells detected clicking in the different buttons.

Some examples ilustrating the automatic detection of cells are shown in figures A.22, A.24 and A.25. The lower panel of figure A.22 shows the cells detected after clicking in the button **HIGH** (notice that the button is activated and the cells are shown as circles in yellow, which is the colour for category "high stained" in my working preferences). If you click again in the button **HIGH**, the button will deactivate and the detected cells by this button will be hidden. The figure A.24 shows the cells added by clicking the button **LOW** in pink colour (upper panel) and the cells added by clicking the button **WS** in blue colour (lower panel). In both cases, the button is activated after clicking it. Although some cells can be detected cliking different buttons, notice that the cells never appear overlapped. The upper panel of figure A.25 shows the detected cells added after clicking the button **Add detected cells** (method EDA) is red colour, which is the colour set to draw points in the working preferences. In this case, only one cell is added. If you have firstly applied this algorithm, the cells detected will be those shown in the lower panel of figure A.25, but in the upper panel of figure A.25, many of them were detected by RDA method before and then they are not added by the overlapping test. So, the order in which the four buttons are activated may provide a different set of cells automatically detected.

Figure A.26 shows another example of operation of cell detection methods. In this case, the RDA method does not detect all cells, which can be detected by EDA method. After the cell detection process is concluded, the cells must be classified. For that, click the button **Run** after the label *Classifier* (see the section A.11 for more details). Figure A.27 shows the result of classification applied to the upper panel of figure A.25. The colours used were set in the working preferences (see section A.5). In our preferences, the high stained cells are set to yellow, low stained cells are shown in pink and cells without staining are in blue colour. If the assignment of staining level is not suitable for the expert, OralImmunoAnalyser allows to change the category of the cells. To do that, the user must select the cells in two ways: 1) activate the select button in the toolbar and click near the point with the left button of the mouse (to select more than one cell, keep pressed the "Crtl" or "Control" key) and 2) use the penultimate button of toolbar to select the cells under the rectangle. Once the cell or cells are selected, click in some of the buttons after



Set calibration: ○ Value: 1,0

Set Parametros: pixels

Min. Diameter: 10

Max. Diameter: 50

Region name: Default name

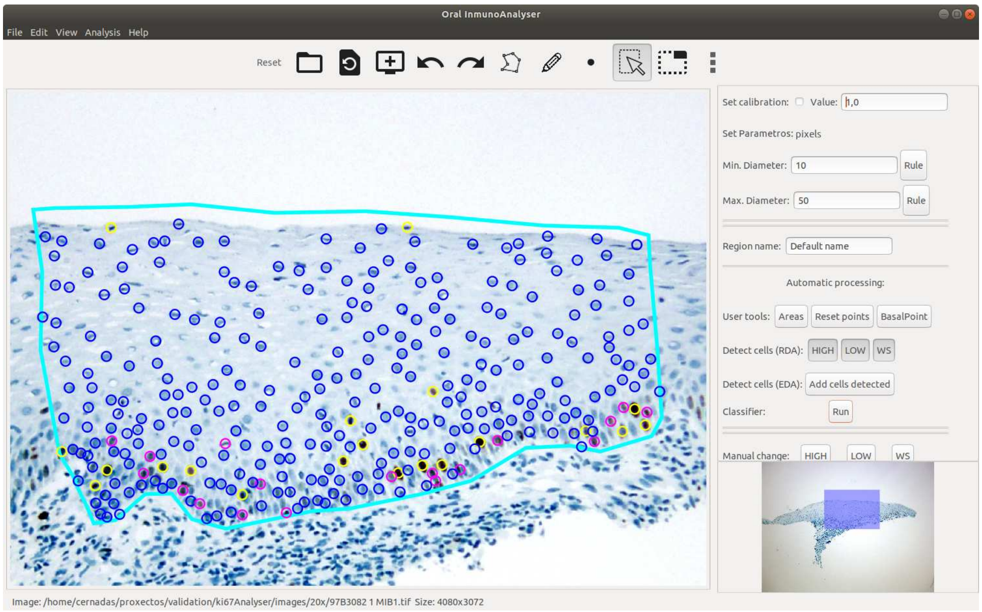
Type of cells to manual draw: ○ HIGH ○ LOW ○ WS ● without

Image: /home/cernadas/proxectos/validation/ki67Analyser/images/20x/97B12485 1 MIB1.tif Size: 4080x3072

Figure A.26: Example of the application of automatic detection of cells: (upper panel and points in yellow, pink and blue) after using RDA method and (lower panel and points in red) after adding the detected cells provided by EDA method.

the label *Manual change* to change the category to the cells. Once the cells are selected, they can be removed clicking the botton “Supr”.

The faster way to add new cells of each category is following the next steps: 1) choose the category of cells to add clicking one of the radio buttons after the label *Type of cells for manual draw*; 2) activate the button to draw points (ninth button) in the toolbar; and 3) click points in



Set calibration: ☐ Value: 1,0

Set Parametros: pixels

Min. Diameter: 10

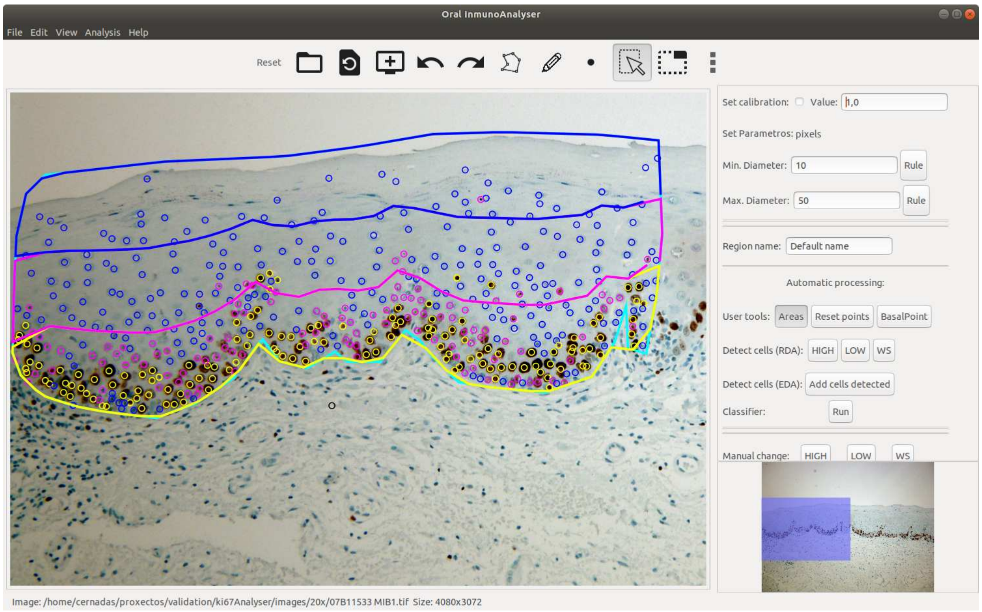
Max. Diameter: 50

Region name: Default name

Image: /home/cernadas/proxectos/validation/ki67Analyser/images/20x/97B3082 1 MIB1.tif Size: 4080x3072

Figure A.27: Image of the upper panel of figure A.25 after using the automatic classification of cells, i. e. after clicking the button **Run** after the label *Classifier*.

the region of analysis (the points added will visualize with the colour of the category chosen). It is recommended to choose the radio button **without** when you finish to add cells.



Set calibration: ☐ Value: 1,0

Set Parametros: pixels

Min. Diameter: 10

Max. Diameter: 50

Region name: Default name

Image: /home/cernadas/proxectos/validation/ki67Analyser/images/20x/07B11533 MIB1.tif Size: 4080x3072

Figure A.28: Show an example of the subareas basal, medial and superior calculated automatically in the region of analysis. The black point outside the region of analysis indicate to the computer the position of the basal subarea.

A.9.2 Counts by subareas

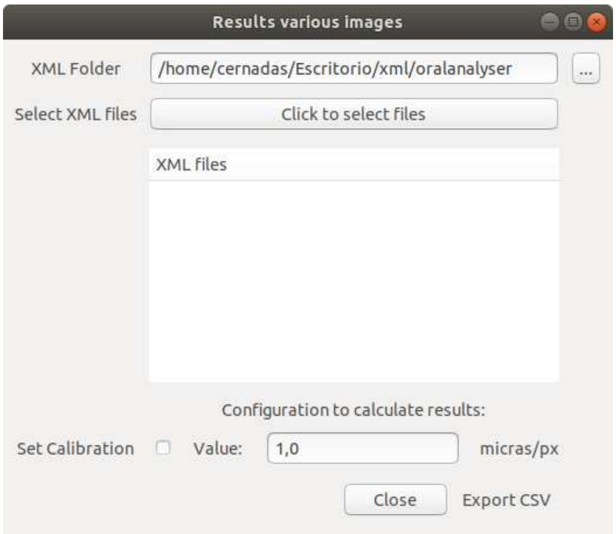
OralImmunoAnalyser allows to count the cells of each level of staining in the region of analysis, but also count the cells in the following subareas: basal, medial and superior. Each subarea covers approximatly the 33% of width of the region going from the basal to the superior part of the region. The subareas are automatically calculated by the computer following the next steps: 1) the user must provide to the computer the side of the region of analysis close to the basal area, marking a point in the image window outside the region of analysis; and 2) click the buton **Areas** after the label *User tools* to visualize the subareas calculated by the computer on the image window (the software used the categories colours to visualize each subarea). Although the subareas are visualized on the image window, their count is not shown in figure A.23 due to computational issues, but this information is exported to CSV files. Figure A.28 shows the subareas calculated by the computer (after marking the basal point) and figure A.20 the count analysis exported to the CSV file. As mentioned, to indicate to the computer which side of the region of analysis is the basal, the user should do the following steps: 1) activate the button **BasalPoint** after the label *User tools* in the lateral panel; 2) while this button is active click a point with the left button of the mouse outside the region of analysis marking the basal side of the region (see the black circle in figure A.28), when the point is marked the button **BasalPoint** will appear deactivated; and 3) click the button **Areas** after the label *User tools* to visualize the subareas calculated by the computer. The algorithm to calculate the subareas is not exhaustive and it could not work suitable if the region of analysis has folds or narrow valleys (as in the example shown in figure A.21).

The button **Reset points** after the label *User tools* is used to clear all the cells into the region of analysis. To do that, you must select the region of analysis and then click this button to remove the cells in the selected region.

A.10 Analysis menu

The only item available is **XML File**, which opens the pop up window of figure A.29. Many times the researchers want to acumulate the quantitative results of various images together. For this purpose, you must do the following steps: 1) process individually each image; 2) save the recognition results (region of analysis, basal point and the category of each cell in the region) in a XML file; and 3) run the submenu **XML File** of menu **Analysis** to calculate the jointly quantitative results.

The first line of window of figure A.29, labelled as *XML Folder*, allows to choose the directory of the XML files (by default, it is the XML directory set in the preferences, but it can be

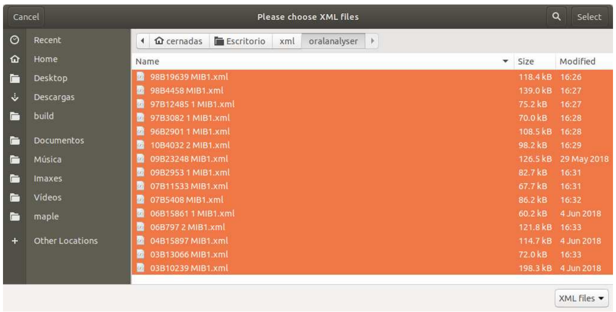


XML Folder /home/cernadas/Escritorio/xml/oralanalyser

Select XML files

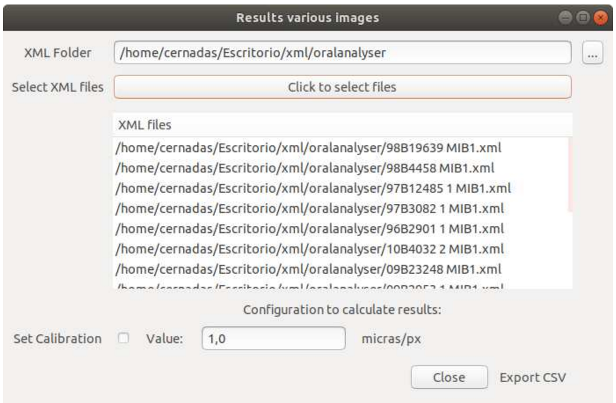
Set Calibration ☐ Value: 1,0 micras/px

Figure A.29: Window open when the submenu **XML File** of menu **Analysis** is chosen.



Name	Size	Modified
98B19639 MIB1.xml	118.4 kB	16:26
98B4458 MIB1.xml	139.0 kB	16:27
97B12485 1 MIB1.xml	75.2 kB	16:27
97B3082 1 MIB1.xml	70.0 kB	16:28
96B2901 1 MIB1.xml	108.5 kB	16:28
10B4032 2 MIB1.xml	98.2 kB	16:29
09B23248 MIB1.xml	126.5 kB	29 May 2018
09B2953 1 MIB1.xml	82.7 kB	16:31
07B11533 MIB1.xml	67.7 kB	16:31
07B5408 MIB1.xml	86.2 kB	16:32
06B15861 1 MIB1.xml	60.2 kB	4 Jun 2018
06B797 2 MIB1.xml	121.8 kB	16:33
04B15897 MIB1.xml	114.7 kB	4 Jun 2018
03B13066 MIB1.xml	72.0 kB	16:33
03B10239 MIB1.xml	198.3 kB	4 Jun 2018

Figure A.30: File chooser dialogue to choose the XML files included in the jointly analysis.



XML Folder /home/cernadas/Escritorio/xml/oralanalyser

Select XML files

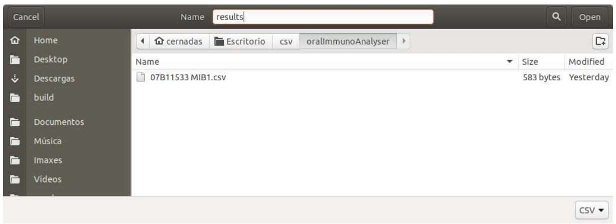
XML files

/home/cernadas/Escritorio/xml/oralanalyser/98B19639 MIB1.xml
/home/cernadas/Escritorio/xml/oralanalyser/98B4458 MIB1.xml
/home/cernadas/Escritorio/xml/oralanalyser/97B12485 1 MIB1.xml
/home/cernadas/Escritorio/xml/oralanalyser/97B3082 1 MIB1.xml
/home/cernadas/Escritorio/xml/oralanalyser/96B2901 1 MIB1.xml
/home/cernadas/Escritorio/xml/oralanalyser/10B4032 2 MIB1.xml
/home/cernadas/Escritorio/xml/oralanalyser/09B23248 MIB1.xml

Set Calibration ☐ Value: 1,0 micras/px

Figure A.31: Window open when the submenu **XML File** of menu **Analysis** is chosen, after selecting the XML files used in the analysis.

changed clicking the three points button). The second line, labelled as *Select XML File*, have the button **Click to select files**. Clicking this button, the file chooser dialogue of figure A.30 is open



Name: results

Name	Size	Modified
07B11533 MIB1.csv	583 bytes	Yesterday

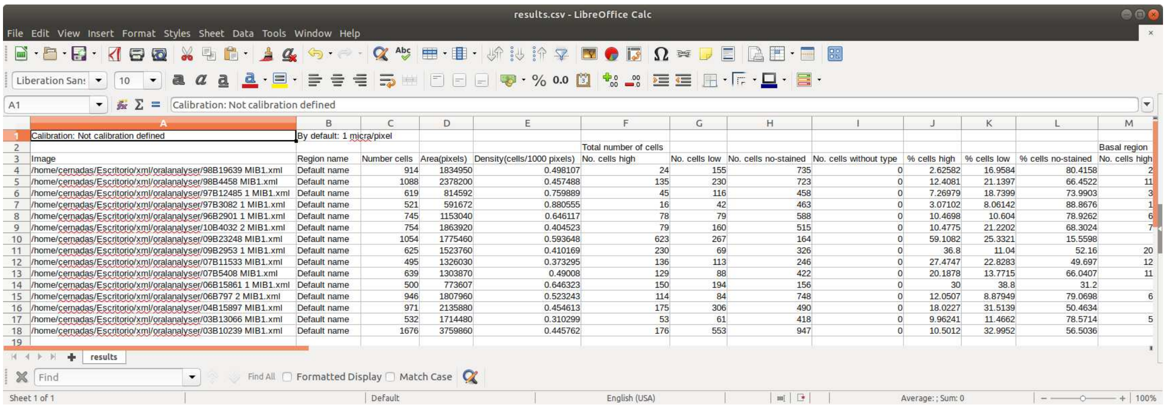
Figure A.32: File chooser dialogue to provide the name of the CSV file.

to choose the XML files used to compute the statistics results. As all file chooser dialogues, a file is selected if you click on the file name with the left button mouse. To choose consecutive files, you click the first file and, keeping the key **Alt** pressed, you click the last one. To choose various files, you click files names keeping the key **Ctrl** pressed. Once the XML files were chosen, click the **Select** button in the bottom of the window. The files selected will be appeared in the Figure A.29, as it can be seen in the Figure A.31. Then, the user can choose the following operations: 1) cancel this operation pressing the **Close** button in the bottom; or 2) press the **Export CSV** to store the joined statistical results in a CSV file. This last operation open a file chooser dialogue, which is shown in figure A.32, providing the name results for the CSV file. This saving process of the results.csv file could require some seconds if there are many XML files to analyse. After saving one CSV file, the window of Figure A.29 goes on open until the user click the **Close** button, in order to do other analysis. One example of the CSV file stored can be observed in the Figure A.33. The information saved for each image are: 1) name of image; 2) name of the region of analysis and its number of cells; 3) area of region and density of cells (if the calibration is set it is appeared in micrometers and otherwise in pixels); 3) the global count of the image and the percentage of cells for each staining level; and 4) this count information for each subarea in the region. As it can be seen in figure A.33, some rows are empty for the counts of the subareas (for example file 10B4032 2 MIB1.xml in row 10). This means that the user has not provided the basal point and then has saved the overlays in the XML file. In this case, the subareas can not calculate the counts by subareas.

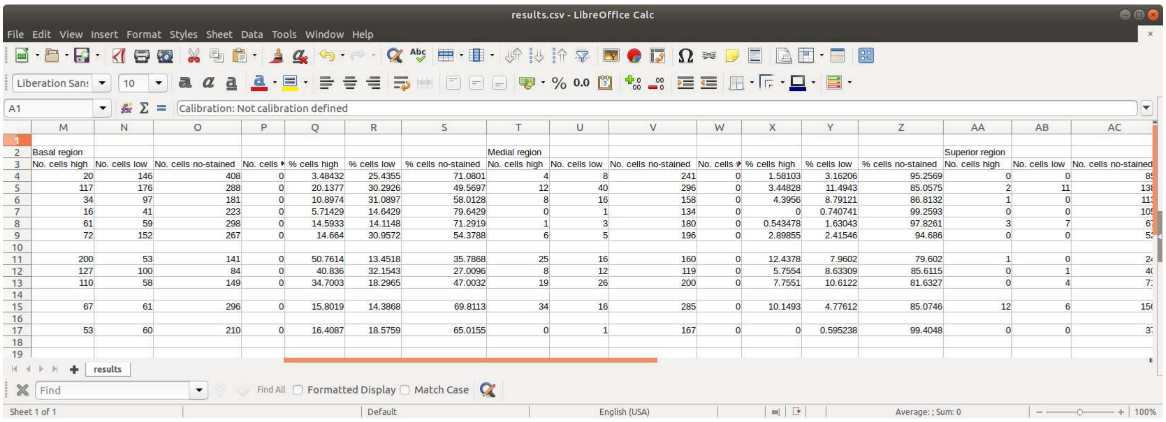
A.11 Classification menu

As mentioned, OralImmunoAnalyser checks if the classifier was loaded when the user run OIA. If there is a classifier trained in the computer, the pop-up message of upper panel of figure A.7 is shown, otherwise, it is shown the pop-up message of lower panel of figure A.7 informing that the classifier must be trained before classifing cells.

The submenus of the Classification menu from the menu bar are:



	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Calibration: Not calibration defined	By default: 1 micra/pixel											
2						Total number of cells							Basal region
3	Image	Region name	Number cells	Area(pixels)	Density(cells/1000 pixels)	No. cells high	No. cells low	No. cells no-stained	No. cells without type	% cells high	% cells low	% cells no-stained	No. cells high
4	/home/cernadas/Escritorio/xml/oralanalyser/98B19639 MIB1.xml	Default name	914	1834950	0.498107	24	155	735	0	2.62582	16.9584	80.4158	2
5	/home/cernadas/Escritorio/xml/oralanalyser/98B4458 MIB1.xml	Default name	1088	2378200	0.457488	135	230	723	0	12.4081	21.1397	66.4522	11
6	/home/cernadas/Escritorio/xml/oralanalyser/97B12485 1 MIB1.xml	Default name	619	814592	0.759889	45	116	458	0	7.26979	18.7399	73.9903	3
7	/home/cernadas/Escritorio/xml/oralanalyser/97B3082 1 MIB1.xml	Default name	521	591672	0.880555	16	42	463	0	3.07102	8.06142	88.8676	1
8	/home/cernadas/Escritorio/xml/oralanalyser/96B2901 1 MIB1.xml	Default name	745	1153040	0.646117	78	79	588	0	10.4698	10.604	78.9262	6
9	/home/cernadas/Escritorio/xml/oralanalyser/10B4032 2 MIB1.xml	Default name	754	1863920	0.404523	79	160	515	0	10.4775	21.2202	68.3024	7
10	/home/cernadas/Escritorio/xml/oralanalyser/09B23248 MIB1.xml	Default name	1054	1775460	0.593648	623	267	164	0	59.1082	25.3321	15.5598	
11	/home/cernadas/Escritorio/xml/oralanalyser/09B2953 1 MIB1.xml	Default name	625	1523760	0.410169	230	69	326	0	36.8	11.04	52.16	20
12	/home/cernadas/Escritorio/xml/oralanalyser/07B11533 MIB1.xml	Default name	495	1326030	0.373295	136	113	246	0	27.4747	22.8283	49.697	12
13	/home/cernadas/Escritorio/xml/oralanalyser/07B5408 MIB1.xml	Default name	639	1303870	0.49008	129	88	422	0	20.1878	13.7715	66.0407	11
14	/home/cernadas/Escritorio/xml/oralanalyser/06B15861 1 MIB1.xml	Default name	500	773607	0.646323	150	194	156	0	30	38.8	31.2	
15	/home/cernadas/Escritorio/xml/oralanalyser/06B797 2 MIB1.xml	Default name	946	1807960	0.523243	114	84	748	0	12.0507	8.87949	79.0698	6
16	/home/cernadas/Escritorio/xml/oralanalyser/04B15897 MIB1.xml	Default name	971	2135880	0.454613	175	306	490	0	18.0227	31.5139	50.4634	
17	/home/cernadas/Escritorio/xml/oralanalyser/03B13066 MIB1.xml	Default name	532	1714480	0.310299	53	61	418	0	9.96241	11.4662	78.5714	5
18	/home/cernadas/Escritorio/xml/oralanalyser/03B10239 MIB1.xml	Default name	1676	3759860	0.445762	176	553	947	0	10.5012	32.9952	56.5036	
19													



	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC
1																	
2	Basal region							Medial region							Superior region		
3	No. cells high	No. cells low	No. cells no-stained	No. cells	% cells high	% cells low	% cells no-stained	No. cells high	No. cells low	No. cells no-stained	No. cells	% cells high	% cells low	% cells no-stained	No. cells high	No. cells low	No. cells no-stained
4	20	146	408	0	3.48432	25.4355	71.0801	4	8	241	0	1.58103	3.16206	95.2569	0	0	8
5	117	176	288	0	20.1377	30.2926	49.5697	12	40	296	0	3.44828	11.4943	85.0575	2	11	13
6	34	97	181	0	10.8974	31.0897	58.0128	8	16	158	0	4.3956	8.79121	86.8132	1	0	11
7	16	41	223	0	5.71429	14.6429	79.6429	0	1	134	0	0	0.740741	99.2593	0	0	10
8	61	59	298	0	14.5933	14.1148	71.2919	1	3	180	0	0.543478	1.63043	97.8261	3	7	6
9	72	152	267	0	14.664	30.9572	54.3788	6	5	196	0	2.89855	2.41546	94.686	0	0	5
10																	
11	200	53	141	0	50.7614	13.4518	35.7868	25	16	160	0	12.4378	7.9602	79.602	1	0	2
12	127	100	84	0	40.836	32.1543	27.0096	8	12	119	0	5.7554	8.63309	85.6115	0	1	4
13	110	58	149	0	34.7003	18.2965	47.0032	19	26	200	0	7.7551	10.6122	81.6327	0	4	7
14																	
15	67	61	296	0	15.8019	14.3868	69.8113	34	16	285	0	10.1493	4.77612	85.0746	12	6	15
16																	
17	53	60	210	0	16.4087	18.5759	65.0155	0	1	167	0	0	0.595238	99.4048	0	0	3
18																	
19																	

Figure A.33: An example of CSV file of the analysis of various XML files imported in LibreOffice Calc. The lower screenshot is the continuation of the upper screenshot.

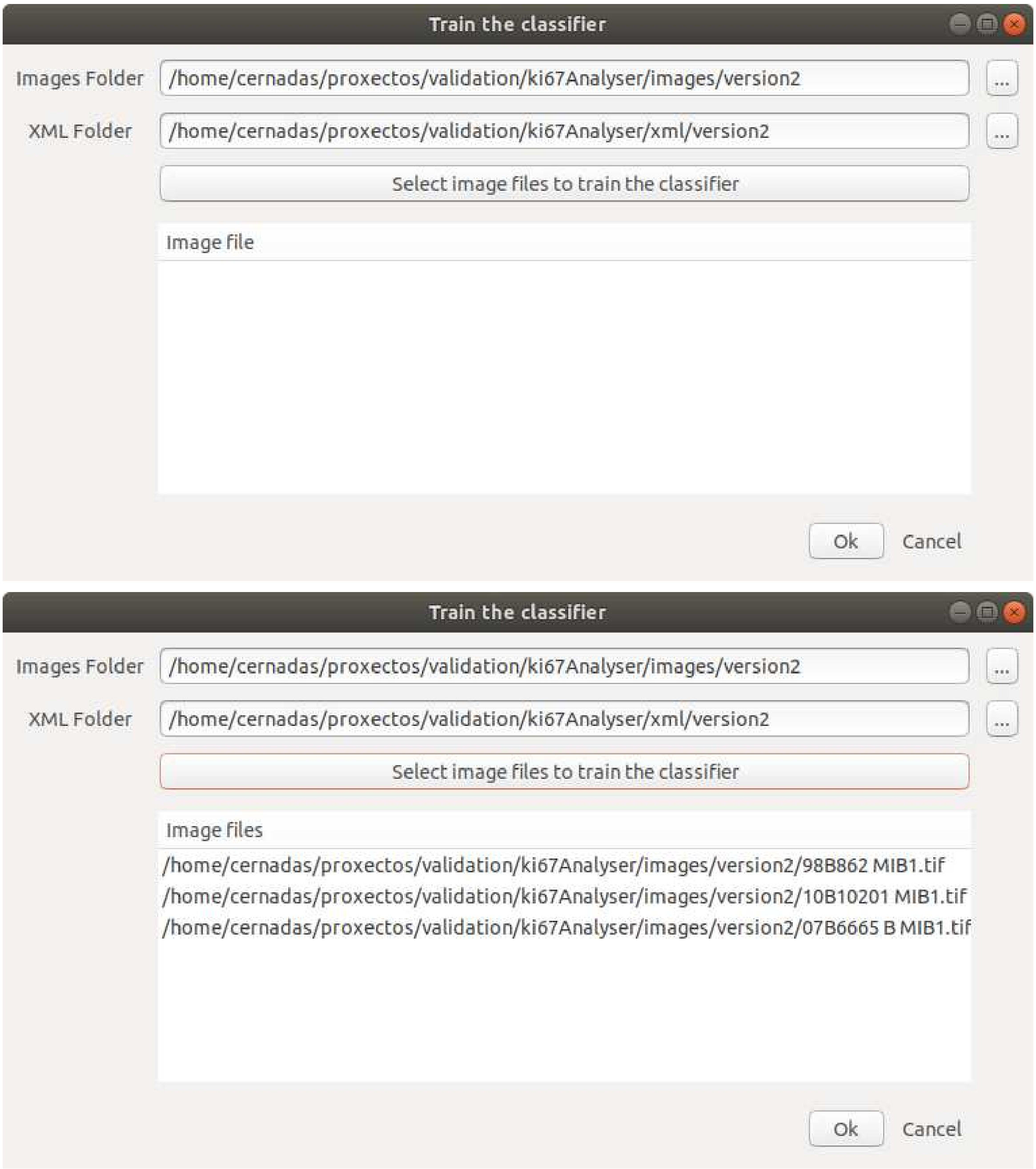

Information

There is no classifier loaded. Train first the classifier

Figure A.34: Pop-up window to inform that there is no classifier loaded.

1. **Classify**: assign a category (high, low or without stained) to each cell. This functionality can also do pressing the button **Run** after the label *Classifier* in the lateral panel (see the section A.9.1). If there is no classifier loaded in OIA software, the pop-up message of figure A.34 is shown. So, it is necessary to train the classifier before using it.
2. **Train classifier**: train the classifier, which is explaining below.

The establishment of the stained level of the cell (into category high stained, low stained and without staining) can be manually set in every time using the GUI (as it was described in section A.9.1). But, OralImmunoAnalyser implement the automatic classification of cells.



Images Folder: /home/cernadas/proxectos/validation/ki67Analyser/images/version2

XML Folder: /home/cernadas/proxectos/validation/ki67Analyser/xml/version2

Image file

Train the classifier

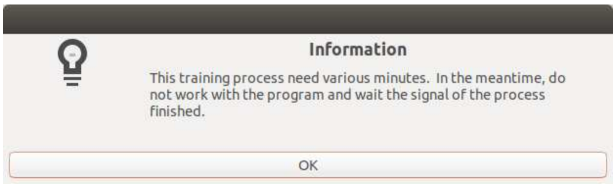
Images Folder: /home/cernadas/proxectos/validation/ki67Analyser/images/version2

XML Folder: /home/cernadas/proxectos/validation/ki67Analyser/xml/version2

Image files
/home/cernadas/proxectos/validation/ki67Analyser/images/version2/98B862 MIB1.tif
/home/cernadas/proxectos/validation/ki67Analyser/images/version2/10B10201 MIB1.tif
/home/cernadas/proxectos/validation/ki67Analyser/images/version2/07B6665 B MIB1.tif

Figure A.35: Window open to train the classifier: before (upper panel) selection the image files and after (lower) selecting the image files to train the classifier.

Before applying the classification operation (using the submenu **Classify** or click the button **Run** in the lateral panel after the label “Classifier”), the classifier need to be trained. The classifier is a supervised machine learning tecnique, which need to learn a function to predict the category of the cell, which have not been seen before by the classifier. To learn this function it is necessary to provide to the classifier a set of cells for which the category is known in a process called training.


Information

This training process need various minutes. In the meantime, do not work with the program and wait the signal of the process finished.

Figure A.36: Pop-up window to inform that the training process need various minutes to finish.

The image characteristics used to discriminate different categories of cells are color features extracted from the image in the position of the cells. So, you must manually analyse a limited number of cells and save the annotated cells overlapped to the image in their corresponding XML files (using the default name, which is the image name with the extension `.xml`) in order to train the classifier. Afterwards, to train the classifier do the following steps:

1. Select the submenu **Train classifier** in menu **Classification**, which opens the window of the upper panel of figure A.35. The first line allows to select the image folder, which is, by default, the image folder specified in the working preferences (see section A.5). The second line allows to select the XML folder, which is also the folder specified in the preferences by default. As we mentioned in section A.5 (see the figure A.10), clicking in the button with three points, you can change the image and XML folders.

2. Click the button **Select images files to train the classifier**, which opens a file choose dialogue as in the figure A.30 to select the image files used to train the classifier. The lower panel of figure A.35 shows the figure of upper panel after the selection of the image files.

3. Click the button **Cancel** to cancel the training process or the button **Ok** to train the classifier. After clicking the button **Ok**, the pop-up window of figure A.36 informs that the training process need various minutes depends on the number of images used in the training, the number of cells and the complexity of the problem to learn for the classifier). While the OralImmunoAnalyser is trainig its classifier, you can not use it. When the training process finished, the pop-up window of upper panel of figure A.7 will be shown.

4. Once the classifier was trained, OIA assign a category to the cell clicking the button **Run** after the label *Classifier* in the lateral panel or going to submenu **Classify** in menu **Classification**. Obviously, you can only apply the classification process if there is a region of analysis selected with cells recognized on the image window.

A.12 Help menu

The **Help** menu is within the menu bar with the submenus **User Manual** and **About Us**. The submenu **About Us** pop-ups a window with a decription of OralImmunoAnalyser and informations about thanks and licence (see Figure A.37). The submenu **User Manual** does not do anything. The help to the users are provided in the current user guide. As well, if you put the mouse pointer on some widget (button, entry, label, etc) of the GUI, a pop-up message are shown with information about the functionality of that widget.

OralImmunoAnalyzer

1.0

Graphical software to analyze quantitatively immunohistochemistry images in the odontology field developped by the research centers CiTIUS (Center for Research in Intelligent techologies) and Stomatology Department of Medicine and Odontology Faculty, both of the University of Santiago de Compostela (Spain)

CITIUS website

Eva Cernadas

Figure A.37: Pop up dialogue to inform about help.

Bibliography

[1] Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine Süsstrunk. SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11):2274–2282, 2012.

[2] Zakaria A. Al-Tarawneh, Maite Pena-Cristóbal, Eva Cernadas, José Manuel Suarez-Peñaranda, Manuel Fernández-Delgado, Almoutaz Mbaidin, Mercedes Gallas-Torreira, and Pilar Gándara-Vila. Oralimmunoanalyser: a software tool for immunohistochemical assessment of oral leukoplakia using image segmentation and classification models. *Frontiers in Artificial Intelligence*, 7, 2024.

[3] Zakaria A. Al-Tarawneh, Ahmad S. Tarawneh, Almoutaz Mbaidin, Manuel Fernández-Delgado, Pilar Gándara-Vila, Ahmad Hassanat, and Eva Cernadas. Cell Detection in Biomedical Immunohistochemical Images Using Unsupervised Segmentation and Deep Learning. *Electronics*, 14(18), 2025.

[4] Mohammad Rizwan Alam, Kyung Jin Seo, Kwangil Yim, Phoebe Liang, Joe Yeh, Chifu Chang, and Yosep Chong. Comparative analysis of Ki-67 labeling index morphometry using deep learning, conventional image analysis, and manual counting. *Translational Oncology*, 51:102159, 2025.

[5] I Amezaga-Fernandez, JM Aguirre-Urizar, JM Suárez-Peñaranda, C Chamorro-Petronacci, I Lafuente-Ibáñez de Mendoza, X Marichalar-Mendia, A Blanco-Carrión, J Antúnez-López, and A García-García. Epidemiological, clinical, and prognostic analysis of oral squamous cell carcinoma diagnosed and treated in a single hospital in Galicia (Spain): a retrospective study with 5-year follow-up. *Med Oral Patol Oral Cir Bucal*, 29(1):e36–e43, 2024.

[6] Thaína A. Azevedo Tosta, Leandro A. Neves, and Marcelo Z. do Nascimento. Segmentation methods of H&E-stained histological images of lymphoma: A review. *Informatics in Medicine Unlocked*, 9:35–43, 2017.

[7] A. Bangor, P. Kortum, and J. Miller. Determining what individual SUS scores mean: Adding and adjective rating scale. *Journal of Usability Studies*, 4(3):114 – 123, 2009.

[8] Ariel E. Bayá, Mónica G. Larese, and Rafael Namías. Clustering stability for automated color image segmentation. *Expert Systems with Applications*, 86:258–273, 2017.

[9] Khushmeen Kaur Brar, Bhawna Goyal, Ayush Dogra, Mohammed Ahmed Mustafa, Rana Majumdar, Ahmed Alkhayyat, and Vinay Kukreja. Image segmentation review: Theoretical background and recent advances. *Information Fusion*, 114:102608, 2025.

[10] John Brooke. SUS: A retrospective. *Journal of Usability Studies*, 8(2):29 – 40, 2013.

[11] J. F. Canny. A computational approach to edge detection. *IEEE T Patt Anal*, 8(6):679–698, 1986.

[12] J. Carletta. Assessing agreement on classification tasks: the kappa statistic. *Comput. Linguist.*, 22(2):249–254, 1996.

[13] E. Cernadas, M. Fernández-Delgado, E. González-Rufino, and P. Carrión. Influence of normalization and color space to color texture classification. *Pattern Recogn.*, 61:120 – 138, 2017.

[14] Eva Cernadas. Applications of computer vision, 2nd edition. *Electronics*, 13(18), 2024.

[15] T.F. Chan and L.A. Vese. Active Contours Without Edges. *IEEE T Image Process*, 10(1):266–277, 2001.

[16] C.C. Chang and C.J. Lin. LIBSVM: A library for support vector machines. *ACM Trans. on Intel. Systs and Technol.*, 2:27:1–27:27, 2011.

[17] Sujata Chaudhari, Nisha Malkan, Ayesha Momin, and Mohan Bonde. YOLO Real Time Object Detection. *International Journal of Computer Trends and Technology*, 68, 2020.

[18] Dorin Comaniciu and Peter Meer. Mean Shift: A Robust Approach Toward Feature Space Analysis. *IEEE T Patt Anal*, 24(5):603–619, 2002.

[19] Adrián Cordido, Eva Cernadas, Manuel Fernández-Delgado, and Miguel A. García-González. CystAnalyser: A new software tool for the automatic detection and quantification of cysts in Polycystic Kidney and Liver Disease, and other cystic disorders. *PLOS Computational Biology*, 16(10):1–18, 10 2020.

[20] Daniel Cremers, Mikael Rousson, and Rachid Direche. A Review of Statistical Aproached to Level Set Segmentation: Integrating Color, Texture, Motion and Shape. *Intl J Comput Vis*, 72(2):195–215, 2007.

[21] Brunno Santos de Freitas Silva, Danielle Coelho Ribeiro Batista, Camila Ferro de Souza Roriz, Lorena Rosa Silva, Ana Gabriela Costa Normando, Alan Roger dos Santos Silva, Maria Alves Garcia Silva, and Fernanda Paula Yamamoto-Silva. Binary and WHO dysplasia grading systems for the prediction of malignant transformation of oral leukoplakia and erythroplakia: a systematic review and meta-analysis. *Clinical Oral Investigations*, 25(7), 2021.

[22] Eliane Pedra Dias, Nathália Silva Carlos Oliveira, Amanda Oliveira Serra-Campos, Anna Karoline Fausto da Silva, Licínio Esmeraldo da Silva, and Karin Soares Cunha. A novel evaluation method for Ki-67 immunostaining in paraffin-embedded tissues. *Virchows Arch*, 479:121–131, 2021.

[23] Richard O. Duda, Peter E. Hart, and David G. Stork. *Pattern Classification*. Wiley-Interscience, New York, NY, USA, 2000.

[24] AK El-Naggar, JKC Chan, JR Grandis, T Takata, and PJ Slootweg. *WHO Classification of Head and Neck Tumours*. IAPC, World Health Organization, 2017.

[25] Jianping Fan, Guihua Zeng, Mathurin Body, and Mohand-Said Hacid. Seeded region growing: an extensive and comparative study. *Pattern Recognition Letters*, 26(8):1139–1156, 2005.

[26] Pedro F. Felzenszwalb and Daniel P. Huttenlocher. Efficient Graph-Based Image Segmentation. *Intl J Comput Vis*, 59(2):167–181, 2004.

[27] Manuel Fernández-Delgado, Eva Cernadas, Senén Barro, and Dinani Amorim. Do we need hundreds of classifiers to solve real world classification problems? *J Mach Learn Res*, 15:3133–3181, 2014.

[28] Antonio Ferrer-Sánchez, Jose Bagan, Joan Vila-Francés, Rafael Magdalena-Benedito, and Leticia Bagan-Debon. Prediction of the risk of cancer and the grade of dysplasia in leukoplakia lesions using deep learning. *Oral oncology*, 132:105967, June 2022.

[29] Margarita Gamarra, Eduardo Zurek, Hugo Jair Escalante, Leidy Hurtado, and Homero San-Juan-Vergara. Split and merge watershed: A two-step method for cell segmentation in fluorescence microscopy images. *Biomedical Signal Processing and Control*, 53:101575, 2019.

[30] P. Gandara-Vila, M. Perez-Sayans, J-M Suarez-Penaranda, M Gallas-Torreira, J Somoza-Martin, M-D Reboiras-Lopez, A Blanco-Carrion, and A Garcia-Garcia. Survival study of leukoplakia malignant transformation in a region of northern Spain. *Med. Oral. Patol. Oral. Cir. Bucal.*, 23:413–420, 2018.

[31] Narjes Ghane, Alireza Vard, Ardeshir Talebi, and Pardis Nematollahy. Segmentation of White Blood Cells From Microscopic Images Using a Novel Combination of K-Means Clustering and Modified Watershed Algorithm. *Journal of Medical Signals & Sensors*, 7(2):92–101, 2017.

[32] Gerardo Gilligan, René Panico, Jerónimo Lazos, Rosana Morelatto, Paola Belardinelli, Maria Ines Criscuolo, Nicolas Bolesina, Ignacio Molina Ávila, Juan Martin Pimentel Solá, Emilce Rivarola, Maria Jimena Morgante, Maria Eugenia Ingrassia Tonelli, Javier Gimenez, Pablo Doratti, Laura Cecilia Werner, Ana Carolina Prado-Ribeiro, Thais Bianca Brandão, Aljomar Jose Vechiato-Filho, Maria Claudia Cuzzullin, Cristina Saldivia-Siracusa, Marcio Ajudarte Lopes, Michelle Agostini, Mario José Romañach, Aline Corrêa Abrahão, Thamyres Campos Fonsêca, Estela Kaminagakura, Ana Lia Anbinder, Kamilla Santos Alves, Maria Leticia de Almeida Lança, Sonia Sacsaquispe-Contreras, Karem Rodriguez Ibazetta, Roberto Gerber-Mora, Saray Aranda Romo, Francisco Javier Tejeda Nava, Karina Cordero-Torres, Ricardo Moreno Silva, Wilfredo Gonzalez-Arriagada, Raiza Toro, Felipe Martins Silveira, Natalia Gonzalez, Pablo Agustin Vargas, Alan Roger Santos-Silva, Ronell Bologna-Molina, Mariana Villarroel-Dorrego, Benjamin Martinez, and Eduardo Piemonte. Oral squamous cell carcinomas and oral potentially malignant disorders: A Latin American study. *Oral Diseases*, 30(5):2965–2984, 2024.

[33] Rafael Gonzalez and Richard Woods. *Digital Image Processing Global Edition*. Pearson Deutschland, 2017.

[34] Miguel A. Gonzalez-Moles, Isabel Ruiz-Avila, Alberto Rodriguez-Archilla, and Ildefonso Martinez-Lara. Suprabasal expression of Ki-67 antigen as a marker for the presence and severity of oral epithelial dysplasia. *Head & Neck*, 22(7):658–661, 2000.

[35] I. González-Ruiz, P. Ramos-García, H. Boujemaoui-Boulaghmoudi, N. Mjouel-Boutaleb, and M. A. González-Moles. Cancer Hallmarks Expression in Oral Leukoplakia: Systematic Review and Meta-Analysis. *Oral Diseases*, n/a(n/a), 2025.

[36] Anubha Gupta, Rahul Duggal, Shiv Gehlot, Ritu Gupta, Anvit Mangal, Lalit Kumar, Nisarg Thakkar, and Devprakash Satpathy. GCTI-SN: Geometry-inspired chemical and

tissue invariant stain normalization of microscopic medical images. *Medical Image Analysis*, 65:101788, 2020.

[37] Metin N. Gurcan, Laura E. Boucheron, Ali Can, Anant Madabhushi, Nasir M. Rajpoot, and Bulent Yener. Histopathological Image Analysis: A Review. *IEEE Reviews in Biomedical Engineering*, 7:97–114, 2014.

[38] Ashley L. Hein, Maheswari Mukherjee, Geoffrey A. Talmon, Sathish Kumar Natarajan, Tara M. Nordgren, Elizabeth Lyden, Corrine K. Hanson, Jesse L. Cox, Annelisse Santiago-Pintado, Mariam A. Molani, Matthew Van Ormer, Maranda Thompson, Melissa Thoene, Aunum Akhter, Ann Anderson-Berry, and Ana G. Yuil-Valdes. QuPath Digital Immunohistochemical Analysis of Placental Tissue. *Journal of Pathology Informatics*, 12(1):40, 2021.

[39] Lei Huang. *Normalization Techniques in Deep Learning*. Springer, 2022.

[40] S Humayun and V Ram Prasad. Expression of p53 protein and ki-67 antigen in oral premalignant lesions and oral squamous cell carcinomas: An immunohistochemical study. *Natl. J. Maxillofac. Surg.*, 2:38–42, 2011.

[41] IARC-OMS. Lip, oral cavity (international agency for research on cancer. world health organization). https://gco.iarc.fr/today/data/factsheets/cancers/1-Lip-oral-cavity-fact-sheet.pdf, 2020. Accessed: 2021-12-07.

[42] H. Irshad, A. Veillard, L. Roux, and D. Racoceanu. Methods for Nuclei Detection, Segmentation, and Classification in Digital Histopathology: A Review – Current Status and Future Potential. *IEEE Reviews in Biomedical Engineering*, 7:97–114, 2014.

[43] Sandra Jardim, João António, and Carlos Mora. Image thresholding approaches for medical image segmentation - short literature review. *Procedia Computer Science*, 219:1485–1492, 2023. CENTERIS – International Conference on ENTERprise Information Systems / ProjMAN – International Conference on Project MANagement / HCist – International Conference on Health and Social Care Information Systems and Technologies 2022.

[44] C. Kervrann, S. T. Acton, J.-C. Olivo-Marin, C.O.S. Sorzano, and M. Unser. Introduction to the Issue on Advanced Signal Processing in Microscopy and Cell Imaging. *IEEE Journal of Selected Topics in Signal Processing*, 10(1):3–5, 2016.

[45] Omar Kujan, Ammar Khattab, Richard J. Oliver, Stephen A. Roberts, Nalin Thakker, and Philip Sloan. Why oral histopathology suffers inter-observer variability on grading oral

epithelial dysplasia: An attempt to understand the sources of variation. *Oral Oncology*, 43(3):224–231, 2007.

[46] Yung-Chieh Lin, Yu-Pao Tsai, Yi-Ping Hung, and Zen-Chung Shih. Comparison between immersion-based and toboggan-based watershed image segmentation. *IEEE Transactions on Image Processing*, 15(3):632–640, 2006.

[47] Zhe Lu, Yi Liu, Junfeng Xu, Hongping Yin, Haiying Yuan, Jinjing Gu, Yan-Hua Chen, Liyun Shi, Dan Chen, and Bin Xie. Immunohistochemical quantification of expression of a tight junction protein, claudin-7, in human lung cancer samples using digital image analysis method. *Comput Meth Prog Biomed*, 155:179–187, 2018.

[48] Almoutaz Mbaidin, Eva Cernadas, Zakaria A. Al-Tarawneh, Manuel Fernández-Delgado, Rosario Domínguez-Petit, Sonia Rábade-Uberos, and Ahmad Hassanat. MSCF: Multi-Scale Canny Filter to Recognize Cells in Microscopic Images. *Sustainability*, 15(18), 2023.

[49] Almoutaz Mbaidin, Sonia Rábade-Uberos, Rosario Dominguez-Petit, Andrés Villaverde, María Encarnación Gónzalez-Rufino, Arno Formella, Manuel Fernández-Delgado, and Eva Cernadas. STERapp: Semiautomatic Software for Stereological Analysis. Application in the Estimation of Fish Fecundity. *Electronics*, 10(12), 2021.

[50] R. Mehrotra, Anurag Gupta, M. Singh, and Rahela Ibrahim. Application of cytology and molecular biology in diagnosing premalignant or malignant oral lesions. *Molecular Cancer*, 5:11 – 11, 2005.

[51] Andressa Fernanda Paza Miguel, DDS André Goulart Poletto, Bubacar Embaló, and Elena Riet Correa Rivero. Association between epithelial-mesenchymal transition markers, proliferative index, and oral epithelial dysplasia: an immunohistochemical study. *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology*, 135(6):904–913, 2023.

[52] Simone Monaco, Nicole Bussola, Sara Buttò, Diego Sona, Flavio Giobergia, Giuseppe Jurman, Christodoulos Xinaris, and Daniele Apiletti. AI models for automated segmentation of engineered polycystic kidney tubules. *Scientific Reports*, 14(1), 2024.

[53] Krishnendu Mondal, Rupali Mandal, and Badal Chandra Sarkar. Importance of Ki-67 Labeling in Oral Leukoplakia with Features of Dysplasia and Carcinomatous Transformation: An Observational Study over 4 Years. *South Asian J Cancer*, 9(2), 2020.

[54] Thayná M L Morais, C. Soares, J. A. Aguirre Urízar, J. Alberdi-Navarro, O. D. de Almeida, and F. Pires. Peri-implant peripheral giant cell lesions: report of 13 new cases

and comparative histological and immunohistochemical analysis with peripheral and central giant cell lesions. *Medicina Oral, Patología Oral y Cirugía Bucal*, 24:e739 – e745, 2019.

[55] T. Mungle, S. Tewary, I. Arun, B. Basak, S. Agarwal, R. Ahmed, S. Chatterjee, A. Maiti, and C. Chakraborty. Automated characterization and counting of Ki-67 protein for breast cancer prognosis: A quantitative immunohistochemistry approach. *Comput meth prog biomed*, 139:149–161, 2017.

[56] Torsten O Nielsen, Samuel C. Y Leung, David L Rimm, Andrew Dodson, Balazs Acs, Sunil Badve, Carsten Denkert, Matthew J Ellis, Susan Fineberg, Margaret Flowers, Hans H Kreipe, Anne-Vibeke Laenkholm, Hongchao Pan, Frédérique M Penault-Llorca, Mei-Yin Polley, Roberto Salgado, Ian E Smith, Tomoharu Sugie, John M. S Bartlett, Lisa M McShane, Mitch Dowsett, and Daniel F Hayes. Assessment of Ki67 in Breast Cancer: Updated Recommendations From the International Ki67 in Breast Cancer Working Group. *JNCI: Journal of the National Cancer Institute*, 113(7):808–819, 2020.

[57] Andrea Nüesch, Maria Paola Ferri, and Christine L. Le Maitre. Application and Validation of Semiautomatic Quantification of Immunohistochemically Stained Sections for Low Cellular Tissue Such as Intervertebral Disc Using QuPath. *JOR SPINE*, 8(1):e70054, 2025. e70054 JSP2-24-0314.R1.

[58] Edward Odell, Omar Kujan, Saman Warnakulasuriya, and Philip Sloan. Oral epithelial dysplasia: Recognition, grading and clinical significance. *Oral Diseases*, 27(8):1947–1976, 2021.

[59] Nobuyuki Otsu. A Threshold Selection Method from Gray-Level Histograms. *IEEE T Syst Man Cyber*, 9(1):62–66, 1979.

[60] E. V. Paravani, Daniea Osella, Javier Eduardo Diaz Zamboni, and Victor Hugo Casco. Diseño, evaluación y validación de un método semiautomático para cuantificación en imágenes de inmunohistoquímica. In *Congreso de métodos numéricos en ingeniería y ciencias aplicadas*, CIMENICS'2010, page 6, 2010.

[61] Slmaro Park, H. Jung, Young-Soo Jung, W. Nam, J. Cha, and Hwi-Dong Jung. Changes in Cellular Regulatory Factors before and after Decompression of Odontogenic Keratocysts. *Journal of Clinical Medicine*, 10, 2020.

[62] Liliana Aparecida Pimenta-Barros, Pablo Ramos-García, Miguel Ángel González-Moles, José Manuel Aguirre-Urizar, and Saman Warnakulasuriya. Malignant transformation of

oral leukoplakia: Systematic review and comprehensive meta-analysis. *Oral Diseases*, 31(1):69–80, 2025.

[63] Oscar Pina and Verónica Vilaplana. Unsupervised Domain Adaptation for Multi-Stain Cell Detection in Breast Cancer with Transformers. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 5066–5074, 2024.

[64] Jesper Reibel. Prognosis of oral pre-malignant lesions: significance of clinical, histopathological, and molecular biological characteristics. *Crit. Rev. Oral. Biol. Med.*, 14:47–62, 2003.

[65] Marina Rodríguez-Candela Mateos, Maria Azmat, Paz Santiago-Freijanes, Eva María Galán-Moya, Manuel Fernández-Delgado, Rosa Barbella Aponte, Joaquín Mosquera, Benigno Acea, Eva Cernadas, and María D. Mayán. Software breastanalyser for the semi-automatic analysis of breast cancer immunohistochemical images. *Scientific Reports*, 14, 2024.

[66] O. Ronneberger, P.Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 9351 of *LNCS*, pages 234–241. Springer, 2015. (available on arXiv:1505.04597 [cs.CV]).

[67] Curtis T. Rueden, Johannes Schindelin, Mark C. Hiner, Barry E. DeZonia, Alison E. Walter, Ellen T. Arena, and Kevin W. Eliceiri. ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics*, 18, 2017.

[68] S. Suzuki and K. Be. Topological structural analysis of digitized binary images by border following. *Comput Vis Graph Image Proc*, 30(1):32 – 46, 1985.

[69] Freja Lærke Sand, Sofie Lindquist, Gitte Lerche Aalborg, and Susanne K. Kjaer. The prognostic value of p53 and Ki-67 expression status in penile cancer: a systematic review and meta-analysis. *Pathology*, 57(3):276–284, 2025.

[70] Jeff Sauro. *A practical guide to the System Usability Scale: Background, Benchmarks & Best Practices*. CreateSpace Independent Publishing Platform, 2011.

[71] T. Seidal, A. Balaton, and H. Battifora. Interpretation and Quantification of Immunostains. *Am J Surg Pathol*, 25:1204–1207, 2001.

[72] Jiaqi Shao, Shuwen Chen, Jin Zhou, Huisheng Zhu, Ziyi Wang, and Mackenzie Brown. Application of U-Net and Optimized Clustering in Medical Image Segmentation: A

Review. *CMES - Computer Modeling in Engineering and Sciences*, 136(3):2173–2219, 2023.

[73] Jianbo Shi and Jiterdra Malik. Normalized Cuts and Image Segmentation. *IEEE T Patt Anal*, 22(8):888–905, 2000.

[74] Connor Shorten and Taghi M Khoshgoftaar. A survey on image data augmentation for deep learning. *Journal of big data*, 6(1):1–48, 2019.

[75] Lucas Antônio Fernandes Torres, Davi Said Gonçalves Celso, Maria L.R. Defante, Victoria Alzogaray, Mayara Bearse, and Ana Claudia Frota Machado de Melo Lopes. Ki-67 as a marker for differentiating borderline and benign phyllodes tumors of the breast: A meta-analysis and systematic review. *Annals of Diagnostic Pathology*, 75:152429, 2025.

[76] Gupta V, Ramalingam K, Yasothkumar D, D Debnath, and V Sundar. Ki-67 Expression as a Prognostic Marker: A Comparative Immunohistochemical Analysis of Oral Epithelial Dysplasia and Oral Squamous Cell Carcinoma. *Cureus*, 15(5), 2023.

[77] I. van der Waal. Potentially malignant disorders of the oral and oropharyngeal mucosa; terminology, classification and present concepts of management. *Oral oncology*, 45 4-5:317–23, 2009.

[78] S. Wang and J.M. Siskind. Image segmentation with ratio cut. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(6):675–690, 2003.

[79] S. Warnakulasuriya. Histological grading of oral epithelial dysplasia: revisited. *The Journal of Pathology*, 194:294–297, 2001.

[80] S. Warnakulasuriya, N. Johnson, and I. van der Waal. Nomenclature and classification of potentially malignant disorders of the oral mucosa. *Journal of oral pathology & medicine : official publication of the International Association of Oral Pathologists and the American Academy of Oral Pathology*, 36 10:575–80, 2007.

[81] S. Warnakulasuriya, O. Kujan, JM Aguirre-Urizar, JV Bagan, MA González-Moles, AR Kerr, G Lodi, FW Mello, L Monteiro, GR Ogden, P Sloan, and NW Johnson. Oral potentially malignant disorders: A consensus report from an international seminar on nomenclature and classification. *Oral Dis.*, October:19, 2020.

[82] Saman Warnakulasuriya, Omar Kujan, José M. Aguirre-Urizar, José V. Bagan, Miguel Ángel González-Moles, Alexander R. Kerr, Giovanni Lodi, Fernanda Weber

Mello, Luis Monteiro, Graham R. Ogden, Philip Sloan, and Newell W. Johnson. Oral potentially malignant disorders: A consensus report from an international seminar on nomenclature and classification, convened by the WHO Collaborating Centre for Oral Cancer. *Oral Diseases*, 27(8):1862–1880, 2021.

[83] Ross Wightman. PyTorch Image Models. `https://github.com/rwightman/pytorch-image-models`, 2019.

[84] Lipeng Xie, Jin Qi, Lili Pan, and Samad Wali. Integrating deep convolutional neural networks with marker-controlled watershed for overlapping nuclei segmentation in histopathology images. *Neurocomputing*, 376:166–179, 2020.

[85] Qing Xu, Zhicheng Ma, Na HE, and Wenting Duan. DCSAU-Net: A deeper and more compact split-attention U-Net for medical image segmentation. *Computers in Biology and Medicine*, 154:106626, 2023.

[86] Takahiro Yagyuu, Chiho Obayashi, Yoshihiro Ueyama, Masato Takano, Yuu Tanaka, Masahiko Kawaguchi, Maiko Takeda, Takahiko Kasai, and Tadaaki Kirita. Multivariate analyses of Ki-67, cytokeratin 13 and cytokeratin 17 in diagnosis and prognosis of oral precancerous lesions. *Journal of Oral Pathology & Medicine*, 44(7):523–531, 2015.

[87] Chang Zhang, Bingjie Li, Xiamei Zeng, XiaoSheng Hu, and Hong Hua. The global prevalence of oral leukoplakia: a systematic review and meta-analysis from 1996 to 2022. *BMC Oral Health*, 23(645):15, 2023.

[88] Zhengxia Zou, Keyan Chen, Zhenwei Shi, Yuhong Guo, and Jieping Ye. Object Detection in 20 Years: A Survey. *Proceedings of the IEEE*, 111(3):257–276, 2023.

[89] Snežana Đokić, Barbara Gazić, Biljana Grčar Kuzmanov, Jerca Blazina, Simona Miceska, Tanja Čugura, Cvetka Grašič Kuhar, and Jera Jeruc. Clinical and Analytical Validation of Two Methods for Ki-67 Scoring in Formalin Fixed and Paraffin Embedded Tissue Sections of Early Breast Cancer. *Cancers*, 16(7), 2024.

List of Figures

List of Tables

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

CAMPUS VIDA
CAMPUS DE EXCELENCIA INTERNACIONAL

Oral cancer can develop from potentially malignant disorders such as oral leukoplakia. Immunohistochemical staining is a complementary technique to improve the diagnosis and prognosis of oral leukoplakia. Cell counting in these images was traditionally performed manually, which was very laborious. OralImmunoAnalyser software is presented, which allows researchers to easily monitor and correct automatically recognized cells before quantification. A comprehensive comparison of 6 unsupervised segmentation methods against 2 supervised deep learning approaches for cell detection is presented. As expected the last ones achieved highest performance as the former, mainly due to incorporate problem knowledge in the training process, but required higher computational resources.