

**CARACTERIZACIÓN DEL MICROAMBIENTE TUMORAL EN BIOPSIAS DE
PACIENTES CON ADENOCARCINOMA GÁSTRICO DE TIPO INTESTINAL
BASÁNDOSE EN LA IDENTIFICACIÓN DEL VIRUS DE EPSTEIN BARR,
MEDIANTE PROCESAMIENTO DE IMÁGENES**



SANTIAGO PLAZAS PEMBERTHY

**Universidad del Cauca
Facultad de Ciencias Naturales, Exactas y de la Educación
Departamento de Física
Popayán
2018**

SANTIAGO PLAZAS PEMBERTHY

TRABAJO DE GRADO, MODALIDAD INVESTIGACIÓN PARA OPTAR AL TÍTULO DE
INGENIERO FÍSICO

**CARACTERIZACIÓN DEL MICROAMBIENTE TUMORAL EN BIOPSIAS DE
PACIENTES CON ADENOCARCINOMA GÁSTRICO DE TIPO INTESTINAL
BASÁNDOSE EN LA IDENTIFICACIÓN DEL VIRUS DE EPSTEIN BARR,
MEDIANTE PROCESAMIENTO DE IMÁGENES**

Directora:

ROSA AMALIA DUEÑAS CUELLAR

PhD, Profesora e investigadora de la Universidad del Cauca

Codirector:

RUBIEL VARGAS CAÑAS

PhD, Profesor e investigador de la Universidad del Cauca

Universidad del Cauca

Facultad de Ciencias Naturales, Exactas y de la Educación

Departamento de Física

Popayán

2018

Nota de Aceptación

Tanto la directora, el co-director y los jurados han evaluado este documento, escuchado la sustentación del estudiante y en su criterio los encuentran satisfactorios, por ende aprueban cumplido este requisito de grado para optar al título de Ingeniero Físico.

Firma Jurado, Cristian Heidelberg Valencia

Firma Jurado, Harold Jofre Bolaños

Firma Director, Rosa Amalia Dueñas Cuellar

Firma Co-Director, Rubiel Vargas Cañas

Notas adicionales

ÍNDICE DE FIGURAS

CAPÍTULO 1	1
Figura 1.1. Mapa conceptual representando el problema de investigación.	4
CAPÍTULO 2	6
Figura 2.0. Santiago Plazas, 2018. Micrografía 100x.	10
Figura 2.1. Diferentes tipos de núcleos celulares.	10
Figura 2.2. Composición del microambiente tumoral del carcinoma gástrico asociado al EBV.	11
Figura 2.3 Comparación entre linfocitos grandes y pequeños.	12
Figura 2.4 Tinción H&E.	13
Figura 2.5 Macrófagos y Fibroblastos en microscopia óptica	14
Figura 2.6. Del objeto real a la imagen digital.	15
Figura 2.7 Pasos fundamentales de un sistema de visión artificial (SVA).	16
Figura 2.8. Enfoque de la investigación en procesamiento digital de muestras histopatológicas en Colombia.	25
CAPÍTULO 3	26
Figura 3.0. Diagrama de la metodología.	26
Figura 3.1. Desarrollo metodológico para el montaje de microarreglos de tejido parafinado.	27
Figura 3.2. Laboratorio perteneciente al grupo de investigación en inmunología y enfermedades infecciosas de Facultad de Medicina de la Universidad del Cauca (GYEi).	28
Figura 3.3. Mapa conceptual representando el procedimiento paso a paso de la tinción mediante hibridación in situ cromogénica.	30
Figura 3.4. Instrumentos utilizados.	30
Figura 3.5 Microscopio utilizado.	33
Figura 3.6. Cámara #2, NIKON, COOLPIX S6300.	34
Figura 3.7. Cámaras # 1 y 3.	35
Figura 3.8. Ejemplo de fotografías de background.	35
Figura 3.9. Proceso de remoción de la sombra del objetivo.	36
Figura 3.10. Esquema de la metodología utilizada para el procesamiento de muestras.	38
Figura 3.11 Ejemplo de aplicación de transferencia de color de Reinhard.	41
Figura 3.12. Ejemplo de topografía de una imagen.	42
Figura 3.13. Ejemplo de aplicación de Watershed.	43
Figura 3.14. Representación del proceso de digitalización de los microarreglos.	44

CAPÍTULO 4	47
Figura 4.1. Corte superficial de tejido gástrico parafinado montado en placa portaobjetos.	47
Figura 4.2. Corte superficial de tejido gástrico teñido por hibridación in situ cromogénica y hematoxilina montado en placa portaobjetos y fijado con cubreobjetos.	47
Figura 4.3. Puntos del microarreglo de la Figura 4.2 bajo el microscopio en diferentes magnificaciones.	48
Figura 4.4. Estructuras que sirven para identificar la presencia del virus Epstein barr en el microambiente tumoral basándose en anotaciones del patólogo.	49
Figura 4.5. Ejemplo de imágenes preliminares.	51
Figura 4.6. Ejemplo de una de las imágenes adquiridas de tejido gástrico canceroso teñido por CISH positivo para EBV.	52
Figura 4.7. Proceso de extracción del fondo.	53
Figura 4.8. Resultados tras aplicar el algoritmo propuesto (Algoritmo 4).	54
Figura 4.9. Umbralizado de la imagen en la Figura 4.8.	56
Figura 4.10. Aplicación inversa de las máscaras 4.9.c y 4.9.d respectivamente.	57
Figura 4.11. Marcadores de la imagen.	58
Figura 4.12. Resultado de segmentación por watershed controlado por marcadores.	59
Figura 4.13. Conteo de células segmentadas.	60
Figura 4.14. Análisis cuantitativo de células con tinción intensa y moderada en los diferentes espacios de color.	62
Figura 4.15 Análisis visual de células con tinción intensa y moderada en diferentes espacios de color.	63
Figura 4.16. Resultado de aplicar el umbral determinado tras los análisis visual y cuantitativo, para identificar y contar células en la imagen.	64
Figura 4.17. Interfaz de la herramienta. 71	71
Figura 4.18. Resultados de la primera encuesta (Tabla 3.5) presentados en un diagrama de cajas y bigotes.	72
Figura 4.19. Resultado de la segunda encuesta (Tabla 3.6).	73

ÍNDICE DE TABLAS

CAPÍTULO 2	6
Tabla 2.1. Métodos recurrentes en procesamiento de imágenes histopatológicas, según la tendencia global. Elaboración propia, basado en Irshad H et. al. 2014 [1].	23
CAPÍTULO 3	26
Tabla 3.1. Estructura de los microarreglos.	27
Tabla 3.2. Clases de cámaras utilizadas y sus respectivos montajes en el microscopio.	34
Tabla 3.3. Valores considerados para cada parámetro de la cámara.	37
Tabla 3.4. Cuantificadores de la identificación del EBV en el microambiente tumoral.	45
Tabla 3.5. Primera encuesta a patólogos.	46
Tabla 3.6. Segunda encuesta a patólogos (encuesta tipo Likert).	46
CAPÍTULO 4	47
Tabla 4.1. Proporción entre tamaños de células del microambiente y área visual del tejido.	50
Tabla 4.2. Comparación de la medida de 50 μm de referencia en píxeles en el centro de la imagen (100%) y en los extremos para las diferentes cámaras.	50
Tabla 4.3. Dispersión de los valores de luminosidad en imágenes de background tomadas con cada una de las tres cámaras.	50
Tabla 4.4. Valores de ajustes de configuración de cámara seleccionados.	51
Tabla 4.5. Configuración de equipos para la digitalización de las placas con tejido.	65
Tabla 4.6 Comparación de valores representativos de configuración (Tabla 4.4) para la adquisición de las imágenes. El # de configuración se utiliza en la Tabla 4.7.	65
Tabla 4.7 Comparación de algunos de los resultados de las pruebas sobre las imágenes adquiridas bajo diferentes configuraciones.	66
Tabla 4.8. Presentación de valores de cuantificación de identificación del EBV calculados para cada punto de los microarreglos	69

CONTENIDO

INTRODUCCIÓN	1
1.1 Definición del problema de investigación.	1
1.2 Objetivos	5
1.2.1 Objetivo General	5
1.2.2 Objetivos Específicos	5
1.3 Presentación del documento	5
MARCO TEÓRICO	6
2.1 Marco conceptual	6
2.1.1 Cáncer.	6
2.1.2 Epstein Barr Virus.	7
2.1.3 Adenocarcinoma gástrico asociado al virus Epstein Barr.	8
2.1.4 Técnicas de tinción.	9
2.1.5 Microambiente tumoral.	10
2.1.6 Procesamiento digital de imágenes.	14
2.2 Marco referencial.	17
2.2.1 Referentes Internacionales	17
2.2.2 Referentes Nacionales	20
2.2.3 Referentes en la Universidad del Cauca	21
2.3 Análisis de la literatura	23
METODOLOGÍA	26
3.1 Procesamiento del tejido	27
3.1.1 Preparación de la biopsia	27
3.1.2 Hibridación in situ cromogénica	28
3.2 Definición de las características del microambiente	32
3.3 Adquisición de imágenes para el desarrollo de la herramienta	33
3.3.1 Parámetros de adquisición	33
3.4 Procesamiento digital de imágenes	37
3.4.1 Preprocesamiento	39
3.4.2 Segmentación	42
3.4.3 Clasificación de marcaciones fuertes y leves	44
3.5 Digitalización de las placas	44
3.6 Caracterización del TME basándose en la identificación de EBV.	45
3.7 Presentación de la herramienta	46
3.8 Evaluación	46

RESULTADOS Y ANÁLISIS	47
4.1 Procesamiento del tejido	47
4.2 Definición de las características del microambiente	49
4.3 Adquisición de imágenes para el desarrollo de la herramienta	50
4.3.1 Mejor magnificación	50
4.3.2 Mejor cámara	50
4.3.3 Configuración ideal de la cámara para obtener las mejores imágenes preliminares.	51
4.4 Procesamiento digital de imágenes	52
4.4.1 Preprocesamiento	52
4.4.2 Segmentación	58
4.4.3 Diferenciación de marcaciones fuertes y leves	61
4.5 Digitalización de las placas	65
4.6 Presentación de resultados de la caracterización	69
4.7 Presentación de la herramienta	70
4.8 Evaluación	71
CONCLUSIONES Y TRABAJOS FUTUROS	74
5.1 Conclusiones	74
5.2 Trabajos futuros	75
REFERENCIAS BIBLIOGRÁFICAS	77

1. INTRODUCCIÓN

1.1 Definición del problema de investigación.

El cáncer es la segunda causa de muerte en el mundo, genera un estimado de 9,6 millones de muertes cada año y representa una de cada seis (16%) de las muertes totales [2]. De éstas, el 70% ocurre en países de bajos y medianos ingresos como Colombia [3], en donde, en 2016 representó el 20% del total de las muertes; de éstas el 18% ocurrieron en el departamento del Cauca, donde de todos los tipos, el cáncer gástrico resultó el más mortífero con un 22% [4,5].

Según la organización mundial de la salud, el cáncer gástrico es una de las enfermedades con más alta mortalidad [2]. En consecuencia, el diagnóstico temprano cobra gran importancia para el tratamiento, ya que permite empezar a medicar al paciente, e indagar en el génesis de la enfermedad. El diagnóstico del cáncer se realiza analizando muestras de tejido (biopsia) bajo la lente de un microscopio óptico [6]; dicha tarea es realizada por un médico especialista en anatomía patológica, quien se encarga de estudiar, analizar las estructuras celulares y dictaminar si han ocurrido cambios morfológicos o estructurales que alteren la funcionalidad de las células, tejidos y órganos; para finalmente concluir si estos cambios son patológicos. Cuando el tejido se identifica como canceroso, el patólogo puede realizar otros exámenes en el microscopio para obtener información adicional sobre el tumor que no puede determinarse con la observación de la muestra teñida de manera rutinaria (hematoxilina - eosina), por ejemplo, el informe de patología puede tener también información obtenida por medio de métodos como: inmunohistoquímica o hibridación *in situ*, entre otros [7,8].

Que el dictamen del médico patólogo sea cualitativo en su totalidad, tiene efectos en su reproducibilidad, es decir, al ser una descripción de lo que cada médico observa, es altamente subjetivo y propicia discordancias entre colegas, e incluso entre posteriores revisiones a manos del mismo observador. Aunque la tasa de desacuerdo es baja, la reproducibilidad general del veredicto está influenciada por diferentes factores, como: ajustes del microscopio, cansancio visual, experiencia, entre otros [9–14].

Un mejor entendimiento del tejido así como la inclusión de componentes cuantitativos y la reducción de la diferencia inter-observacional se ha logrado desde diferentes enfoques metodológicos, entre los más relevantes se encuentran análisis físico-ópticos a través de *scattering* [15], patrones 'speckle'

[16,17]; microscopía confocal [18], automatización de procesos de tinción [19] y análisis digital de imágenes. Entre estas, el análisis digital de imágenes es catalogado como el nuevo paradigma de la patología moderna [20] gracias a su aplicabilidad y versatilidad.

Teniendo en cuenta lo anterior, y dado que el cáncer gástrico es uno de los de mayor frecuencia en el Departamento del Cauca, este proyecto, en conjunto con el grupo de investigación en inmunología y enfermedades infecciosas (GYEi) de la Universidad del Cauca, buscó desarrollar una herramienta de análisis de imágenes en el marco del proyecto 4487 de la vicerrectoría de Investigaciones (VRI) "Asociación de los herpes virus Epstein barr y Citomegalovirus con el adenocarcinoma gástrico de tipo intestinal en pacientes del Departamento del Cauca" para facilitar la identificación de uno de los factores etiológicos del cáncer gástrico, el cual está relacionado con la infección por el virus Epstein Barr (EBV) [21–26]. Este virus infecta a células epiteliales (células de tejido superficial) y a células B (células del sistema inmune) donde permanece por largos periodos de tiempo. Se debe mencionar que, de manera convencional, la relación entre el EBV y el cáncer gástrico se identifica con la ayuda de métodos de tinción diferencial (hibridación *in situ*), es decir, la muestra cambia de color en los sitios donde se encuentran células infectadas por el virus; así, el médico patólogo establece un veredicto, que es en naturaleza cualitativo y depende principalmente de su experiencia, seguido de la calidad de la muestra y de los equipos que tenga a disposición.

Incluir en el análisis de patología convencional un componente cuantitativo haciendo uso del procesamiento digital de imágenes, crea una línea de base que puede ser utilizada como punto de partida en la discusión de resultados entre colegas; esto con el fin de brindar un mejor diagnóstico a cada paciente, ya que a partir de los datos recolectados se podría definir aún mejor la relación del EBV con el cáncer gástrico. Además, mejorar visualmente la calidad de la imagen de la muestra analizada, brinda mayor comodidad para el análisis del patólogo, lo que impacta directamente la calidad del diagnóstico, permitiendo que se enfoquen los esfuerzos en la discusión de los resultados. De no contar con la adición del procesamiento digital de imágenes, la extracción de componentes cuantitativos representa un desafío para cualquier patólogo, ya que se trata de una tarea repetitiva, que consume el tiempo del especialista, y que requiere un ojo agudo y una gran capacidad de memoria; cualidades que se pueden complementar satisfactoriamente mediante el uso de la tecnología. Es así como la implementación de un sistema de visión artificial no reemplaza ninguna función del médico, al contrario, le brinda herramientas adicionales que permiten disminuir el tiempo de análisis e incrementar la replicabilidad del proceso de diagnóstico, haciéndolo más preciso.

Como ya se mencionó, cerca del 10% de los casos de cáncer gástrico se encuentran asociados a la infección por EBV [27], sin embargo esta asociación no es diagnosticada en el departamento del Cauca, por lo que se incurre en falencias tales como una apropiada terapia viral, que podría, de cierta manera, disminuir la transformación maligna del epitelio gástrico. Es ampliamente conocido que el diagnóstico temprano en conjunto con la correcta clasificación de los subtipos de cáncer incrementa la probabilidad de supervivencia y disminuye en igual proporción los costos asociados al tratamiento; en general el diagnóstico temprano mejora considerablemente la calidad de vida de los pacientes [2]. Consecuentemente, desarrollar un sistema de visión artificial valiéndose tanto del conocimiento médico patológico como del procesamiento digital de imágenes agilizará y mejorará el diagnóstico del adenocarcinoma gástrico asociado a EBV; y llevará el análisis de imágenes histopatológicas de lo analógico a lo digital, permitiendo que se adquieran ventajas como el uso de la nube (internet), para por ejemplo, aumentar la facilidad de compartir biopsias entre expertos y estudiantes de todo el mundo [20].

Estandarizar el procedimiento de obtención y análisis de imágenes, junto a la inclusión de componentes cuantitativos aporta objetividad al veredicto médico, y facilita el entendimiento del paciente sobre lo encontrado en la biopsia [20]. En este sentido, se ha demostrado que la herramienta del procesamiento digital de imágenes en microscopía es eficaz para mejorar procesos diagnósticos en segmentación de glóbulos rojos [28], segmentación de leucocitos en sangre periférica [29], segmentación y seguimiento de células [30], entre otras aplicaciones.

Teniendo en cuenta lo anterior y debido a que el análisis de biopsias de adenocarcinoma gástrico asociado a EBV todavía no se ha visto beneficiado por la implementación de esta herramienta, y dada su mencionada importancia, este proyecto tiene por objetivo *“Caracterizar el microambiente tumoral en biopsias de pacientes con adenocarcinoma gástrico de tipo intestinal basándose en la identificación del virus de Epstein Barr, mediante procesamiento de imágenes”*. Considerando este contexto, surge la siguiente pregunta de investigación: ¿Cómo pueden incorporarse las técnicas de procesamiento digital de imágenes en la caracterización del microambiente tumoral de biopsias de pacientes con adenocarcinoma gástrico de tipo intestinal basándose en la identificación del virus de Epstein Barr?

Un resumen conceptual del problema de investigación se muestra en la Figura 1.0.

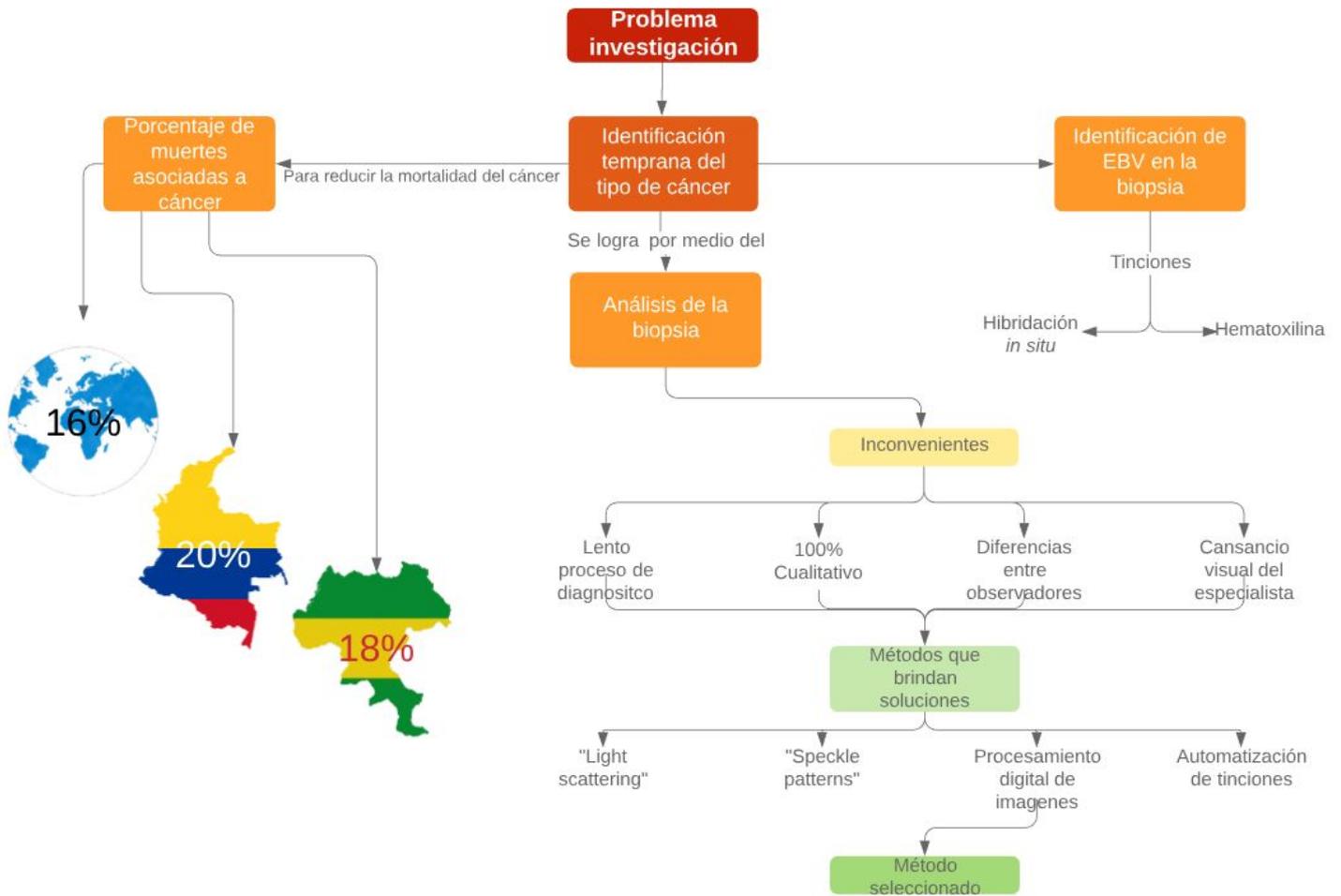


Figura 1.0. Mapa conceptual representando el problema de investigación. Los métodos tradicionales para la correcta identificación del tipo de cáncer gástrico a través de la identificación de EBV tienen grandes inconvenientes que pueden ser mitigados al incorporar procesamiento digital de imágenes. Elaboración propia.

1.2 Objetivos

1.2.1 Objetivo General

Caracterizar el microambiente tumoral en biopsias de pacientes con adenocarcinoma gástrico de tipo intestinal basándose en la identificación del virus de Epstein Barr, mediante procesamiento digital de imágenes.

1.2.2 Objetivos Específicos

- Identificar las condiciones óptimas para la captura de la información a procesar mediante la observación del proceso actual.
- Implementar técnicas de procesamiento de imágenes que permitan identificar células B y células tumorales infectadas con el virus de Epstein Barr presentes en el microambiente de muestras de adenocarcinoma gástrico de tipo intestinal.
- Evaluar la eficacia del método de procesamiento de imágenes en biopsias de adenocarcinoma gástrico de tipo intestinal.

1.3 Presentación del documento

La organización de este documento se planteó en cinco capítulos siguiendo un hilo lógico para la exposición del trabajo de investigación:

Primero, en la introducción se expone el contexto, el problema de investigación y se plantean los objetivos a cumplir. Posteriormente, en el marco teórico se explican detalladamente los fundamentos de la temática, empezando por cáncer gástrico, luego, Virus de Epstein Barr, después, técnicas de tinción y microambiente tumoral, para finalizar con procesamiento digital de imágenes. Más adelante se incluye el marco referencial que permite tener un panorama general del estado del arte y el análisis de los trabajos más relevantes. Posteriormente en el tercer capítulo, que es el pilar principal del documento, se aborda la metodología de investigación explicando detalladamente cada paso del trabajo y cómo se desarrolló. En el cuarto capítulo se encuentran los principales resultados obtenidos y su interpretación. Por último en el quinto capítulo se revelan las conclusiones finales y se presenta una discusión sobre trabajos futuros a partir de esta investigación.

2. MARCO TEÓRICO

2.1 Marco conceptual

2.1.1 Cáncer.

Por definición el cáncer es el resultado de lesiones genéticas que dan lugar a un fenotipo alterado en una célula afectada. Después de superar el crecimiento normal y controles reguladores de muerte, una célula cancerosa puede seguir evolucionando y adquiriendo mutaciones adicionales para sostener su supervivencia en medio de los intentos del organismo huésped para combatir los efectos deletéreos que ésta causa y restablecer el orden en el tejido afectado [31]. El término cáncer se define también como un grupo de enfermedades que se caracterizan por el crecimiento descontrolado y la diseminación de células anormales (tumor), una vez que la propagación no es controlada, una serie de eventos desfavorables pueden ocurrir en el organismo, y en muchos casos resultar en muerte. El cáncer es causado por factores externos (agentes físicos, químicos o biológicos) y factores internos (mutaciones heredadas y espontáneas, hormonas, factores epigenéticos y condiciones inmunológicas, entre otros). Estos factores causales pueden actuar en conjunto o en secuencia para iniciar o promover la carcinogénesis [32].

Cáncer gástrico.

El cáncer gástrico (CG) es una neoplasia que se origina en el estómago y abarca diferentes tipos histológicos entre los que se encuentra el adenocarcinoma, el linfoma no Hodgkin y los tumores carcinoides, sin embargo el más frecuente es el adenocarcinoma que a su vez se agrupa en dos tipos, intestinal y difuso, según la clasificación de Lauren [33]. El desarrollo de esta patología está asociado con agentes infecciosos como la bacteria *H. pylori* [34] y el virus Epstein barr [35], además de factores genéticos [36], ambientales [37], comportamentales (alto consumo de sal en la dieta, el hábito de fumar, el consumo de alcohol, baja ingesta de fibra en la dieta) [38,39], étnicos y sociales (nivel socioeconómico) [40].

Según las estimaciones de GLOBOCAN 2018 (base de datos del observatorio global del cáncer para estimar la incidencia y mortalidad del cáncer), a nivel mundial el CG es responsable por más de 1.000.000 casos nuevos de cáncer, de los cuales, 783.000 terminan en muerte, lo que lo convierte en el quinto cáncer con diagnóstico más frecuente y en la tercera causa de muerte por cáncer [41]. Es de resaltar que, mientras que en algunos continentes este tipo de cáncer ha

disminuido, Suramérica se encuentra en el tercer lugar de incidencia después de Asia oriental y Europa oriental [42]. Colombia se considera como un país de alta incidencia y mortalidad por CG a nivel mundial. El tipo de CG más frecuente es el adenocarcinoma (90%), que es un tipo de tumor que se define como '*una neoplasia del tejido epitelial*' [33,43,44] (crecimiento anormal del tejido que recubre las cavidades y superficies estructurales del cuerpo) y que además tiene un origen glandular o posee características glandulares. En el Suroccidente colombiano, caracterizado por sus regiones montañosas, las tasas de CG son elevadas [45], sin embargo, en el departamento del Cauca son pocos los estudios que visibilizan esta problemática.

Por lo tanto, determinar la causa de la patología es importante ya que permite clasificar correctamente el subtipo de cáncer y proporciona una hoja de ruta para la estratificación y los ensayos de terapias dirigidas. En este sentido el proyecto Atlas Genoma del Cáncer (TCGA) propuso en el año 2014 una clasificación molecular para el CG, dividiéndolo en cuatro subtipos: positivo para virus del Epstein Barr (CaGVEB), inestabilidad de microsatélites, genómicamente estable y con inestabilidad cromosómica [35].

2.1.2 Epstein Barr Virus.

El virus Epstein Barr (EBV) es un tipo de herpes muy común; se transmite a través de la saliva, transfusiones sanguíneas, contacto sexual, trasplantes o por vía perinatal. Su prevalencia es casi total en la población (90 - 99%), especialmente en poblaciones con condiciones socioeconómica bajas; por ejemplo, en países en vía de desarrollo el 90% de los niños se infectan desde edades tempranas [46]. La infección primaria es usualmente asintomática, pero resulta en una infección vitalicia, porque el virus persiste en las células B de memoria [47,48]. No obstante, el virus puede reactivarse por varias causas, produciendo daños en el organismo; incluso, como ya ha sido mencionado, se asocia con la génesis del CG [47,49,50].

Entrando en detalle, el virus ingresa al organismo y se instala en dos huéspedes específicos, las células B (célula del sistema inmune) y las células epiteliales, a las que ingresa a través de la interacción entre ciertas proteínas de membrana y receptores de las células atacadas, lo que les permite la entrada a la célula [51]. Una vez dentro, el virus ingresa hasta el núcleo y utiliza la maquinaria de la célula infectada para replicar su material genético y crear millones de copias de sí mismo, matando la célula al final del proceso, lo que desencadena una reacción inmune del cuerpo (fiebre, etc), esto es conocido como ciclo lítico. En células epiteliales este proceso ocurre directamente después de la infección, mientras

que actualmente está en discusión si en las células B sólo ocurre después de pasar de un estado de latencia a uno activo [52].

La segunda posibilidad es que el virus permanezca latente en el núcleo del huésped indefinidamente, lo que le permite mezclar su material genético (inclusión cromosomal), y transmitir la infección a nuevas células a través del proceso de división celular, en un proceso conocido como ciclo lisogénico [53].

A la fecha, se conocen tres estados de latencia para el EBV, dependiendo del estado de latencia sólo una porción de los genes del virus se expresan [54], cada programa de latencia conduce a la producción de un conjunto limitado y distinto de proteínas y ARN virales [55]. Por ejemplo, EBNA1, "*Epstein-Barr virus nuclear antigen 1*" y EBER (ARN pequeño codificado por el virus) son comunes a los tres estados de latencia, y son los blancos moleculares más comunes para identificar la presencia del virus en diversas muestras biológicas [56].

2.1.3 Adenocarcinoma gástrico asociado al virus Epstein Barr.

El adenocarcinoma gástrico asociado a EBV (CaGEBV) es un subtipo de CG que presenta características clínicas y genéticas particulares, que tienen origen a partir del virus Epstein Barr, el cual se vale de varios mecanismos para desencadenar el proceso carcinogénico [57]. El CaGEBV se presenta entre el 2 - 20% de la población mundial y anualmente se diagnostican de 200.000 a 876.000 nuevos casos de CaGVEB, de los cuales se producen 69.081-90.000 muertes cada año [58,59]. En Colombia, la tasa más alta de mortalidad por CaGVEB se reporta en regiones montañosas y la más baja en regiones planas [60]. El CaGEBV cuenta con una baja tasa de metástasis a ganglios linfáticos y una infiltración densa y difusa de células T citotóxicas o linfocitos atípicos en las biopsias de mucosa gástrica [61].

El linfocito B infectado por EBV se activa y multiplica sin ser debidamente regulado por el sistema inmunológico (linfocitos T) en una frecuencia del 5-15% de los casos [62,63]. A pesar de que la célula B libera anticuerpos, estos no reaccionan con los antígenos del EBV. De esta forma se da un incremento en el número de células B portadoras de EBV, que atraen células T, cuya función es inhibir la activación de las células B. En consecuencia, en las zonas con mayor infección por EBV se encuentra una gran cantidad de linfocitos B y T [64]. La corrupción del material dentro del núcleo de las células infectadas por EBV ocasiona errores en la reproducción celular, lo que contribuye a la generación de cánceres de tipo epitelial, como el adenocarcinoma gástrico. En este orden de ideas numerosas investigaciones vinculan al EBV con la carcinogénesis [21-26].

Como se mencionó anteriormente las células gástricas malignas y positivas para EBV expresan una serie de genes [62], de los cuales EBERs en específico han sido utilizado con mayor frecuencia como marcador diagnóstico, ya que se encuentran expresados en todas las neoplasias asociadas al mismo [56], además se ha demostrado que estos micro-ARN están involucrados en los fenómenos de oncogénesis, adhesión celular, señales de transducción, regulación de la apoptosis, expresión y secreción del factor de crecimiento similar a la insulina y del factor de crecimiento autocrino de las células gástricas cancerígenas [58].

2.1.4 Técnicas de tinción.

Hematoxilina y eosina.

La técnica de tinción mediante el uso de hematoxilina y eosina (H&E) es la más utilizada en histología y patología rutinaria. La hematoxilina es un tinte básico y colorea las estructuras ácidas de un color azul morado. Núcleos y ribosomas tienen una fuerte afinidad a este tinte debido a su alto contenido de ADN y ARN. La eosina en contraste es un tinte ácido que tiñe las estructuras básicas de rojo o rosado. La mayoría de las proteínas citoplasmáticas son básicas, por ende son reconocidas por la eosina [65].

Hibridación *in Situ*.

La hibridación *in situ* (In situ es latín para “en su lugar original”) es una técnica que utiliza una sonda (secuencia de ácidos nucleicos complementaria a la secuencias de ADN o ARN que quiere ser identificada) para localizar o detectar donde están presentes secuencias específicas de ADN o ARNm (ARN mensajeros) en muestras de un tejido fijado, células o cromosomas. La estrategia actual utiliza, además de la sonda de ADN o RNA, un sistema de detección con anticuerpos, lo que aumenta la accesibilidad, sensibilidad y especificidad de la técnica [66].

La técnica de hibridación *in situ* colorimétrica (CISH), se basa en la adición de un cromógeno a la sonda o al anticuerpo complementario. Esta técnica se utiliza para localizar información molecular en tejidos, tipos de células, cromosomas u otras estructuras específicas, por lo tanto para su interpretación se requiere evaluar señales dentro de un contexto morfológico y requiere la preservación de un marco estructural reconocible [67].

En resumen, el proceso de hibridación *in situ* comienza con la preparación de la muestra (tejido) a través de un pretratamiento térmico y la digestión de la enzima endógena peroxidasa; posteriormente se realiza el proceso de desnaturalización (separación de hebras complementarias, si se trabaja con ADN) y se adiciona la

sonda para que se dé la hibridación (unión de la cadena marcada con cromógeno a su cadena complementaria). Posteriormente se adicionan enzimas que reaccionan con las cadenas de la sonda y finalmente cromógenos que reaccionan con las enzimas. Estas reacciones se efectúa a través de anticuerpos que detectan antígenos específicos (como ocurre en la inmunohistoquímica) [67].

En todos los casos el resultado de la prueba de CISH debe ser leído y examinado por un médico patólogo calificado, quien basándose en la información visual que pueda obtener del lente del microscopio y su propia experiencia, dará un veredicto [66]. Un ejemplo de tejido teñido por CISH se presentan en la Figura 2.0.

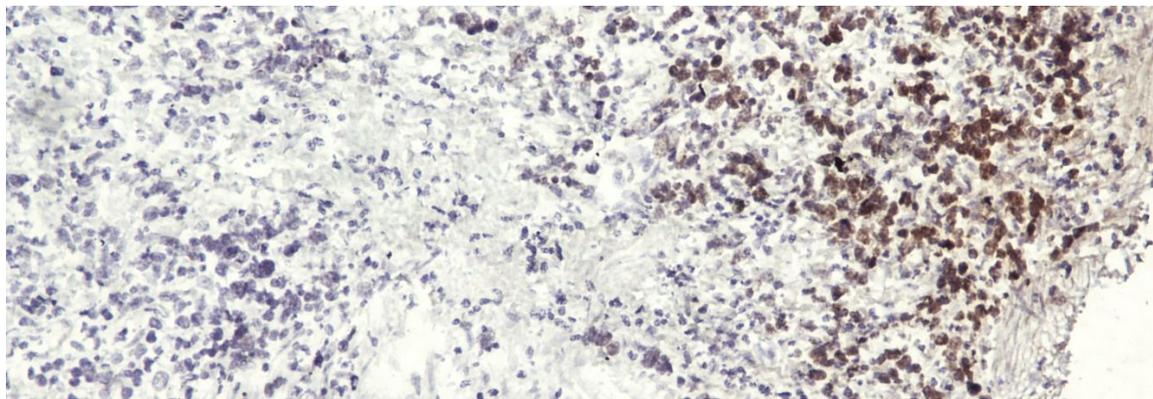


Figura 2.0. Santiago Plazas, 2018. Micrografía 100x. Tejido teñido por hibridación in situ, positivo para EBV (células coloreadas de café).

2.1.5 Microambiente tumoral.



(a)

(b)

(c)

(d)

Figura 2.1. Diferentes tipos de núcleos celulares. (a) Linfocito; (b) Epitelial; (c) Epitelial cancerígeno; (d) Epitelial (mitosis) En general el núcleo de los linfocitos es regular y de menor tamaño que el núcleo de la célula epitelial. Tomado de Irshad H et. al. 2014 [1].

El microambiente tumoral (TME) del CG es el campo de batalla durante el proceso neoplásico, donde células inmunes y otras células del tejido conectivo (estromales) pueden favorecer o mitigar la proliferación de células tumorales, en la Figura 2.1 se observan algunas de estas células. En conjunto con el patrón de expresión de los genes, el microambiente tumoral está profundamente relacionado con la progresión del tumor. El TME del CG consiste en varios tipos

de células estromales, “células derivadas de la médula ósea” (BMDCs), “mastocitos” (MCs), “linfocitos que infiltran los tumores” (TILs), “macrófagos asociados al tumor” (TAMs) y “fibroblastos asociados a cáncer” (CAFs) [68].

En el microambiente tumoral del CG asociado a EBV (GC_TME_EBV+), el número de células no cancerígenas es mayor que en el microambiente tumoral de CG no asociado a EBV (GC_TME_EBV-) [69]. Sin embargo el tamaño y composición del TME puede variar significativamente entre casos específicos [70]. Una representación esquemática de las células presentes en el GC_TME_EBV+ se muestra en la Figura 2.2.

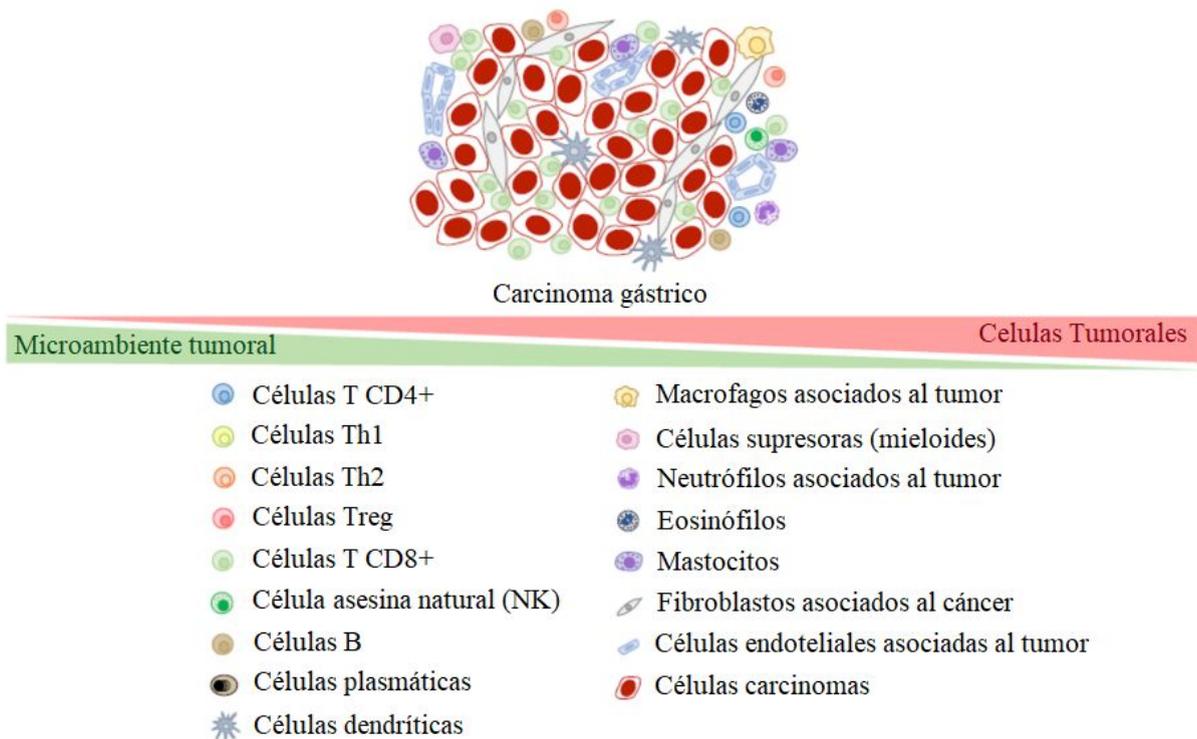


Figura 2.2. Composición del microambiente tumoral del carcinoma gástrico asociado al EBV. Se presentan las diferentes estructuras celulares que se pueden encontrar en el microambiente tumoral gástrico asociado a EBV. Modificado de Tan GW et al. 2018 [70].

Morfología de los linfocitos B y T.

Los linfocitos fueron descritos por primera vez en 1774 [71] y en los estudios clásicos de sangre y tejidos fueron definidos como células esféricas u ovoides con diámetros entre los 6 y 15 μm al ser aplanados en portaobjetos de vidrio. Algunos de estos estudios describieron dos tipos de linfocitos basados en el tamaño: linfocitos grandes con diámetros de 9 a 15 μm y linfocitos pequeños con diámetros de 6 a 9 μm [72], como se observa en la Figura 2.3. Sin embargo los linfocitos también se encuentran clasificados por su función específica en el sistema inmune: asesinos naturales (NK), linfocitos B y linfocitos T. Se

encontraron gran número de linfocitos grandes en pacientes con infecciones virales agudas, en cambio en pacientes con enfermedades autoinmunes se observó un incremento de linfocitos pequeños, por lo que se vinculó el tamaño del linfocito con su reactividad o activación [73].

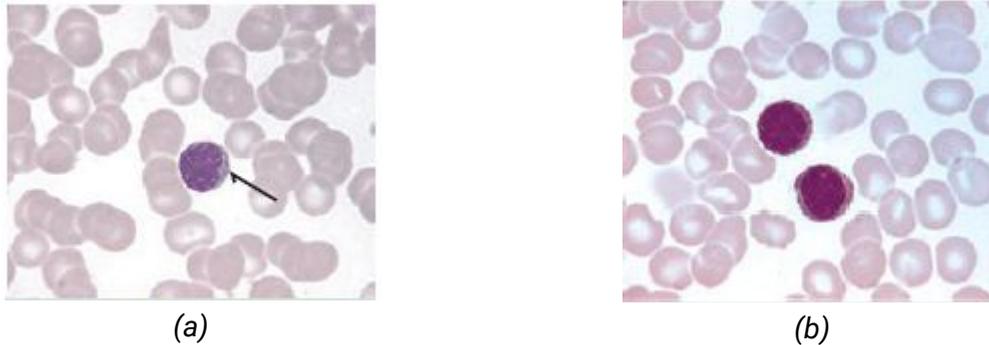


Figura 2.3 Comparación entre linfocitos grandes y pequeños. (a) *Linfocito pequeño o no reactivo;* (b) *Linfocitos grandes o reactivos;* Se observa una diferencia en la coloración del linfocito pequeño, el cual tiene un color más pálido a comparación de los linfocitos reactivos. Además los linfocitos reactivos tienden a crecer más que los no reactivos, por lo tanto se los llama linfocitos grandes. Tomado de Kierszenbaum AL et al. 2015 [74].

Hoy se conoce que a partir de la activación de los linfocitos ocurren cambios en su morfología, ya que la estimulación de estos está asociada a una compleja secuencia de eventos morfológicos y bioquímicos. La activación de los linfocitos B y T resulta en la transformación de linfocitos pequeños (en reposo), en células grandes proliferantes con citoplasma altamente basofílico (se denomina basófilo a cualquier célula que se tiñe fácilmente con colorantes), con cromatina condensada de forma irregular o difusa, y contornos nucleares redondos o ligeramente irregulares (Figura 2.4). Los núcleos de estas células se pueden observar en microscopía de luz, pero no sobresalen demasiado [73].

La mayoría de los linfocitos encontrados normalmente son pequeños con un núcleo ovalado en forma de riñón que se tiñe de púrpura debido a la densa concentración de cromatina que ahí reside y su núcleo ocupa aproximadamente el 90% del área celular [73]. Como se mencionó anteriormente, los principales linfocitos son los B y T, cuya mayor diferencia morfológica es que el promedio de diámetros de los linfocitos B es mayor, sin embargo esta diferencia es más pequeña que la variabilidad biológica natural que puede presentar cada célula [75].

Los linfocitos B se encuentran dispersos, usualmente en el estroma y los linfocitos T se encuentran distribuidos entre las células tumorales, aunque también se encuentran en el estroma. Sin embargo los NK no se detectan en el GC_TME_EBV+. El virus EBV desencadena infiltración de linfocitos T en el TME, como consecuencia se localizan muchos más linfocitos T en el GC_TME_EBV+

que en el GC_TME_EBV-- [69]. Los linfocitos activados proliferan y maduran en linfocitos efectores y células de memoria. Las células efectoras incluyen células T auxiliares, células T citotóxicas y células plasmáticas [73,76].

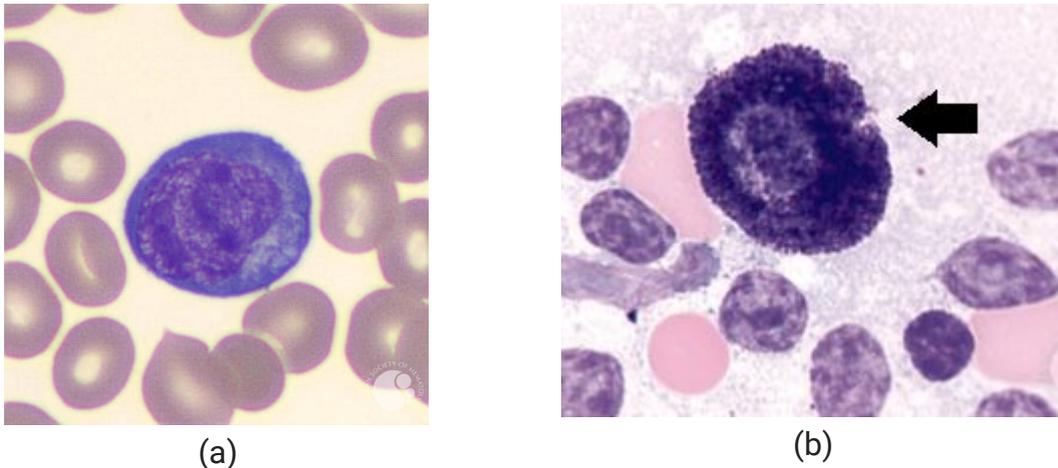


Figura 2.4 Tinción H&E. (a) *Linfocito grande/reactivo* [73]. (b) *Mastocito* [77]. En el mastocito se puede observar granulado el citoplasma y un núcleo de menor tamaño.

Morfología de los mastocitos.

La morfología de estas células, originarias de la médula ósea, varía de acuerdo a su hábitat, cuando se encuentran en tejido conectivo laxo son redondeados, pero cuando se encuentran cercanos a los vasos sanguíneos se tornan ovoides y alargados, en las fibras dérmicas poseen forma de huso (instrumento utilizado para hilar), de estrella o filiforme. Tienen un diámetro medio de 12.6 - 13.5 μm , el núcleo es relativamente grande con diámetros de 4-7 μm , con cantidades de cromatina variable. Uno, dos o más núcleos pueden estar presentes [78], se observa un mastocito en la Figura 2.4.

Morfología de los macrófagos asociados al tumor.

Estos son macrófagos que han sido reclutados al microambiente tumoral por acción de quimiocinas. En CG, los macrófagos son más comúnmente vistos en el estroma [79], y se reportan en cantidades significativamente menores en GC_TME_EBV+ en comparación a GC_TME_EBV-- [80]. Los macrófagos son células de grandes dimensiones (diámetro entre 25 a 50 μm) y poseen un núcleo grande, único y central (Figura 2.5).

Morfología de los fibroblastos asociados al cáncer.

Los fibroblastos son las células que proporcionan la estructura en forma de entramado (estroma) a los tejidos, siendo las células más comunes del tejido

conectivo y pueden producir colágeno y fibras elásticas. Aquellos asociados al cáncer son fibroblastos activados en el TME que tienen un inmunofenotipo diferente en comparación con los fibroblastos inactivos. Son un tipo de célula abundante en el el TME y su cantidad no puede disminuir ya que no pueden hacer apoptosis. Bajo el microscopio óptico tiene forma de huso con un núcleo elíptico; generalmente su citoplasma no resalta en este tipo de microscopía [74].

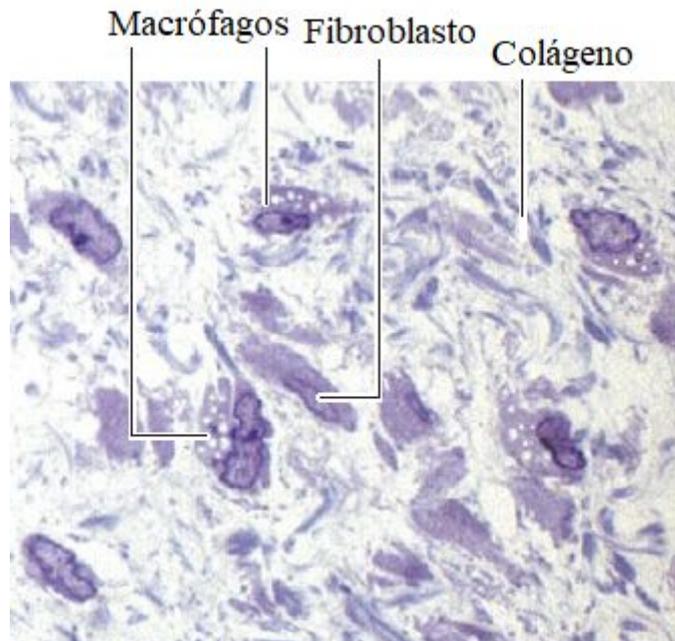


Figura 2.5 Macrófagos y Fibroblastos en microscopia óptica.
Modificado de Kierszenbaum AL et al. 2015 [74].

2.1.6 Procesamiento digital de imágenes.

El procesamiento de imágenes, hace parte de una disciplina llamada visión artificial, que trata de emular uno de los sentidos más importantes del ser humano, es decir lograr que una máquina obtenga información concreta del entorno físico utilizando receptores ópticos; incluyendo las fases de adquisición de datos y de procesamiento, en el caso general, los ojos y todos sus componentes son emulados por cámaras fotográficas digitales, y el cerebro que entiende la información, es emulado por un programa de computadora [81].

En concreto, la imagen digital está compuesta por píxeles y es de naturaleza discreta, es decir cada píxel sólo puede tomar una cantidad finita de valores y una imagen solo tener una cantidad finita de pixeles, así entre más pixeles estén representando una escena real, mayor será la resolución de la imagen; el valor del píxel representa el voltaje que se suministra a una pantalla en cada punto, en otras palabras esto es la intensidad de brillo de la pantalla en ese punto. Normalmente los monitores cuentan con tres canales, rojo, verde y azul, porque imitan la forma de ver del ser humano, quien tiene receptores visuales para esos

colores en específico. En consecuencia la mayoría de las imágenes poseen tres capas, o canales, que juntos dan la apariencia de formar los demás colores, sin embargo existen imágenes que poseen información espectral adicional en forma de más canales [82]. En la Figura 2.6 se observa el proceso de digitalización de un elemento real.

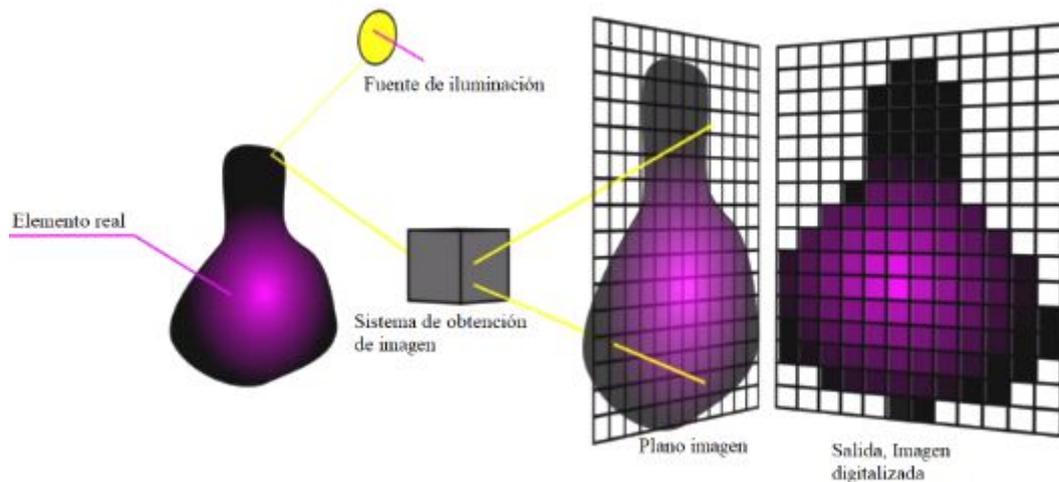


Figura 2.6. Del objeto real a la imagen digital. Se transforma una imagen real e infinita en un mapa de bits finito, basándose en la intensidad y frecuencia de la luz detectada. Modificado de González y Woods 2017 [83].

Gracias a la estructura matricial de las imágenes digitales, se traslada el trabajo de procesar imágenes al dominio de las matemáticas, especialmente del álgebra lineal, que ofrece los mecanismos para transformar una imagen en otra, sin embargo debido a la complejidad y tamaño de las operaciones requeridas, se delegan a un computador. Por consiguiente el procesamiento digital de imágenes se realiza mediante programación, basándose en la creación y aplicación de algoritmos matemáticos con el fin de obtener imágenes mejoradas [84].

En contraste los sistemas de visión artificial (SVA) buscan obtener datos precisos de una escena, extraer información de bajo nivel (información que comprenden las máquinas) y representarla en alto nivel (información que comprenden los seres humanos), por ejemplo partiendo de la fotografía de una mano, extraer información y representar la cantidad de dedos en un número conciso. El procesamiento de imágenes está incluido en los SVA, ya que permite convertir información visual en información de bajo nivel, es decir transformar imágenes complejas en imágenes simples, que pueden ser interpretadas. Los sistemas de visión artificial permiten obtener resultados consistentes en tareas repetitivas, como por ejemplo contar objetos, identificar colores, formas o movimiento [85].

El desarrollo de SVA sigue los pasos lógicos evidenciados en la Figura 2.7 y se basa en el conocimiento del tema y dominio del problema, como pilares de un SVA exitoso.

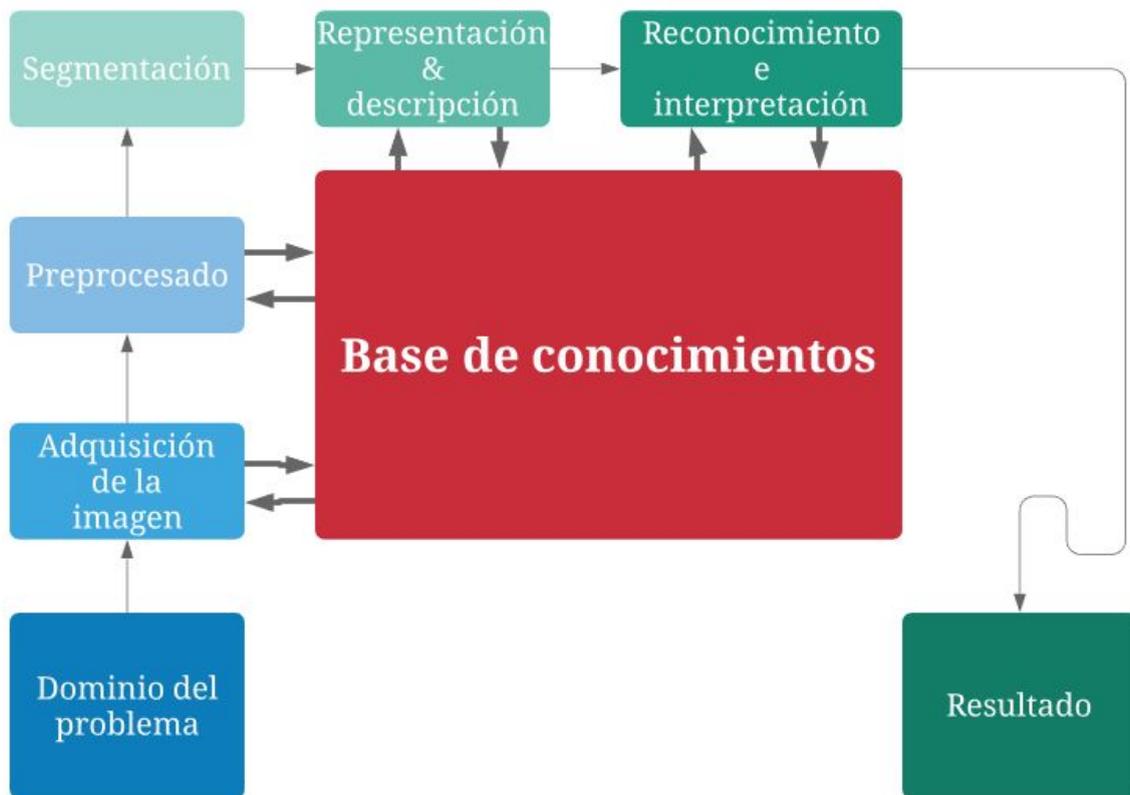


Figura 2.7 Pasos fundamentales de un sistema de visión artificial (SVA). Todos los pasos tienen como eje central la base de conocimientos y se permite regresar a un paso anterior conforme se comprende mejor la base. Modificado de González y Woods 2017 [83].

Un SVA necesita imágenes digitales como sustento, por lo que el proceso de adquisición de dichas imágenes es considerado un punto crítico. “Mejores” imágenes requieren menos procesamiento para obtener resultados. Sin embargo, no existe la imagen perfecta, por lo que se pre-procesan las imágenes obtenidas, “mejorando” las imágenes adquiridas (por ejemplo cambios de contraste). Una vez se tienen imágenes idóneas, se procede a segmentar, es decir, a dividir la imagen en sus partes constituyentes. Posteriormente, a cada una de las partes se la describe mediante el uso de características representativas. Para finalizar, se busca reconocer un grupo de objetos que compartan características similares, ya sea basándose en las descripciones de objetos bien diferenciados, en métodos computacionales o técnicas de inteligencia artificial [86].

2.2 Marco referencial.

2.2.1 Referentes Internacionales

El uso masivo del diagnóstico asistido por computadora (CAD, computer-assisted diagnosis) puede ser rastreado a la década de los 90s con el surgimiento de la mamografía digital [87]; convirtiendo hoy el uso de CAD en una tarea rutinaria en la detección del cáncer de mama [88]. A partir de esta década se observó un incremento en el desarrollo e investigación de herramientas CAD en diferentes áreas de la medicina, principalmente gracias a los avances en la capacidad de procesamiento y almacenamiento de información. Como consecuencia, actualmente, en lugar de usar microscopia convencional, cada vez más patólogos realizan diagnósticos interactuando con imágenes digitales [20].

A pesar de los avances en la investigación sobre CaGEBV, no se encuentra bibliografía de investigaciones que desarrollen un proceso de visión artificial para analizar microambientes tumorales de CG asociados a EBV (GC_TME_EBV+). No obstante existen investigaciones que han estudiado el (GC_TME_EBV+) utilizando herramientas de automatización y análisis de imágenes encontradas en el mercado, como '*Vectra 3 Automated Quantitative Pathology Imaging system*' [89]. Este proyecto de investigación es novedoso por el tipo de aplicación, y es por naturaleza altamente compatible con estudios que trabajaron preprocesamiento, morfología, detección, segmentación y clasificación de células tanto en imágenes histopatológicas como citológicas.

En el mundo, la investigación se ha enfocado en tejidos diferentes y en partes específicas del proceso de visión artificial. Entre los tipos de tejidos se incluyen análisis de imágenes de: cáncer de mama [90,91], tejido cardiovascular [92], folículos ováricos [93], hígado [94], pulmón [95] y próstata [96]. Entre las partes específicas del proceso de visión artificial se pueden contar: preprocesamiento de la imagen [97–99], detección automática y segmentación [28–30,100], extracción de características [101] y clasificación [102,103]. Gran parte de estos estudios se enfocan en el análisis de imágenes citológicas [1,104], sin embargo aquellos que se enfocan en las imágenes histopatológicas pueden ser rastreados a los trabajos de Weind *et al.* [105], Hamilton [106] y Bartels [107], quienes sentaron los precedentes para el análisis espacial de las estructuras presentes en este tipo de imágenes [104]. A continuación se presentan algunos de los trabajos más recientes y relevantes para esta investigación.

Normalización de color

Debido a la falta de estandarización de la tinción H&E entre diferentes laboratorios, la eficacia de los algoritmos de análisis de placas se ve muy reducida. En consecuencia Piórkowski y Gertych [97] presentan en su artículo un enfoque desde la normalización del color para segmentar núcleos en imágenes de tejido teñido por H&E. Buscaron entre los espacios de color, vectores que al transferirse a la imagen original H&E mediante normalización de Reinhard [108], mejoraran el rendimiento de un método de segmentación nuclear de referencia (basado en deconvolución de color y watershed controlado por marcadores en el espacio lab). Los autores analizaron las dependencias entre los colores rojo y azul en las imágenes normalizadas y encontraron que restando el canal azul al rojo, en las áreas donde el rojo domina sobre el azul, se pueden detectar los límites de los núcleos. Por lo tanto para encontrar el mejor vector de normalización probaron 100.000 vectores aleatorios sobre tres imágenes H&E aleatorias, aplicaron el método de segmentación de referencia para cada imagen normalizada y validaron los resultados al comparar con un “control verdadero” de la segmentación original corregida manualmente por médicos patólogos. La implementación de este método disminuyó la cantidad de píxeles falsos positivos y ayudó a diferenciar mejor núcleos agrupados, obteniendo una precisión del 94% en las imágenes seleccionadas.

Detección de núcleos

La detección de núcleos es crucial en análisis cuantitativos, en consecuencia Kowal *et al.* [109,110], propusieron un método basado en el marco de reconocimiento Bayesiano, que encuentra regiones esféricas con distribuciones de intensidad características de los núcleos aproximándose por generación de discos, en un proceso de marcación de puntos. Este método es una alternativa a los métodos usualmente usados con watershed para plantación de semillas, como umbralizado o gradiente de intensidad. En su trabajo, los autores plantearon el uso de la geometría estocástica para detectar semillas en los núcleos, en donde dada la matriz Y de intensidades de una imagen, como entrada al sistema, y una función X de características de discos (coordenadas del centro y radio), buscaron construir una función de probabilidad de masa dependiente de X y Y , la cual maximizaron tomando en cuenta restricciones sobre la interacción entre los discos, basadas en conocimiento *a priori* del tejido.

Kowal y Korbicz [109,110] utilizaron dos métodos para determinar la función de probabilidad de masa empírica, primero, basándose en el método de segmentación de Otsu, determinaron las regiones nucleares y utilizaron la operación morfológica de dilatación para determinar las regiones de fondo, con

un disco del tamaño del núcleo típico como elemento estructurante. El segundo método se basa en la segmentación manual de las regiones pertenecientes a núcleos y fondo para entrenar un sistema de inteligencia artificial. Para maximizar la función utilizaron un algoritmo determinístico basado en el método *Steepest Ascent*. Probaron la efectividad del método reconociendo núcleos en imágenes citológicas de tejido de cáncer de mama y obtuvieron 88,5% de núcleos detectados correctamente, 15,2% falsos positivos y 11,5% núcleos no detectados. En imágenes de tejido sano (93,2%, 10,4% y 6,8%) respectivamente.

Por otro lado, Höfener *et al.* [111] implementaron un sistema de detección de núcleos celulares utilizando aprendizaje profundo de máquina a través de la identificación de varios parámetros esenciales, con los cuales configuraron un mapa de probabilidades (PMap). El trabajo de Höfener *et al.* [111] se enfocó en la identificación de los parámetros de construcción de la red neuronal, que obtuvieron mejor valor-F (*F1-score*) en la detección de núcleos. Utilizaron dos bases de datos validadas por patólogos experimentados, para entrenar y probar redes neuronales completamente convolucionales (FCN), con diferentes parámetros. Una de las bases estaba compuesta de imágenes H&E de adenocarcinoma colorrectal y la otra de imágenes de tejido teñido de cáncer de mama; las cuales separaron manualmente, en imágenes de células individuales con anotaciones en los núcleos. Estas imágenes sirvieron para entrenar las FCN, de las cuales Höfener *et al.* [111] obtuvieron mapas de probabilidades. En el PMap, a cada pixel de entrada se le asignó un valor en representación de la probabilidad de pertenecer o estar cerca a un núcleo, posteriormente utilizaron una función de máximos locales para detectar cada núcleo individual. Escogieron como detecciones positivas aquellas que se encontraran dentro del radio promedio de las células estudiadas y que estuvieran verificadas por un patólogo. Como resultado de la mejor configuración de parámetros, obtuvieron un valor-F de 0,816 para las imágenes H&E, y 0,819 para las imágenes de cáncer de mama.

Segmentación

Singh *et al.* [112] desarrollaron un método para segmentar núcleos de células fuertemente aglomeradas, en donde fallan los métodos tradicionales. El método consiste en tres etapas, primero se clasifica el *foreground* o región celular y el fondo, segundo se detectan las semillas iniciales para el crecimiento de regiones y tercero se utiliza el crecimiento de regiones, en conjunto con distribuciones gaussianas, como información de entrada para la inicialización de un algoritmo 'level-set'.

Singh utilizó umbralización global y después umbralización local para refinar las secciones *foreground*, teniendo en cuenta que la ventana local debe ser un poco más grande que el núcleo para extraer eficientemente las células del fondo. Para el segundo paso, plantó las semillas manualmente y cortó la imagen para que la semilla inicial quede en la posición central. En el tercer paso parte de las semillas plantadas para crecer regiones con píxeles vecinos con propiedades similares, como regla utilizó la desviación estándar y la conexión en ocho. Singh basó los contornos iniciales del algoritmo de *level-set* en el resultado del crecimiento de regiones y en información de la distribución gaussiana para mejorar la precisión en la detección de núcleos. Singh afirma que su método tiene un rendimiento mucho mejor que las técnicas actuales, obtuvo una tasa de precisión media del 82,2% (media de detecciones correctas) y una tasa de precisión general del 81,3% (detección correcta/control verdadero), se compara con 62,3% & 67,1% [113] y 60,8% & 60,5% [114] respectivamente.

2.2.2 Referentes Nacionales

En Colombia la investigación de herramientas CAD en imágenes histopatológicas se ha dirigido principalmente a carcinoma de células basales [115–119], cáncer de mama [120,121] y próstata [121,122]. En cuanto al proceso de visión artificial, la gran mayoría estudia clasificación [123–126] precedido por preprocesado [127–130].

Normalización de color

Dorado *et al.* [128] desarrollaron un método de separación de color de imágenes de tejidos teñidos mediante H&E, proyectando linealmente la representación RGB en una superficie discriminante personalizada. En su trabajo Dorado *et al.* [128] proyecta el espacio RGB de una imagen de entrada a una familia de superficies que conectan las distribuciones de una serie de planos $[(R + B) \cdot 2 = B; G]$ que dividen la nube de tonos H&E. Utilizaron dicha proyección para agrupar tanto las muestras de hematoxilina como las de eosina, de donde fueron derivadas las bases de color. Posteriormente, para validar sus resultados umbralizaron las imágenes resultantes mediante el método de Otsu y compararon su rendimiento con dos bases de imágenes anotadas manualmente por patólogos. Obtuvieron una sensibilidad de detección del 71% y 67% en la primer y segunda base de imágenes respectivamente.

Detección de núcleos

Corredor *et al.* [131] trabajaron en la detección automática de linfocitos en imágenes de cáncer de pulmón, buscando obtener elementos cuantitativos sobre

la infiltración de linfocitos como contribución en la graduación del tumor. En el método propuesto, se parte directamente por una segmentación automática basada en watershed. Posteriormente se utilizan la forma del núcleo, textura y color para clasificar cada célula como linfocito o no linfocito, utilizando un clasificador entrenado de máquinas de vectores de soporte. Validaron los resultados utilizando una base de datos con 3.420 estructuras clasificadas (linfocitos y no-linfocitos) de 13 regiones de interés de imágenes digitalizadas de cáncer de pulmón. Encontraron además que la estrategia que presentaron tiene un valor-F 30% mayor que mediante la utilización de aprendizaje profundo.

Segmentación

Mejia *et al.* [132] propusieron un método de segmentación y clasificación de células cervicales con el fin de disminuir la naturaleza cualitativa del diagnóstico de lesiones neoplásicas. En este trabajo se parte de la interacción de un experto con la imagen, en la que el experto selecciona áreas de interés (ROI) que contengan una célula aislada. En la ROI el experto selecciona también el núcleo celular. Proceden aplicando un detector de bordes *Canny* para segmentar el núcleo, para segmentar el citoplasma, dividen la imagen en sus tres componentes RGB más la imagen en grises, aplican ecualizado del histograma a las cuatro capas y calculan la magnitud del gradiente, utilizan el promedio del gradiente de todas las capas para resaltar el borde del citoplasma. Para el caso de sobre-segmentación, utilizan un método iterativo de comparación de las regiones en la imagen hasta quedar solo con tres regiones, correspondientes a núcleo, citoplasma y fondo. Analizan 35 células, a las que les miden diversas características, que entregan a una máquina de soporte de vectores para seleccionar aquellas características que mejor representan las células seleccionadas. Para finalizar encontraron que las características más relevantes para la clasificación de células cervicales son los tamaños relativos, formas del núcleo y el citoplasma, dejando de lado características como textura y rugosidad.

2.2.3 Referentes en la Universidad del Cauca

En la Universidad del Cauca se ha trabajado en procesamiento de muestras asociadas a cáncer de mama [133], próstata [134], CG [16,17], entre otros. A continuación se muestran dos de los trabajos más relevantes en el contexto de esta investigación.

En el trabajo de grado presentado por Dorado y Escobar [133], se propone un método de visión artificial para determinar reactividad hormonal en tejido de cáncer de mama. Parten procesando muestras mediante la técnica de inmunohistoquímica, de donde proceden a tomar imágenes utilizando un

microscopio óptico de alta resolución con una cámara integrada de 2MP. Dorado y Escobar aplicaron a todas las imágenes realce adaptativo de contraste, corrimiento negativo de histograma en 20 valores de gris y un filtro de mediana 3x3. Después utilizaron un modelo de probabilidad de pertenencia de color, realizando algunas comparaciones entre los planos R,G,B diferencian células reactivas de no reactivas. Por otro lado para estandarizar su base de imágenes compararon el valor de asimetría de cada una, y propusieron un umbral de 90000, en caso de superado aplicaron una "look up table" (LUT) con una operación sigmoide, y en caso de no superarlo, una LUT arcotangente hiperbólica. Para el proceso de segmentación utilizaron el método de Otsu seguido de operaciones morfológicas de apertura y cierre, para finalizar aplicaron *watershed*. Validaron sus resultados comparando la cantidad de núcleos reactivos vs no reactivos encontrados por el sistema con las cantidades encontradas por el patólogo y obtuvieron un 90% de aciertos.

En otro trabajo de grado, Paredes [134] buscó determinar el grado gleason en biopsias diagnosticadas como adenocarcinoma de la próstata. En un primer paso, procesó tejidos cancerosos de próstata mediante la técnica de fijación con parafina, cortó los bloques parafinados y montó los tejidos en placas portaobjetos, a continuación rehidrató las muestras y aplicó tinción H&E. En un segundo paso, adquirió imágenes del tejido utilizando un microscopio óptico con cámara acoplada, sin embargo para cada toma fotográfica, varió los ajustes de los parámetros de la cámara. En un tercer paso, procesó las imágenes adquiridas a 400X para análisis citológico y las adquiridas a 100X para análisis estructural. Para el preprocesado citológico utilizó el espacio de color YCbCr, el cual ecualizó mediante modificación del histograma realizando ecuaciones booleanas entre los planos. Para el segmentado general de las células utilizó el método de Otsu y cambio de contraste controlado por el usuario, además para eliminar ruido, usó un filtro de mediana y cierre morfológico con *kernel* elíptico de tamaño del núcleo promedio. Utilizó interacción con el usuario como mecanismo de identificación de núcleos, que posteriormente realizó con una función gamma y segmentó por color en el espacio RGB. Para finalizar el análisis citológico, utilizó un algoritmo de extracción de componentes conectados y extrajo características que pasó a una máquina de soporte vectorial. Obtuvo una exactitud del 86.9% en determinación de núcleos hipercromáticos. Además cabe señalar que Paredes indicó en la sección de trabajos futuros, la necesidad de desarrollar un protocolo para la estandarización de la toma de imágenes histológicas y la necesidad de mejorar la etapa de preprocesado mediante clasificadores de color.

2.3 Análisis de la literatura

Las aplicaciones CAD en patología son muy diversas y cubren casi todos los tejidos, especialmente sanguíneos y mamarios, la mayoría de estas se centran en citología (análisis de células individuales) puesto que las imágenes citológicas pueden ser obtenidas fácilmente a través de exámenes poco invasivos, además las células en estas imágenes se encuentran naturalmente separadas y no se aprecian estructuras complicadas en el tejido. En contraste el análisis de imágenes histopatológicas (tejido que preserva su estructura original) presenta un desafío para la segmentación de las células, dado que, usualmente la mayoría de estas células hacen parte de complejas e irregulares estructuras histológicas.

Sin importar el tejido la segmentación de células está muy bien documentada, sin embargo cada aplicación particular sigue ofreciendo dificultades por la variación en las imágenes, causada principalmente por pequeñas diferencias en la preparación, como concentración de colorantes, lisura del corte, presencia de artefactos o daños en la muestra; problemas en la adquisición de la imagen, como pérdida de calidad por compresión, ruido digital y características específicas del sistema de adquisición, microscopio, aberraciones ópticas, etc.

Entre los diferentes tipos de células, aquellas de mayor interés son los linfocitos y las células epiteliales, que pueden lucir muy diferentes conforme a distintos factores como el tipo de núcleo, ciclo de vida y malignidad de la enfermedad, que además están usualmente organizadas en racimos superpuestos y poseen aspectos muy heterogéneos. Estas características hacen que la segmentación, detección y clasificación requieran de una solución robusta desde el procesamiento de imágenes que incluya el proceso de obtención de la muestra y conocimiento médico como etapas primordiales.

En la Tabla 2.1 se mencionan y discuten algunos de los métodos de procesamiento utilizados en la mayoría de los trabajos a nivel internacional, estos son recurrentes y aparecen en diferentes combinaciones según propósito y aplicación [1].

Tabla 2.1. Métodos recurrentes en procesamiento de imágenes histopatológicas, según la tendencia global. Elaboración propia, basado en Irshad H et. al. 2014 [1].

Método de procesamiento	Procedimiento
Umbralización	Otsu, umbralizado adaptativo local.
Análisis morfológico	Erosión, dilatación, apertura, cierre, <i>white and black top hat</i> y gradientes morfológicos.

Crecimiento de regiones	<i>Seed points</i> y clasificación de píxeles vecinos.
<i>Watershed</i>	Segmentación según el análisis topográfico de la imagen.
<i>Active contour models and level sets</i>	Segmentación al determinar curvas de nivel de funciones potenciales
<i>K means clustering</i>	División de la imagen en K conglomerados basándose en el valor del píxel, textura y localización.
Modelos probabilísticos	Modelado de la imagen a partir de su distribución de probabilidad.
<i>Graph cuts</i>	Algoritmos de segmentación que siguen la teoría de grafos [135].

Entre estos métodos, cabe destacar Otsu para umbralizar, pues resulta efectivo lidiando con la variación en el histograma, como lo muestran los diferentes trabajos expuestos, sin embargo al ser un método computacional, debe usarse con cautela, pues en el caso particular de imágenes que no poseen un histograma bimodal, se observan malas umbralizaciones. Por otro lado Mejia *et al.* [132] se enfocaron en segmentar basándose en la detección de bordes y gradientes, ese enfoque es interesante, aunque utilizar el color como método de segmentación genera resultados más consistentes.

Por otro lado, el método “*watershed*”, a pesar de ser uno de los métodos más utilizados, generalmente es muy difícil de aplicar debido a que genera sobre segmentación y depende en gran medida de los llamados ‘*seed points*’, marcadores reales muy diferenciados en la imagen, y que como se muestra en el trabajo de Singh [112], muchas veces son plantados manualmente, lo que en cierta forma destruye el propósito de agilizar el proceso diagnóstico; este acercamiento es más bien utilizado en investigación, donde se busca identificar características específicas de las células para su posterior clasificación, por lo que dedican más tiempo a plantar semillas individuales.

Los métodos de convolución de color y transferencia de color resultan muy efectivos y facilitan la segmentación, sobretodo en imágenes con tinción inmunohistoquímica y H&E debido a que el conjunto de vectores funcionales es bien conocido para estas imágenes, y ya que algunas imágenes de CISH no son muy diferentes a las inmunohistoquímicas, se facilita la aplicación de estos métodos.

Se encontró que la inclusión de segmentación manual, selección manual de áreas de interés y en general etapas interactivas en varios de los trabajos, demuestra

que el procesamiento digital de imágenes puede ser utilizado como una herramienta de ayuda no automatizada, dejando los procesos más complejos al ser humano. Además la inclusión de patólogos expertos en la verificación de la mayoría de los trabajos encontrados, demuestra que el conocimiento médico de base es fundamental y debe estar acompañando el proceso desde el inicio de cada proyecto, como norma general, la verificación de los resultados se logra al comparar el número de células detectadas por el programa con el número de células detectadas por el experto.

En Colombia, la investigación está claramente enfocada en clasificación, donde aparecen métodos recurrentes como “*bag of words*” y “*deep learning*”. A continuación se presenta en la Figura 2.8 la distribución, según enfoque, de los artículos encontrados a través de scopus, mediante la cadena de búsqueda: “TITLE-ABS-KEY (histopathology) AND (LIMIT-TO (AFFILCOUNTRY, “Colombia”)) AND (LIMIT-TO (SUBJAREA, “COMP”) OR LIMIT-TO (SUBJAREA, “MATH”) OR LIMIT-TO (SUBJAREA, “ENGI”) OR LIMIT-TO (SUBJAREA, “MATE”) OR LIMIT-TO (SUBJAREA, “PHYS”))”.

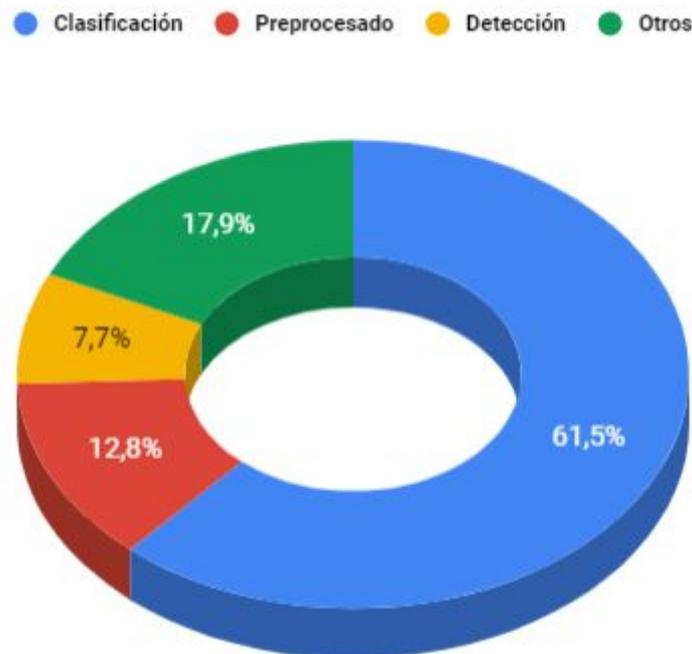


Figura 2.8. Enfoque de la investigación en procesamiento digital de muestras histopatológicas en Colombia. Se observa gran interés en la comunidad científica colombiana en desarrollo y prueba de algoritmos de clasificación.

Se identificaron los elementos más relevantes para esta investigación, sin embargo se sugiere referirse a los diferentes estudios de revisión bibliográfica como los que presentaron Gloria Bueno *et al.* [20], Gurcan Metin [104], Antoine Veillard [1] y Fuyong Xing [136].

3. METODOLOGÍA

El desarrollo metodológico de este trabajo, que se explica en forma de diagrama en la Figura 3.0, partió de la realización de una exhaustiva revisión bibliográfica (como se muestra en el marco referencial), con el objetivo de definir un plan de acción para procesar una colección de 32 bloques de tejido parafinado de pacientes caucanos con adenocarcinoma gástrico de tipo intestinal y dos placas control, obtenidos en el marco del proyecto con código 4487 de la vicerrectoría de Investigaciones “Asociación de los herpes virus Epstein barr y Citomegalovirus con el adenocarcinoma gástrico de tipo intestinal en pacientes del Departamento del Cauca”. Según el plan de acción identificado y en conjunto con el Grupo de investigación en Inmunología y Enfermedades Infecciosas (GYEi), se procedió con la preparación de los bloques de parafina que contenían las biopsias de CG (corte en micrótopo y montado en placa) y posterior tinción mediante la técnica de hibridación *in situ* cromogénica. Posteriormente se digitalizaron imágenes del tejido, que sirvieron de base para caracterizar el microambiente tumoral mediante procesamiento digital, utilizando una combinación de algoritmos, siguiendo el método de Otsu, normalización de color y enfoque watershed para preprocesar, detectar y segmentar las células teñidas. Para finalizar se validaron los resultados con ayuda de médicos patólogos.



Figura 3.0. Diagrama de la metodología. El proceso se basa en anotaciones del patólogo.

3.1 Procesamiento del tejido

3.1.1 Preparación de la biopsia

Para empezar, se organizó el tejido parafinado en microarreglos, para lo cual, un profesor y patólogo experimentado de la Universidad del Cauca, analizó todos los bloques de tejido parafinado y señaló controles, zonas tumorales y no tumorales en cada uno de ellos. Se utilizó una jeringa adaptada, para extraer bocados de las zonas marcadas, posteriormente se insertaron en bloques de parafina previamente preparados (Figura 3.1). Luego, se siguió la estructura presentada en la tabla 3.1, según número de paciente y tipo de zona (tumoral o no tumoral).

Tabla 3.1. Estructura de los microarreglos.

# Microarreglo	Caso #(i)		Caso #(i +1)
	--	No tumoral	No tumoral
	Control	Tumoral	Tumoral
	--	Tumoral	Tumoral



(a)



(b)



(c)



(d)



(e)



(f)

Figura 3.1. Desarrollo metodológico para el montaje de microarreglos de tejido parafinado. (a) Bloque de tejido parafinado; (b) Instrumento sacabocados; (c) Bocado de tejido en la punta del instrumento; (d) Bloque de tejido parafinado después de sacar los bocados; (e) Bloque de parafina que recibe los bocados; (f) Bloque microarreglo.

Cada bloque microarreglo fue sellado utilizando finas capas de parafina, además para asegurar los bocados en el nuevo bloque, se dejaron reposar 24 horas a cuatro grados celsius. Posteriormente, se entregaron los bloques al laboratorio de histología de la Universidad del Cauca, en donde, con la ayuda de un micrótopo, se realizaron cortes superficiales de 3 a 5 μm de espesor, que se montaron en placas portaobjetos cargadas positivamente.

3.1.2 Hibridación *in situ* cromogénica

Una vez en posesión de las placas portaobjetos con los cortes superficiales parafinados, se procedió a teñir el tejido mediante la técnica de Hibridación *in situ* cromogénica (CISH) siguiendo una variación del método utilizado por Kaizaki *et al.* [137] evidenciado en la Figura 3.3 y utilizando el kit comercial (SKU: 111) de BIOCARE MEDICAL® en el laboratorio de GYEi mostrado en la Figura 3.2.

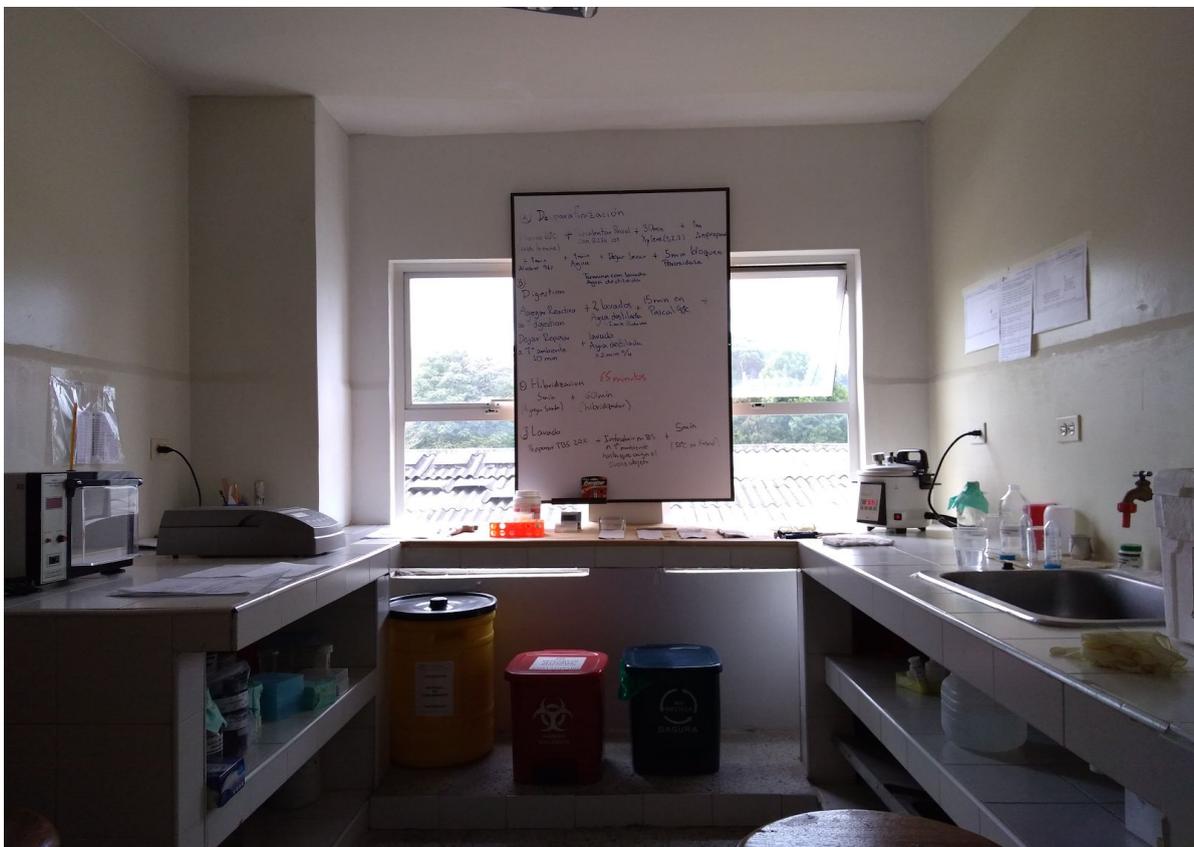
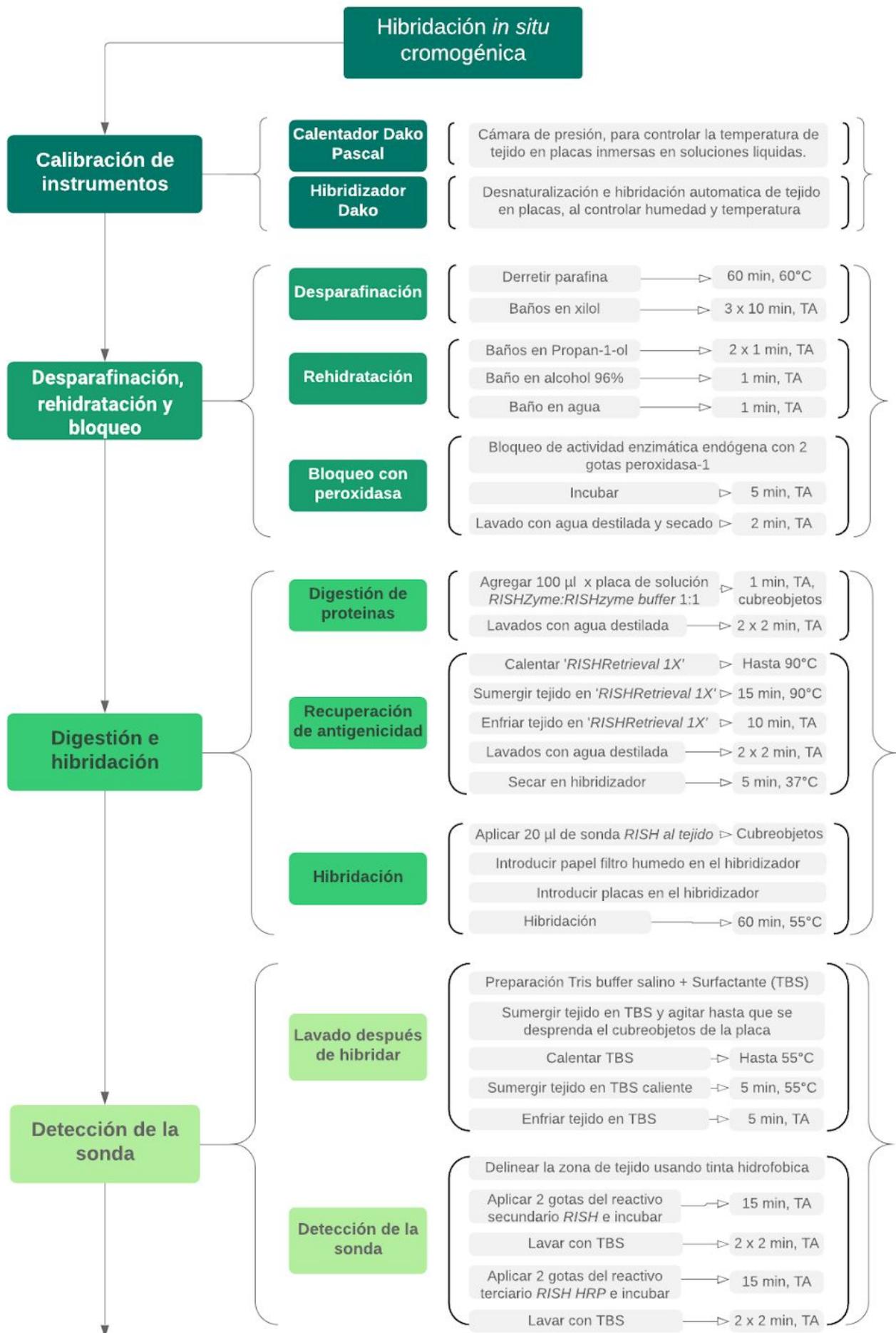


Figura 3.2. Laboratorio perteneciente al grupo de investigación en inmunología y enfermedades infecciosas de Facultad de Medicina de la Universidad del Cauca (GYEi). En este laboratorio se realizó la tinción por hibridación *in situ* cromogénica.



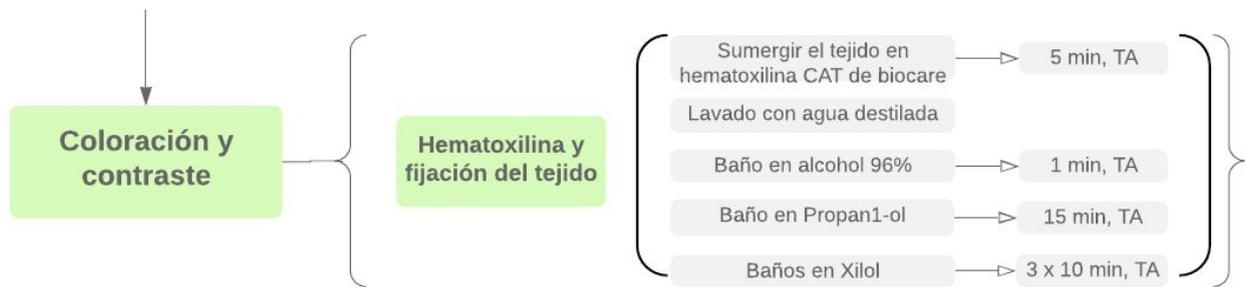


Figura 3.3. Mapa conceptual representando el procedimiento paso a paso de la tinción mediante hibridación in situ cromogénica. Elaboración propia, basado en el trabajo con -GYEi-

Calibración de instrumentos

Durante la tinción se utilizaron reactivos sensibles a los cambios de temperatura, por esto, para realizar el procedimiento, el proyecto adquirió en calidad de préstamo dos instrumentos facilitados por la Compañía de Patólogos del Cauca, un calentador *Dako Pascal* y un *Hibridizador Pascal*, (Figura 3.4). El primer instrumento se usó para controlar la temperatura de dos soluciones (*RISHRETRIEVAL 1X* y *Buffer TBS*) almacenadas en coplin de plástico, dentro de los cuales estaban las placas. El segundo instrumento se usó para realizar el proceso de hibridación al controlar la temperatura de una cámara. Sin embargo al no tener conocimiento del estado de los instrumentos en cuestión, se procedió a caracterizarlos mediante comparación directa de la temperatura y tiempo con una termocupla tipo T, siguiendo la calibración de fábrica. Una explicación más completa de la caracterización de los instrumentos se puede encontrar en el Anexo 1.

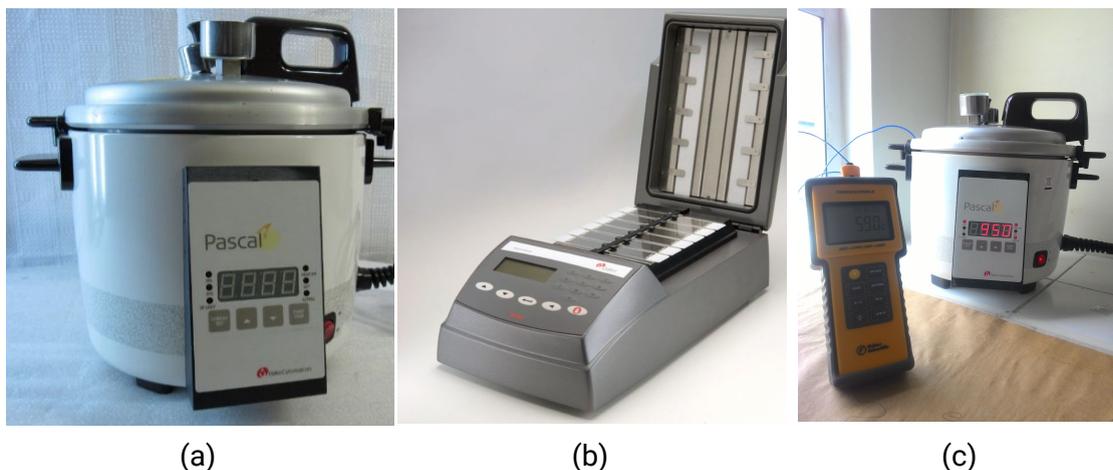


Figura 3.4. Instrumentos utilizados. (a)Calentador Dako Pascal; (b)Dako Hybridizer (Hibridizador); (c) Calibración.

Desparafinación, rehidratación y bloqueo

Para realizar la hibridación es necesario que las placas se encuentren desparafinadas, por lo que en primer lugar, se derritió la parafina de las mismas elevando su temperatura a 60°C durante 60 minutos en un horno. Seguidamente, se disolvió la parafina derretida sumergiéndolas en tres baños consecutivos de xilol por 10 minutos cada uno. A continuación, se rehidrató el tejido siguiendo tres pasos, primero, sumergiéndolo dos veces en *Propan-1-ol* durante un minuto cada vez, segundo, sumergiéndolo en alcohol al 96% durante un minuto, y tercero sumergiéndolo en agua durante un minuto. Después de esto, se dejó secar el tejido para posteriormente bloquear la actividad enzimática endógena del tejido, agregando dos gotas de *peroxidasa-1*, e incubando a temperatura ambiente por cinco minutos con el objetivo de evitar la reacción del sustrato por afinidad. Para finalizar se lavó con agua destilada y se dejó secar a temperatura ambiente.

Digestión e hibridación

Se preparó un reactivo de digestión, en solución 1:1, *RISHzyme™* : *RISHzyme™ Buffer*, utilizando 50 µL de enzima y 50 µL de buffer por cada placa. En seguida, a temperatura ambiente se agregó el reactivo preparado y se cubrió cada placa con un cubreobjetos, pasado un minuto, se lavó con agua destilada dos veces por un minuto cada vez. A continuación se mezcló *RISH™ Retrieval 10X* y agua desionizada, en solución 1:10, utilizando 5 y 45 mL respectivamente con lo que se obtuvo 50 mL de *RISH™ Retrieval 1X*.

Posteriormente se calentó la solución *RISH™ Retrieval 1X* en un coupling dentro del instrumento "*Dako Pascal*" durante 50 minutos, hasta que la solución alcanzó 90°C, luego, se verificó el pH de la solución (7,0) y se recuperó la antigenicidad de la muestra sumergiendo las placas en dicha solución, la cual se volvió a introducir en el instrumento *Pascal* durante 15 minutos más. Seguidamente se dejó enfriar por 10 minutos, se lavó con agua destilada por dos minutos y se dejó secar en el hibridizador a 37°C. Una vez estuvieron secas, se aplicaron 20 µL de sonda *RISH™ EBER DNA* a cada placa y se cubrieron usando cubreobjetos nuevos de 22 x 22 mm. Posteriormente se introdujeron al hibridizador, que realizó hibridación a 55°C durante 60 minutos.

Detección de la sonda.

Se preparó una solución de Tris-buffer-salino + surfactante (TBS), con 500mM Tris, 60mM KCl, 2.8M NaCl, 1.0% Tween-20 en disolución por agua desionizada con ajuste de pH a 7,4.

Terminada la hibridación y con el objetivo de remover residuos de la sonda en el tejido, se lavaron las placas con la solución TBS, a temperatura ambiente, agitando hasta haber removido los cubreobjetos. Simultáneamente se calentó un coupling con TBS en *Dako Pascal* durante 24 minutos, hasta que la solución alcanzó 55°C. Luego, se sumergieron las placas en la solución caliente, la cual se volvió a introducir en *Dako Pascal* por cinco minutos más. Para terminar el lavado, se escurrieron las placas y se dejaron secar en el hibridizador a 37°C.

Para detectar la sonda, se debe marcar las cadenas de la sonda primaria con un agente cromogénico, para esto se agregaron dos gotas del reactivo *RISH™ Secondary Reagent* por cada placa y se incubó en una cámara húmeda por 15 minutos a temperatura ambiente. Después, se hicieron dos lavados con TBS de 2 minutos cada uno y se dejó secar. Posteriormente se agregaron dos gotas del reactivo *RISH™ HRP Tertiary Reagent*, repitiendo el proceso anterior de incubación, lavado y secado. Para terminar se aplicaron 50 µL de reactivo *Betazoid* (32 µL de cromógeno Diaminobenzidina [DAB] por cada 1,0 ml de tampón/buffer de sustrato DAB) por cada placa, se incubó por cinco minutos en una cámara húmeda a temperatura ambiente y se lavó con agua destilada.

Coloración de contraste y deshidratación.

Acto seguido, se sumergieron las placas en CAT hematoxilina durante cinco minutos, lo que le dio coloración de contraste al fondo del tejido, inmediatamente después se lavaron con agua destilada para remover el exceso de hematoxilina. Posteriormente se deshidrataron las placas, siguiendo el proceso inverso a la rehidratación, pasando primero por un baño de un minuto en alcohol al 96%, dos baños de un minuto en *Propan-1-ol* y tres baños de 10 minutos en xilol. Finalmente, se aplicó una gota de medio de montaje permanente (Entellan® 107961) en cada cubreobjetos y se almacenaron las placas.

3.2 Definición de las características del microambiente

En esta etapa del proceso, se realizaron reuniones con un patólogo experto en la detección de EBV para determinar las características más importantes a identificar en las placas. En dichas reuniones, el patólogo ubicó y marcó las secciones más importantes del tejido en varias placas con ayuda de un microscopio óptico, indicando las características a encontrar para identificar las células infectadas y las células no infectadas.

3.3 Adquisición de imágenes para el desarrollo de la herramienta

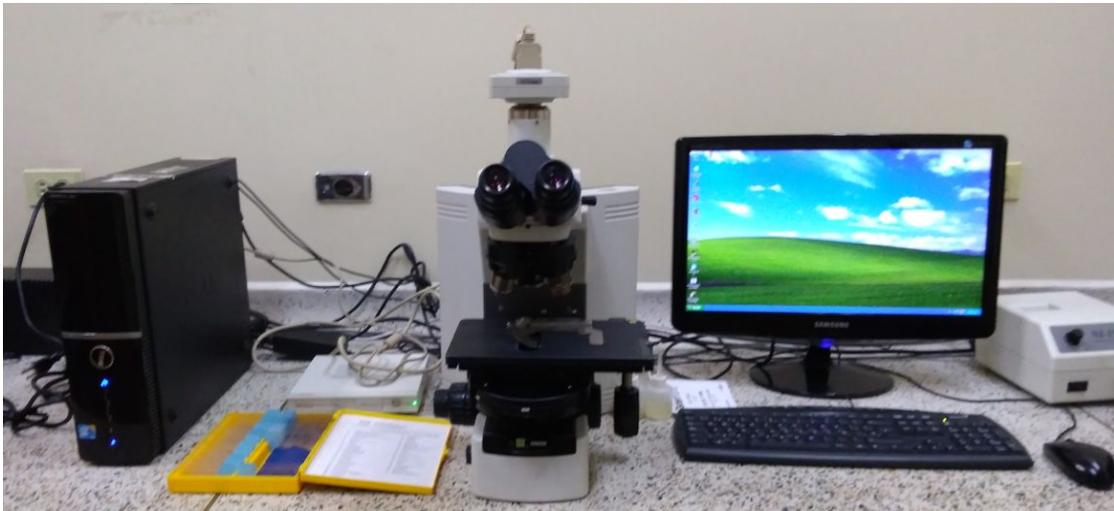


Figura 3.5 Microscopio utilizado. Marca Nikon Eclipse 80i, con cámara Nikon DS-MV-U2.

En esta etapa, se observó el tejido en las placas y se adquirieron imágenes con el objetivo de determinar un conjunto de parámetros, valores y *settings* de toma óptimos para empezar a desarrollar el sistema de visión artificial.

Para adquirir dichas imágenes se utilizó un microscopio óptico de alta resolución. Marca NIKON, modelo Eclipse 80i, con cámara digital integrada de 2 Mpx, marca NIKON, modelo DS-MV-U2 y software NIS-Element Sf (Figura 3.5). Que fue prestado al proyecto por la Unidad de Microscopía Electrónica de la Universidad del Cauca.

3.3.1 Parámetros de adquisición

Con ayuda de un patólogo, se marcaron zonas de interés en las placas, y se anotaron características generales para la correcta identificación de la tinción por CISH. Posteriormente, basándose en dichas anotaciones, se identificó una configuración óptima del sistema de adquisición de imágenes mediante el método experimental, controlando deliberadamente los siguientes parámetros: magnificación, cámara, iluminación, velocidad de obturación, balance de blancos y contraste. En el método experimental aplicado, se controló primero la magnificación y se escogió el tipo de cámara a utilizar, después, se controlaron los demás parámetros.

Identificación de la cámara a utilizar y magnificación de la imagen

Tabla 3.2. Clases de cámaras utilizadas y sus respectivos montajes en el microscopio.

#	Modelo de cámara	Montaje	Resolución
1	NIKON, DS-MV-U2	Acoplado interno al microscopio	2 Mpx
2	NIKON, COOLPIX S6300	Trípode, fotografía del ocular	16 Mpx
3	Smartphone ASUS ZENFONE 4	Fotografía directa del ocular	13 Mpx

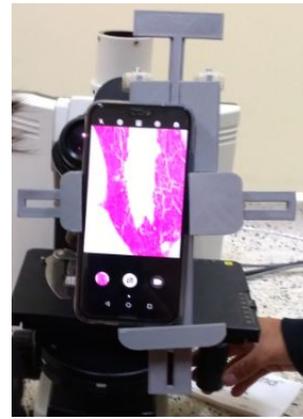


Figura 3.6. Cámara #2, NIKON, COOLPIX S6300. Se observa en montaje con trípode fotografiando directamente al ocular del microscopio.

Para comenzar, se observó por el microscopio la zona de interés utilizando un ocular de 10X y cuatro objetivos diferentes (4X, 10X, 40X y 100X). Se identificó el mejor objetivo basándose en la proporción entre tamaño de la célula y área visual del tejido. A continuación, se utilizaron tres tipos de cámaras y montajes (Tabla 3.2, Figura 3.6, Figura 3.7) para tomar fotografías de una regla patrón, de 1 mm de largo, dividida en 100 partes iguales, de tal forma que entre cada división se pudieran medir exactamente 10 μm . Este patrón de calibración sirvió para medir la cantidad de píxeles que representaban 50 μm en diferentes localidades de la imagen, empezando por el centro y alejándose de forma radial, con el fin de identificar aberraciones morfológicas causadas por el paso de la luz a través de los diferentes lentes, comparando la diferencia entre los valores extremos y el valor central.



(a)



(b)

Figura 3.7. Cámaras # 1 y 3. (a) Cámara NIKON, DS-MV-U2 acoplada al microscopio; (b) Smartphone ASUS ZENFONE 4 en montaje comercial para microscopio, zoom digital en la imagen mostrada en pantalla.

Posteriormente, para evaluar heterogeneidad en la iluminación generada por desalineaciones de la cámara con el lente del microscopio, se tomaron 10 fotografías de zonas sin tejido o *background* homogéneas en su composición, es decir, que deberían presentar un color sólido en la imagen. Las fotografías se tomaron en distintos lugares de una misma placa, con el fin de reducir los efectos de posibles artefactos (e.g., polvo) en la misma. Además, el valor de iluminación se cambió para cada toma siguiendo una función aleatoria, de tal modo que se desvinculó el efecto de un valor de iluminación específico en la heterogeneidad de la iluminación. Este procedimiento se realizó para cada una de las tres cámaras presentadas en la Tabla 3.2.

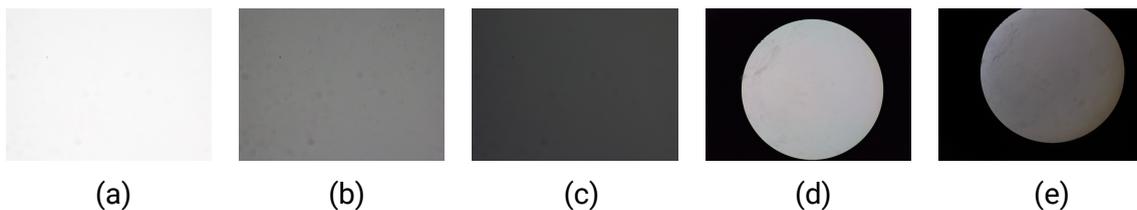


Figura 3.8. Ejemplo de fotografías de background. (a-c) NIKON DS-MV-U2; (d) Smartphone ASUS ZENFONE 4; (e) NIKON COOLPIX S6300. Se requiere que el background sea uniforme

Como se observa en la Figura 3.8, las fotografías tomadas con las cámaras 2 y 3 incluyen la sombra del ocular, para remover este artefacto, se procesaron todas las fotografías tomadas con estas cámaras mediante un algoritmo de umbralización llevando la zona de sombra a valores de cero y la zona *background* a 255. Posteriormente se aplicó un algoritmo de optimización para encontrar automáticamente las coordenadas del rectángulo de mayor área inscrito en la elipse. Para finalizar se recortó la imagen original en las coordenadas encontradas para remover por completo la sombra del objetivo, perder la mínima

información posible de la imagen y conservar la facilidad de procesamiento de una imagen rectangular (Figura 3.9).

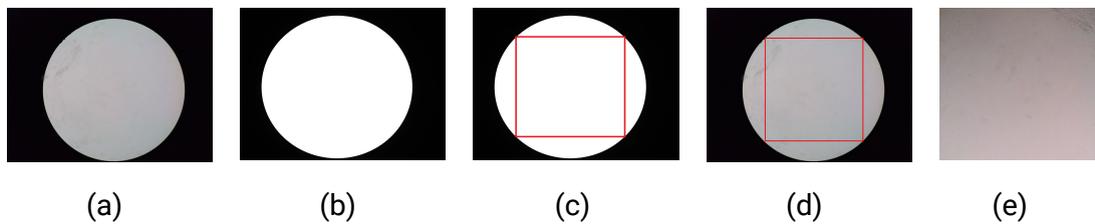


Figura 3.9. Proceso de remoción de la sombra del objetivo. (a) Imagen original de las cámaras # 2 o 3. (b) Segmentado de la sección elíptica que corresponde al campo de visión del microscopio, dejando por fuera la sombra del ocular. (c) Obtención del rectángulo de mayor área inscrito en la elipse segmentada. (d) Extracción de las coordenadas de los vértices del rectángulo inscrito. (e) Recorte de la imagen original en las coordenadas encontradas.

Más adelante, para determinar la heterogeneidad en la iluminación de una imagen, se utilizó la desviación estándar del valor de luminosidad (*Lightness*) de los píxeles como una medida representativa que indicó qué tan dispersos estaban los datos con respecto a la luminosidad media. Con el fin de encontrar un valor que definiera la dispersión de los valores de luminosidad se utilizó el Algoritmo 1, presentado a continuación:

1. Convertir la imagen de entrada del espacio de color BGR¹ (*Blue, Green, Red*) al espacio de color HLS (*Hue, Lightness, Saturation*)
2. Dividir la imagen en sus capas constituyentes, H, L, S.
3. Desechar las capas H y S, manteniendo la capa L.
4. Calcular la desviación estándar de cada fila de la imagen y registrar cada valor en un vector llamado H. Hacer lo mismo para las columnas y registrar en un vector llamado V.
5. Crear un vector llamado PH, donde se registrará en una posición el promedio de los valores del vector H. De igual manera crear el vector PV.
6. Repetir los pasos 1 a 4 para cada una de las 10 imágenes, llenando los vectores PH y PV, con los promedios de los valores de los vectores H y V.
7. Calcular el promedio del vector PH y registrar en una variable llamada DPH.
8. Calcular el promedio del vector PV y registrar en una variable llamada DPV.

Algoritmo 1. Algoritmo para identificar la dispersión de los valores de luminosidad de los fondos.

Este algoritmo se ejecutó tres veces, una vez por cada set de imágenes de cámaras diferentes y entregó como resultado la dispersión de los valores de luminosidad en la horizontal (DPH) y en la vertical (DPV) para cada cámara. Se

¹ En Python, el orden de los espacios de color es diferente al usual. RGB => BGR, HSL => HLS.

escogió la cámara que presentó menor aberración morfológica y menores valores DPH/DPV.

Identificación de la configuración ideal de la cámara para obtener las mejores imágenes preliminares.

Para comenzar, se partió de la configuración base utilizada en la unidad de microscopía para el uso del microscopio *Nikon Eclipse 80i* : Iluminación total y apertura total del diafragma, después, basándose en observaciones de un patólogo experto sobre la apariencia de la imagen, se optó por utilizar 6 puntos en la escala de iluminación del microscopio y 6 puntos en la escala de apertura del diafragma (Tabla 3.3). Con esta configuración del microscopio se tomaron 50 fotografías con la cámara seleccionada en el punto anterior, a tres zonas de interés, obteniendo imágenes en todas las posibles combinaciones de los valores considerados. Se separaron las imágenes en tres carpetas, cada una correspondiente a una zona de interés y se guardaron los detalles de la configuración en la metadata de cada imagen.

Posteriormente, con el fin de identificar las mejores imágenes preliminares, se depuraron las imágenes por su apariencia, mediante un proceso subjetivo basado en anotaciones del patólogo, que identificó aquellas imágenes en las que mejor se pudiera reconocer las células teñidas de café (infectadas por EBV) de otras células y del fondo.

Tabla 3.3. Valores considerados para cada parámetro de la cámara.

Velocidad de obturación	Contraste	Balance de blancos
1/60s	Bajo	Automático
1/90s	Medio	Sin balance de blancos
1/125s	Alto	-----
1/180s	Enhanced	-----
1/250s	Dinámico	-----

3.4 Procesamiento digital de imágenes

La herramienta que se desarrolló, tuvo por objetivo facilitar el conteo de células infectadas por EBV, así como cuantificar marcaciones inespecíficas, es decir marcaciones no nucleares y cuantificar la cantidad de tejido sano. Para lograr esto se siguió el esquema presentado en la Figura 3.10.

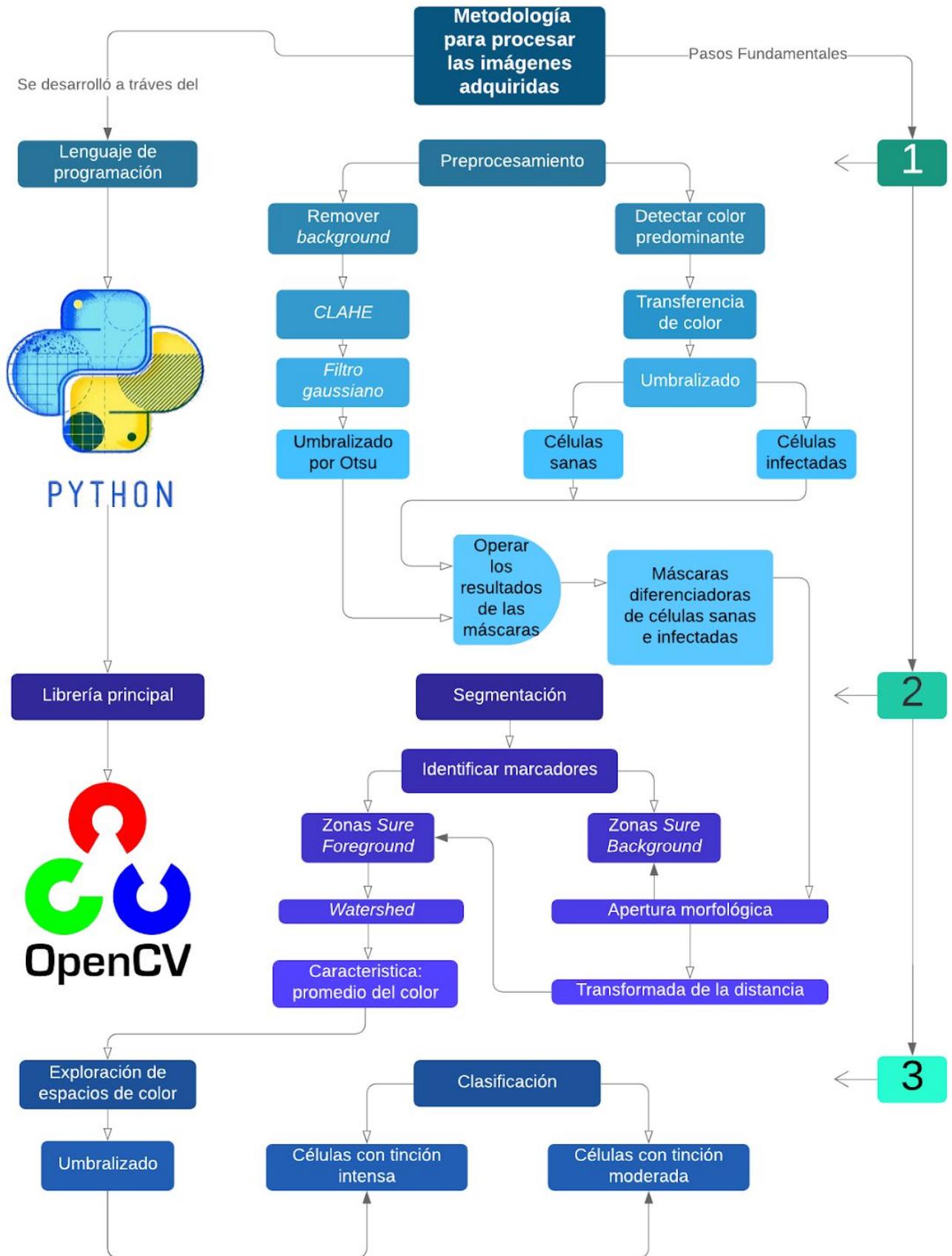


Figura 3.10. Esquema de la metodología utilizada para el procesamiento de muestras. Se ilustra la la secuencia de algoritmos utilizados para tareas específicas en el procesamiento de imágenes de GC_TME basándose en la identificación del virus de epstein barr. Elaboración propia.

Se escogió Python como el lenguaje de programación idóneo para desarrollar la herramienta debido a su robustez, a la eficacia de sus librerías para el manejo de imágenes y su naturaleza como software de código abierto, ya que la utilización y modificación de este software no requiere de licencias pagas ni derechos restrictivos que impidan la libre comercialización o utilización de los programas escritos en este lenguaje.

3.4.1 Preprocesamiento

Para iniciar, se utilizaron las imágenes seleccionadas en el punto anterior como entradas al sistema, para representar las imágenes en su forma matricial se utilizó la función de lectura de imagen de la librería *OpenCV*, obteniendo matrices con tres planos de profundidad en el espacio de color BGR (azul, verde y rojo), además se utilizó la librería *numpy* que permitió realizar operaciones matemáticas a las matrices adquiridas.

A continuación, se realizaron dos procesos paralelos. Aplicación del método de otsu para remover el fondo de la imagen y utilización de una variación de la transferencia de color de Reinhard para detectar tanto marcaciones de color café (positivo para EBV), como células sanas en las imágenes de la biopsia. Posteriormente se combinaron los resultados de estos dos procesos mediante la operación lógica AND (*bitwise*).

Remoción del Fondo de la Imagen

En primer lugar se incrementó el contraste mediante la técnica *Contrast Limited Adaptive Histogram Equalization (CLAHE)*, consistente en ecualizar separadamente histogramas de pequeñas ventanas de la imagen, que al reunirse forman la imagen original ecualizada. En segundo lugar, se pasó la imagen por un filtro gaussiano, que hace uso de una función del mismo nombre para difuminar la imagen, removiendo el ruido efectivamente. Para finalizar, se utilizó el método de Otsu para umbralizar, es decir separar el fondo (*background*) y el área conformada por células (*foreground*).

El método de otsu tiene por objetivo hallar un valor de umbralización para el que la suma de la dispersión de los pixeles a cada lado del umbral sea mínima. Para esto, Otsu itera sobre todos los posibles valores de umbralizado y calcula, para cada valor umbral una medida de la dispersión (varianza ponderada) de los niveles de gris de los pixeles a cada lado del umbral. En concreto Otsu basa su método en funciones de densidad de intensidades de una imagen con intensidades en el rango $[0, L-1]$ separadas por el valor umbral k , para *background*

'Cb' y *foreground* 'Cf'. Se pueden apreciar más detalles del método de Otsu en el Algoritmo 2.

Método de umbralización de Otsu

1. Denotar las intensidades de una imagen MxN, con número de píxeles n_i por :
 $\{0,1,2,3,4,\dots, L-1\}$

2. Componentes del histograma normalizado :

$$p_i = \frac{n_i}{M \cdot N}, \quad \sum_{i=0}^{L-1} p_i = 1, \quad p_i \geq 0.$$

3. Funciones de densidad de intensidades, en dos clases Cb y Cf :

Background Cb, intensidades en el rango $[0,k]$

Foreground Cf, intensidades en el rango $[k+1,L-1]$

4. Probabilidades de las clases Cb, Cf :

$$P_{Cb}(k) = \sum_{i=0}^k p_i \quad P_{Cf}(k) = \sum_{i=k+1}^{L-1} p_i$$

P_i es la probabilidad de la intensidad i en la imagen

5. Intensidad media de las clases Cb, Cf :

$$m_G = \sum_{i=0}^{L-1} i \cdot p_i = P_{Cb}(k) \cdot m_1(k) + P_{Cf}(k) \cdot m_2(k)$$

6. Intensidad media de la imagen :

$$m_G = \sum_{i=0}^{L-1} i \cdot p_i = P_{Cb}(k) \cdot m_1(k) + P_{Cf}(k) \cdot m_2(k)$$

7. Varianza entre clases

$$\sigma_B^2(k) = P_{Cb}(k) \cdot [m_1(k) - m_G]^2 + P_{Cf}(k) \cdot [m_2(k) - m_G]^2$$

8. El valor de k es seleccionado por búsqueda secuencial :

$$k^* = \text{Max}_{0 \leq k \leq L-1} \{\sigma_B^2(k)\}$$

Algoritmo 2. Método de umbralización de Otsu.

Detección de células marcadas

Se utilizó una variación del método de transferencia de color de Reinhard [108], el cual consiste en utilizar el espacio de color L^*a^*b , que tiene canales descorrelacionados de color, para imponer características cromáticas de una imagen en otra, en otras palabras, transferir algunos aspectos de la distribución de los píxeles en el espacio L^*a^*b a otra imagen. Primero se explica el método original y a continuación, el método modificado.

Reinhard consiguió transferir los colores entre imágenes restando a cada pixel del canal L^*a^*b de la imagen objetivo, la media del canal correspondiente de la misma imagen y escalando los canales de la imagen objetivo por la proporción de la

desviación estándar del objetivo sobre la desviación estándar de la fuente, multiplicado por los canales objetivo. Para finalizar se agregó la media de los canales de la imagen fuente a los canales de la imagen objetivo, respectivamente. Reinhard convirtió la imagen resultante de nuevo al plano BGR para su despliegue (Figura 3.11).

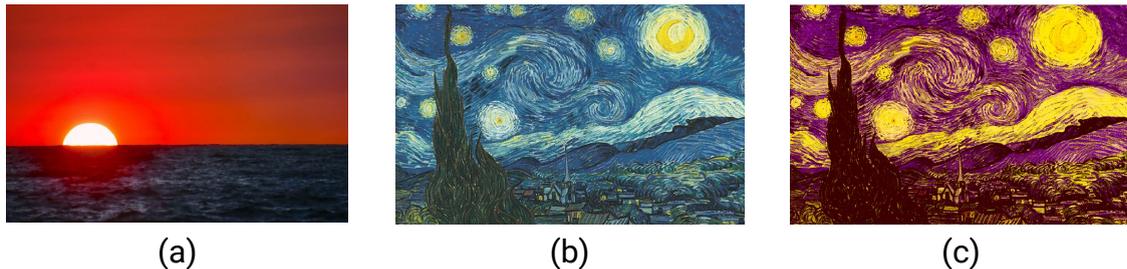


Figura 3.11 Ejemplo de aplicación de transferencia de color de Reinhard. (a) Imagen fuente; (b) Imagen objetivo; (c) Imagen objetivo con los colores de la imagen fuente.

Primer paso	Tercer paso	Quinto paso
Pasar imagenes a: $L*a*b$	$LTar = (LStdTar / LStdSrc) * LTar$	Recombinar $LTar, aTar, bTar$
Imagen Objetivo = $LTar, aTar, bTar$	$aTar = (aStdTar / aStdSrc) * aTar$	Pasar resultado a BGR
Imagen Fuente = $LSrc, aSrc, bSrc$	$bTar = (bStdTar / bStdSrc) * bTar$	Convenciones
Segundo paso	Cuarto paso	Mean = Obtener la media
$LTar = LTar - LMeanTar$	$LTar = LTar + LMeanSrc$	Std = Desviación Estándar
$aTar = aTar - aMeanTar$	$aTar = aTar + LMeanSrc$	Tar = <i>Target</i> (Objetivo)
$bTar = bTar - bMeanTar$	$bTar = bTar + LMeanSrc$	Src = <i>Source</i> (Fuente)

Algoritmo 3. Algoritmo de transferencia de color de Reinhard [108].

Primer paso	Tercer paso	Quinto paso
Pasar imagen objetivo a: $L*a*b$	$LTar = (LStdTar) * LTar$	Recombinar $LTar, aTar, bTar$
Imagen Objetivo = $LTar, aTar, bTar$	$aTar = (aStdTar) * aTar$	Pasar resultado a BGR
	$bTar = (bStdTar) * bTar$	
Segundo paso	Cuarto Paso	Convenciones
$LTar = LTar - LMeanTar$	If $LTar > 0$, $LTar = 255$, else : $LTar = 0$	If => Condicional
$aTar = aTar - aMeanTar$	If $aTar > 0$, $aTar = 255$, else : $aTar = 0$	else => En caso contrario
$bTar = bTar - bMeanTar$	If $bTar > 0$, $bTar = 255$, else : $bTar = 0$	

Algoritmo 4. Algoritmo Reinhard modificado. Se utilizó para el preprocesamiento de las imágenes objetivo adquiridas en la sección 3.2.

Las modificaciones hechas al algoritmo original (Algoritmo 3) consistieron en: no utilizar una imagen fuente, en vez de eso, en el tercer paso, reemplazar los valores de desviación estándar de los canales de la imagen fuente $StdSrc$ por la unidad (1). Segundo, agregar un paso extra después del tercero, para llevar a cero los valores de negativos de $LTar, aTar$ y $bTar$, y llevar a 255 los valores positivos. Tercero, eliminar el cuarto paso. Para finalizar, se repite el quinto paso,

recombinar las capas L_{Tar} , a_{Tar} , b_{Tar} en una imagen L^*a^*b que después se convierte al espacio BGR para su despliegue. Como se observa en el Algoritmo 4.

El algoritmo 4 se aplicó para obtener mayor contraste de las partes de la imagen (marcación positiva, negativa y background), es decir, concentrar los rangos de valores de los píxeles pertenecientes a cada parte. Posteriormente, se umbralizó la imagen resultante con un filtro pasa bandas utilizando un rango de valores calculado a partir de la comparación del histograma de la imagen completa y de secciones segmentadas manualmente pertenecientes a cada tipo de zona.

Combinación de resultados

A continuación para obtener una máscara de las componentes de la imagen positivas para EBV se realizó la operación lógica AND entre la imagen binaria sin fondo y la imagen binaria con las células marcadas.

3.4.2 Segmentación

Para segmentar se siguió el esquema de aplicación de *watershed* controlado por marcadores, partiendo de la imagen binaria obtenida en la sección 3.4.1.c.

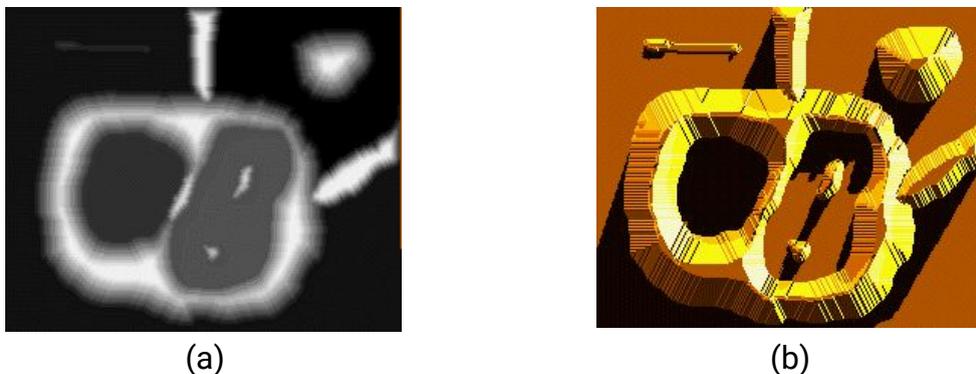


Figura 3.12. Ejemplo de topografía de una imagen. (a) Niveles de gris de una imagen; (b) Representación topográfica. Se representa la topografía en función de la intensidad.

El algoritmo de *watershed* trata a la imagen como una superficie topográfica, como se observa en la Figura 3.12, donde los niveles de intensidad representan montañas, el nivel cero de intensidad representa al nivel del mar, que se usa para calcular la altura de las montañas. De ahí se empiezan a llenar los valles con “agua” hasta que las aguas de diferentes fuentes se toquen. Las paredes de las represas representan los bordes de los objetos y cada lagos separado representa un componente segmentado de la imagen.

En el esquema controlado por marcadores, estos representan fuentes de agua ubicadas estratégicamente dentro de cada objeto para lidiar con la diferencia de altura y sobre segmentación. Primero, se deben obtener los marcadores en la

imagen, para ello se crean zonas *sure foreground*, zonas que se sabe que pertenecen al objeto a segmentar y zonas *sure background*, zonas que se sabe que no pertenecen al objeto a segmentar, después se hallan las zonas desconocidas, restando *sure foreground* a *sure background*. En la Figura 3.13 se puede observar un ejemplo de segmentación por *watershed* controlado por marcadores, las líneas rojas delimitan cada objeto segmentado.

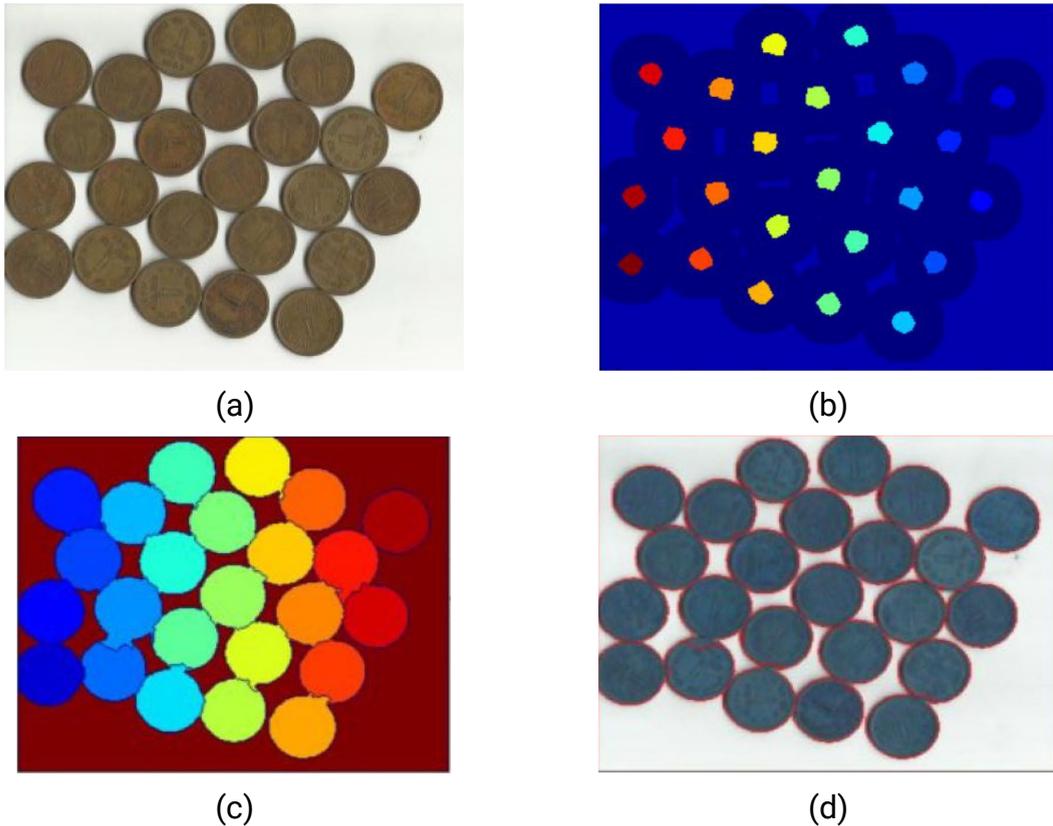


Figura 3.13. Ejemplo de aplicación de Watershed. (a) Imagen original; (b) Designación de zonas *sure foreground* en colores, *sure background*, en azul claro y zonas desconocidas en azul oscuro; (c) Segmentación por *watershed*; (d) Imagen original en escala de grises segmentada por *watershed*. Tomado de la documentación de *openCV* [138].

Una vez identificadas estas zonas, se deben separar los marcadores por componentes conectados, incluyendo además a todo el *sure background* como un componente único (así no esté conectado entre sí), después se puede entregar el mapa de componentes conectados al algoritmo de *watershed*, para que utilice dichas regiones como partida y pueda expandir estas por las zonas desconocidas, logrando segmentar la imagen.

En este orden de ideas, para identificar el *sure background* se aplicó apertura morfológica utilizando una elipse del tamaño del núcleo promedio como elemento estructurante, después se aplicó dilatación morfológica. Para identificar el *sure foreground*, se aplicó la transformada de la distancia al resultado de la apertura

morfológica y se umbralizó la imagen resultante con un valor del 10% del máximo de la distancia. Una vez identificados los marcadores, se siguió el procedimiento explicado en el párrafo anterior.

3.4.3 Clasificación de marcaciones fuertes y leves

Para identificar diferencias cuantificables entre las células con marcación fuerte y (leve/moderada) se analizaron los niveles de gris para dos grupos seleccionados de células, uno para cada tipo de marcación. Los grupos se encuentran constituidos por células previamente segmentadas y el análisis se realizó en diferentes espacios de color.

Una vez identificado el espacio de color más adecuado se calculó el promedio aritmético del valor de los píxeles pertenecientes a cada célula mediante la aplicación de máscaras específicas para cada célula identificada. Finalmente se escogieron umbrales para bandas que se aplicó al valor del promedio de las células, basándose en la media y desviación estándar del vector de promedios calculados para cada tipo de tinción, con lo que se determinaron marcaciones moderadas y fuertes.

3.5 Digitalización de las placas

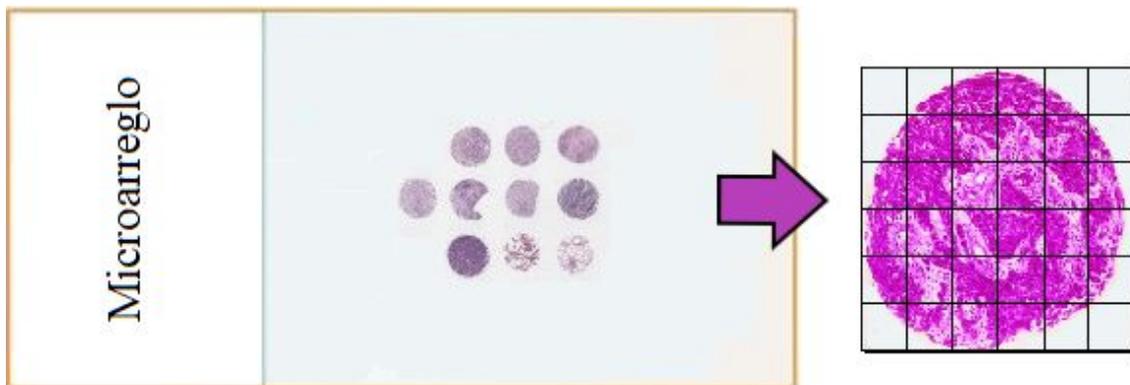


Figura 3.14. Representación del proceso de digitalización de los microarreglos. Cada punto del microarreglo se registra en una carpeta independiente y cada cuadro de la cuadrilla representa una imagen de dicha carpeta. *Elaboración propia.*

Una vez desarrollada la estructura funcional de la herramienta, se realizaron tres pruebas a la base de imágenes adquiridas en la sección 3.2. En la primera prueba se sobrepuso la máscara de marcaciones sobre la imagen original, para observar marcaciones no reconocidas por la herramienta. En la segunda prueba se hizo una comparación visual de las marcaciones específicas detectadas contra marcaciones inespecíficas. En la tercera prueba se corrió la herramienta sobre las imágenes para comparar el número de marcaciones detectadas. Finalmente, la

realización de estas pruebas determinó los parámetros y condiciones óptimas para la captura de las imágenes a procesar.

Utilizando los parámetros determinados, se procedió a digitalizar las placas recorriendo cada muestra en el microarreglo en forma secuencial, avanzando de izquierda a derecha y en sentido descendente, ajustando el enfoque para cada toma. Se registró cada muestra del microarreglo en carpetas independientes, junto con las coordenadas correspondientes en el microarreglo. En la Figura 3.14 se puede observar cómo se digitaliza un punto del microarreglo.

3.6 Caracterización del TME basándose en la identificación de EBV.

Según Bonilla *et al.* [139], desde una perspectiva investigativa, una caracterización es un tipo de descripción cualitativa que recurre a lo cuantitativo con el fin de profundizar el conocimiento sobre algo. En este caso se buscó caracterizar el microambiente tumoral a través de la cuantificación de células infectadas con el virus Epstein Barr (EBV).

Partiendo de las imágenes segmentadas, para cada imagen adquirida se calculó el número de células infectadas, el número de células no infectadas, el número de células con tinción intensa, el número de células con tinción moderada y los correspondientes porcentajes de cada tipo de célula frente al total de células detectadas, estos cuantificadores se muestran en la Tabla 3.4.

Tabla 3.4. Cuantificadores de la identificación del EBV en el microambiente tumoral.

#	CUANTIFICADOR
(a)	Número de células no marcadas.
(b)	Número de células marcadas.
(c)	Número de células con tinción intensa.
(d)	Número de células con tinción moderada.
(e)	Porcentaje de células no marcadas del total de células identificadas.
(f)	Porcentaje de células marcadas del total de células identificadas.
(g)	Porcentaje de células con tinción intensa del número de células marcadas.
(h)	Porcentaje de células con tinción moderada del número de células marcadas.
(i)	Porcentaje de células con tinción intensa del total de células identificadas.
(j)	Porcentaje de células con tinción moderada del total de células identificadas.

Posteriormente, se calcularon los valores mostrados en la Tabla 3.4 para cada imagen perteneciente a un mismo punto del microarreglo y se organizaron en matrices. Después, para cuantificar las células infectadas con el EBV en todo el punto del microarreglo, se calculó la suma de cada columna de valores y a partir de estos valores se calcularon los porcentajes. Este proceso se repitió para cada punto de tejido en los microarreglos.

3.7 Presentación de la herramienta

A partir de los resultados obtenidos en las secciones anteriores y a partir de reuniones con un patólogo, se determinaron las necesidades del usuario final y se determinaron las imágenes y resultados a desplegar y de qué manera. Para desarrollar la interfaz de la herramienta se utilizó el modulo Tkinter de Python.

3.8 Evaluación

Para evaluar la herramienta desarrollada, se compararon los resultados obtenidos por la herramienta con los resultados obtenidos por patólogos de forma tradicional. Para obtener los últimos se aplicó una encuesta voluntaria (Tabla 3.5) de forma individual a varios patólogos, en donde, teniendo en cuenta su poca disponibilidad de tiempo, se les pidió que a partir de una imagen de referencia se hiciera la cuenta del número de células totales, el número de células presentando marcación positiva y de estas contar el número de marcaciones intensas y el número de marcaciones moderadas o leves.

Tabla 3.5. Primera encuesta a patólogos.

Realice los siguientes conteos en las fotografías seleccionadas.			
Células totales	Células marcadas de café	Células marcadas con tinción intensa	Células con tinción moderada/leve

Tabla 3.6. Segunda encuesta a patólogos (encuesta tipo Likert)².

A las siguientes preguntas responder conforme con la siguiente escala.				
Muy en desacuerdo	En desacuerdo	Ni de acuerdo ni en desacuerdo	De acuerdo	Muy de acuerdo
1. Creo que herramientas de este tipo me benefician de alguna manera.				
2. Creo que esta herramienta es fácil de usar.				
3. Creo que esta herramienta es útil para identificar células marcadas de café.				
4. Utilizaría esta herramienta.				

Finalmente, para medir la satisfacción del gremio de patólogos respecto a la herramienta desarrollada, se socializaron los resultados de la investigación, se aplicó una segunda encuesta de tipo Likert (Tabla 3.6) y se realizaron conversatorios con varios patólogos.

² Likert es una de las maneras más fiables de medir opiniones y percepciones [140,141], en esta se ofrece un rango de respuestas de 5 o 7 puntos para medir satisfacción.

4. RESULTADOS Y ANÁLISIS

4.1 Procesamiento del tejido

Como resultado del procesamiento del tejido se obtuvieron placas portaobjetos con cortes superficiales de tejido gástrico teñido mediante hibridación *in situ* cromogénica y hematoxilina, como se muestra en la Figura 4.2. Se observa que la apariencia macroscópica de los microarreglos dista de la ideal, resultado de complicaciones en el proceso de corte de los bloques de microarreglos (Figura 3.1.f y Figura 4.1). Como consecuencia, algunos puntos del microarreglo se desprendieron de la placa en los procesos de lavado y se observó poca uniformidad del corte en la mayoría de los puntos.

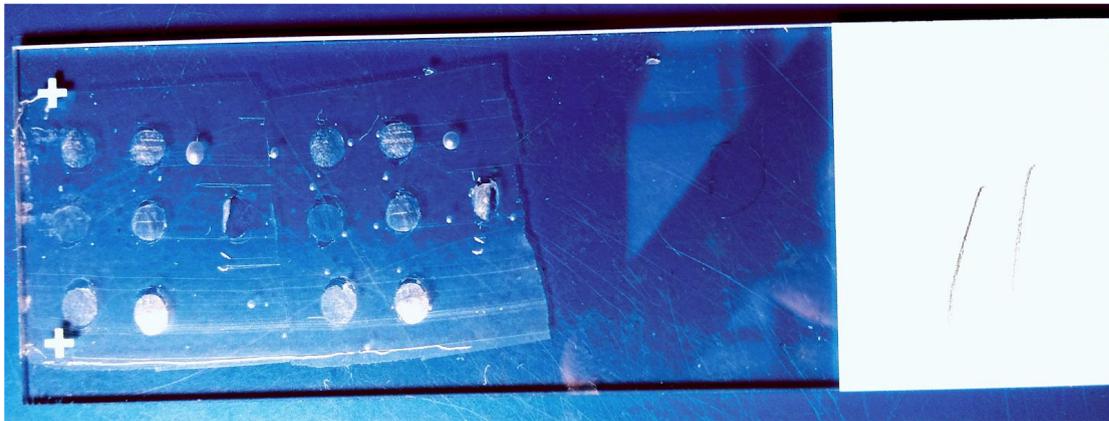


Figura 4.1. Corte superficial de tejido gástrico parafinado montado en placa portaobjetos. Observar la pérdida de algunos puntos del microarreglo, así como la desalineación de la estructura original planteada para el microarreglo.

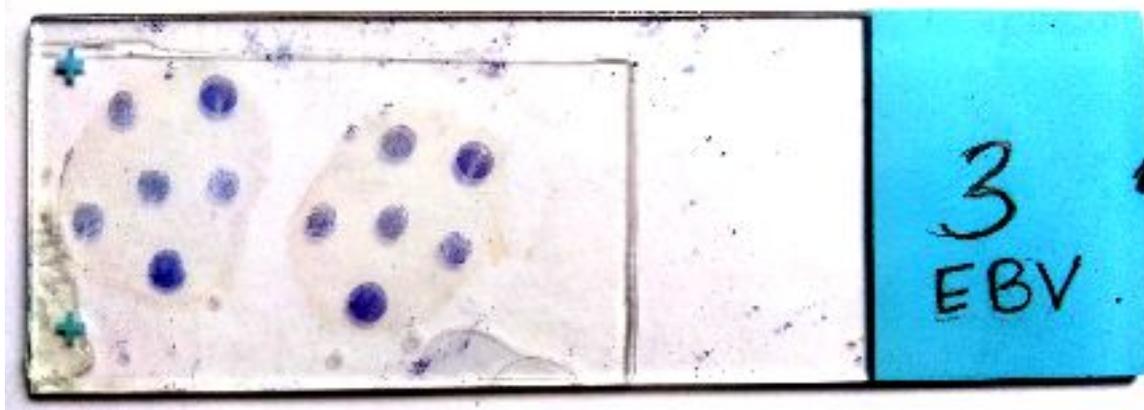


Figura 4.2. Corte superficial de tejido gástrico teñido por hibridación *in situ* cromogénica y hematoxilina montado en placa portaobjetos y fijado con cubreobjetos. Prestar atención a la coloración morada que adquirió la muestra al ser teñida con hematoxilina, es importante resaltar que la tinción específica por hibridación *in situ* solo es visible con ayuda del microscopio.

A pesar de las complicaciones en el procesamiento del tejido, se observa en la Figura 4.3 que la apariencia microscópica de los puntos del microarreglo es aceptable, además se logró eludir la poca uniformidad del corte ajustando manualmente el enfoque del microscopio para cada sección de la placa (práctica usual en patología), con lo que se pudo observar exitosamente la mayoría del tejido procesado. Cabe resaltar que, como se aprecia en la Figura 4.3, distinguir marcaciones positivas en bajos aumentos es muy complicado, notándose poca diferencia entre 4.3 (a) y 4.3 (b), por lo que normalmente los patólogos utilizan imágenes con mayor magnificación como la 4.3 (c) y 4.3 (d) en donde se puede diferenciar células cafés y células azules/moradas.

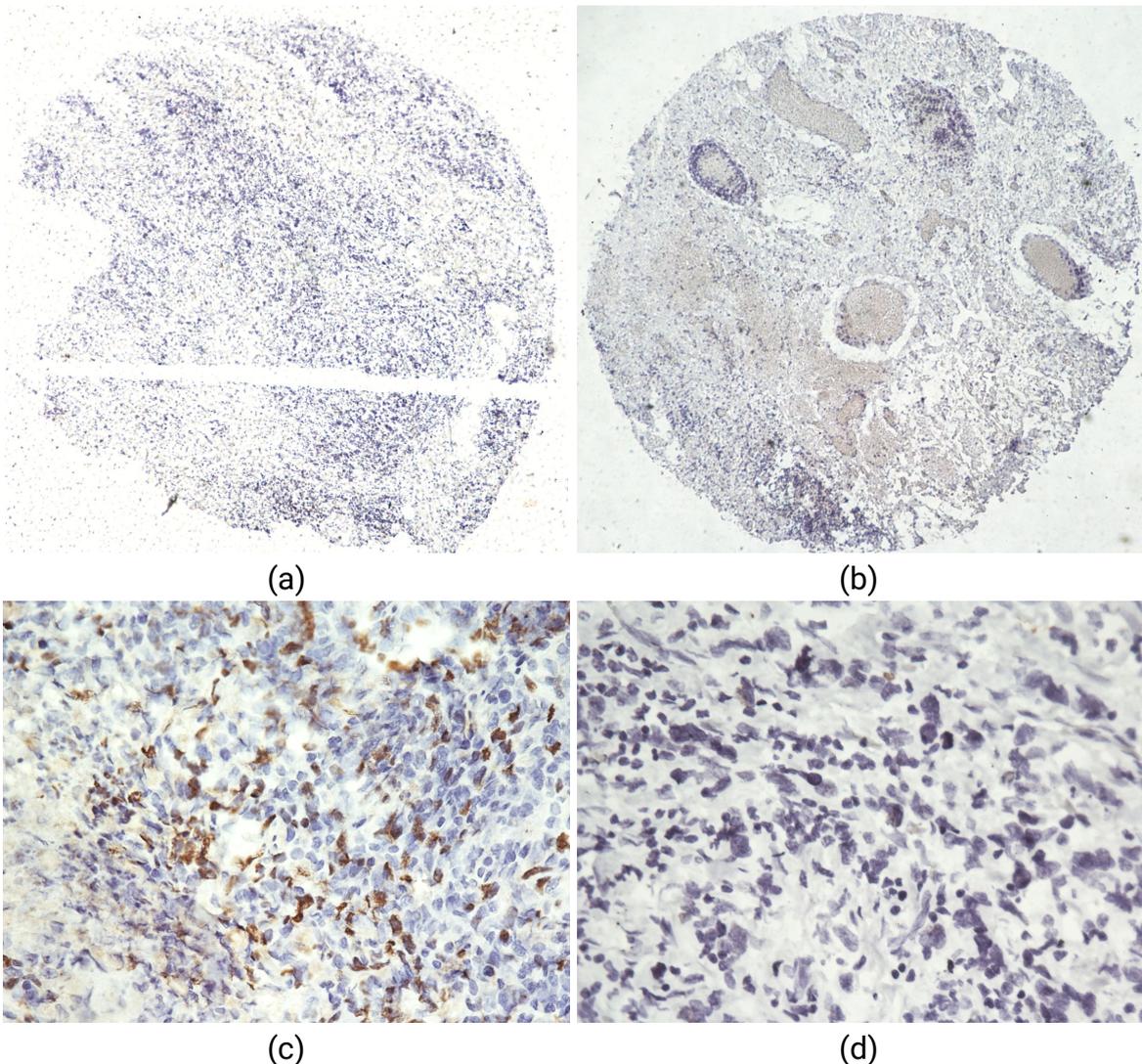


Figura 4.3. Puntos del microarreglo de la Figura 4.2 bajo el microscopio en diferentes magnificaciones. (a & b) magnificación 40X; (c & d) magnificación 400X. Nótese la apariencia del tejido en cuestión bajo la lente del microscopio; las células generalmente se encuentran apiñadas, agrupadas de forma estrecha, lo que impide diferenciar una célula de otra de forma trivial. Comúnmente se utiliza 40X para el estudio de las estructuras del tejido y 400X para el estudio de las células.

4.2 Definición de las características del microambiente

Se identificaron cuatro estructuras base para la caracterización del microambiente tumoral basándose en la tinción para EBV: células no infectadas, células infectadas, células con tinción intensa y fondo o background (Figura 4.4). A partir de la caracterización de estas cuatro estructuras, el patólogo expresó la necesidad de identificarlas y cuantificarlas a través de una herramienta software.

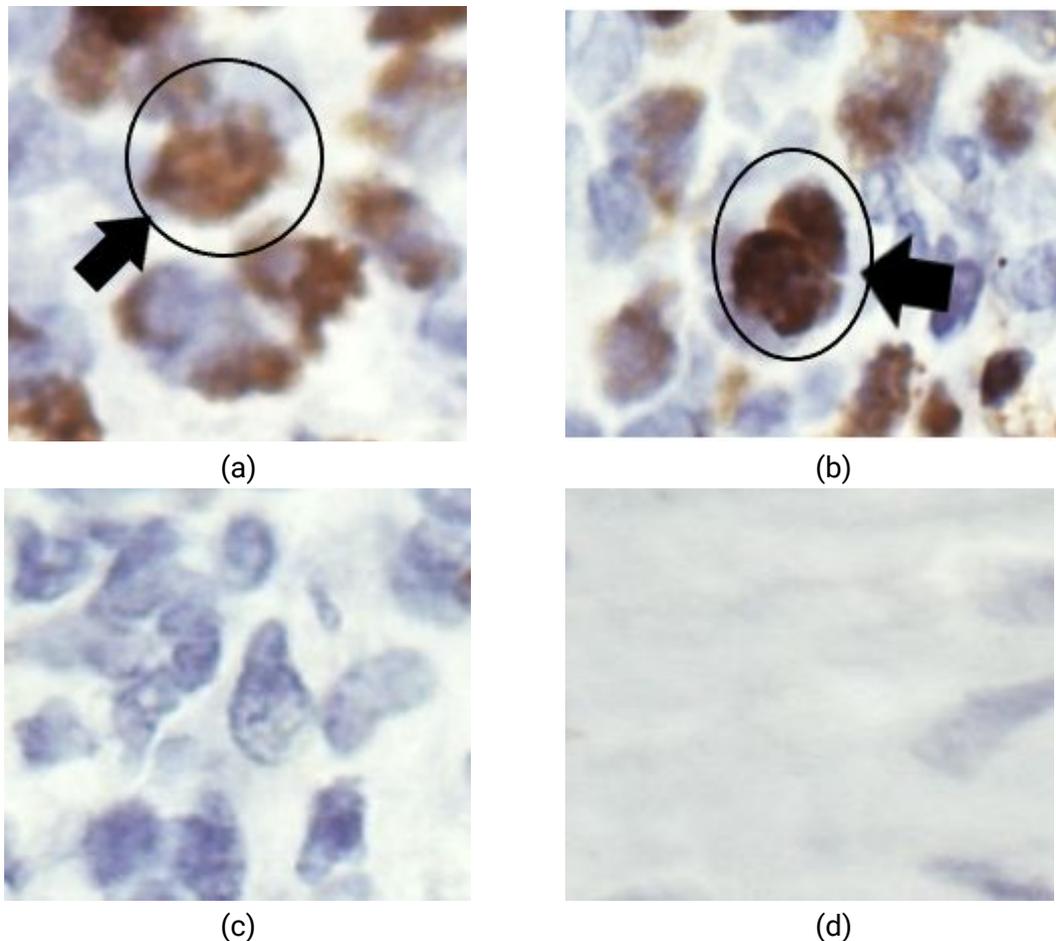


Figura 4.4. Estructuras que sirven para identificar la presencia del virus Epstein barr en el microambiente tumoral basándose en anotaciones del patólogo. (a) *Célula infectada, es principalmente café;* (b) *Célula fuertemente infectada, se observa una coloración más oscura que en la célula infectada;* (c) *Célula no infectada, es principalmente azul;* (d) *Fondo, de apariencia difusa y morada.*

Como se explicó anteriormente, el EBV infecta a células B y a células epiteliales por igual, que por naturaleza se presentan en conglomeración y tienen una alta variabilidad natural de características, por lo que se consideró que separar estos dos tipos de células representaría en sí mismo otro trabajo de investigación, por lo tanto la herramienta elaborada no distingue entre estas células en el momento de calcular el número total de células infectadas por EBV.

4.3 Adquisición de imágenes para el desarrollo de la herramienta

4.3.1 Mejor magnificación

Tabla 4.1. Proporción entre tamaños de células del microambiente y área visual del tejido.

Rango de áreas de las células presentes en el microambiente ($78.53 \mu\text{m}^2$ - $425 \mu\text{m}^2$) [142]		
Magnificación del objetivo (Ocular 10x/20 mm)	Área visible aproximada $\pi*(\text{Field of view})^2$	(Área célula/Área visible)*100 (límite inferior - límite superior)
4X	$19634954 \mu\text{m}^2$	$3.99 \times 10^{-4} \%$ - $2.16 \times 10^{-3} \%$
10X	$3141593 \mu\text{m}^2$	$2.49 \times 10^{-3} \%$ - 0.0135%
40X	$196350 \mu\text{m}^2$	0.0399% - 0.2160%
100X	$31415 \mu\text{m}^2$	0.2490% - 1.3528%

Se escogió utilizar la combinación de ocular 10X y objetivo de 40X, ya que con la magnificación 400X, se encontró la mejor proporción entre tamaños de células y área visible, como se muestra en la Tabla 4.1. Con la utilización de los objetivos de 4X y 10X la proporción es demasiado baja, por lo que no es posible apreciar células individuales, sin embargo el uso de estos objetivos es ideal para ubicar la muestra y hacer análisis estructural del tejido respectivamente. Por otro lado aunque el objetivo de 100X permite ver cada célula individual y en detalle, tiene una proporción muy alta y no permite analizar suficientes células para hacer un adecuado barrido del tejido.

4.3.2 Mejor cámara

Tabla 4.2. Comparación de la medida de $50 \mu\text{m}$ de referencia en pixeles en el centro de la imagen (100%) y en los extremos para las diferentes cámaras.

115%	115%	115%		101%	101%	101%		105%	105%	105%
115%	100%	115%		101%	100%	101%		105%	100%	105%
115%	115%	115%		101%	101%	101%		105%	105%	105%
NIKON COOLPIX S6300				NIKON DS-MV-U2				ASUS, ZENFONE 4		

Tabla 4.3. Dispersión de los valores de luminosidad en imágenes de background tomadas con cada una de las tres cámaras.

Promedio de desviaciones en la horizontal	6.8913	2.2374	14.2008
Promedio de desviaciones en la vertical	7.2547	2.1855	17.3574
(a) COOLPIX S6300; (b) DS-MV-U2; (c) ZENFONE4	(a)	(b)	(c)

Se determinó que la cámara NIKON DS-MV-U2 presentó menor aberración morfológica y dispersión de valores de luminosidad (Tablas 4.2 y 4.3). Por defecto, ésta cámara se encuentra acoplada internamente al microscopio, de tal

forma que en esta configuración el arreglo óptico es más preciso que en cualquier montaje que se pueda hacer en el ocular y a diferencia de las otras dos cámaras, posee un campo de visión más angosto, que limita el efecto de las aberraciones morfológicas. A pesar de tener una menor resolución (2Mpx), la calidad de las imágenes se consideró superior. Por lo tanto se escogió la cámara NIKON DS-MV-U2 para adquirir las imágenes a procesar.

4.3.3 Configuración de la cámara para obtener las imágenes preliminares.

Se seleccionaron los valores de configuración mostrados en la Tabla 4.4 para obtener imágenes con apariencias similares a las de la Figura 4.5. Cabe resaltar que se identificó 1/60s como la mejor velocidad de obturación en combinación con la iluminación del microscopio de seis puntos propuesta en la metodología. Estas configuraciones preliminares se usaron de base para diseñar el algoritmo de identificación.

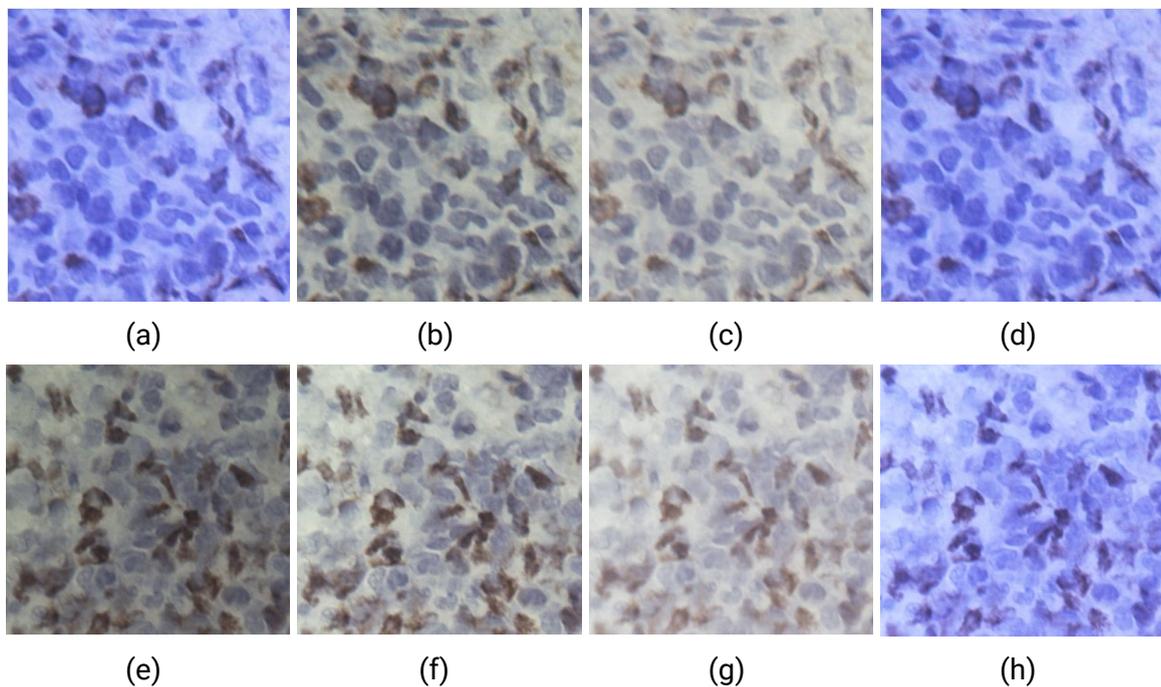


Figura 4.5. Ejemplo de imágenes preliminares. Mejores imágenes seleccionadas según diferenciación visual de las células marcadas. Se muestran los valores utilizados para configuración de la cámara en la tabla 4.4.

Tabla 4.4. Valores de ajustes de configuración de cámara seleccionados.

Letra asignada	Vel. obturación	Contraste	Balace de blancos
(a)	1/60s	HIGH	NO
(b)&(f)	1/60S	ENHANCED	SÍ
(c) & (g)	1/60S	HIGH	SÍ
(d) & (h)	1/60s	ENHANCED	NO
(e)	AUTO	ENHANCED	SÍ

4.4 Procesamiento digital de imágenes

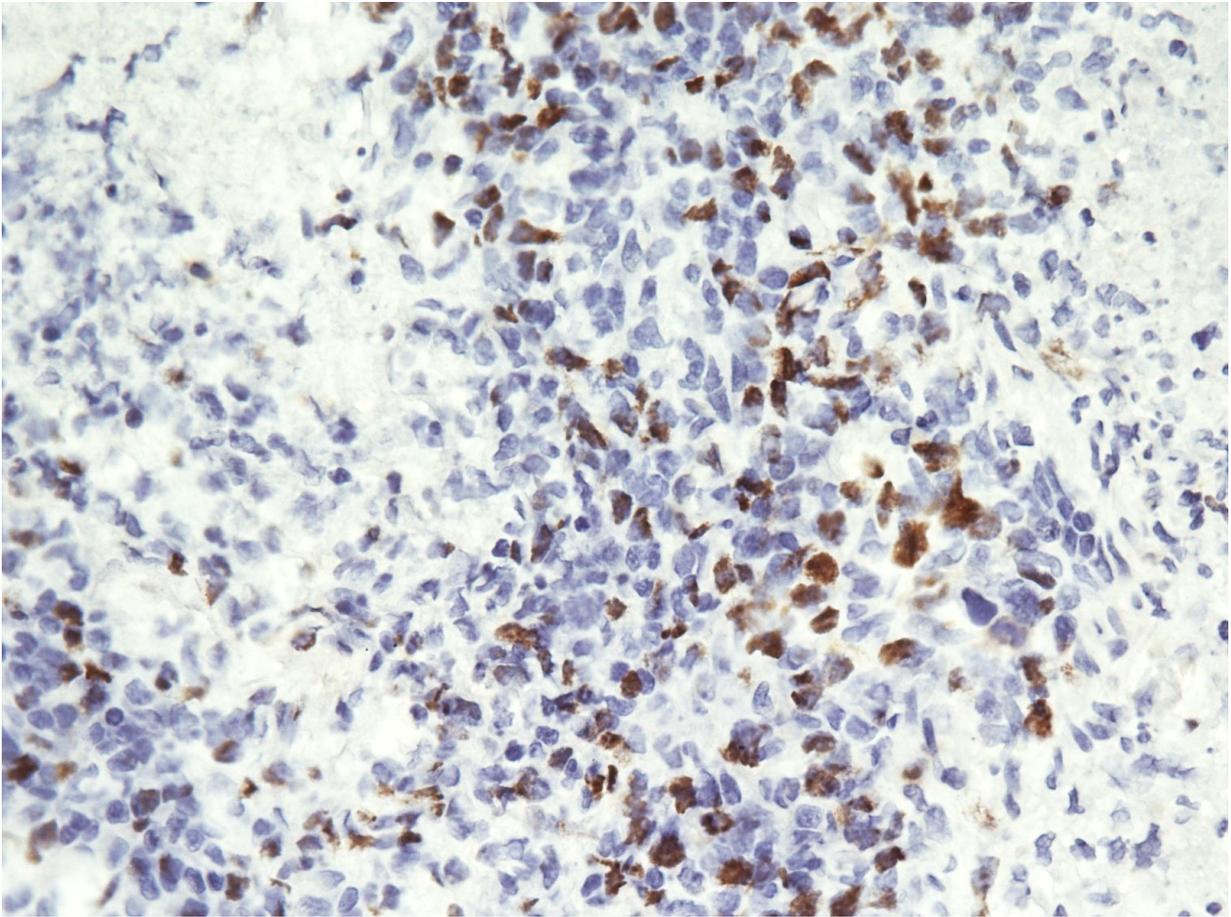


Figura 4.6. Ejemplo de una de las imágenes adquiridas de tejido gástrico canceroso teñido por CISH positivo para EBV. El software obtuvo resultados consistentes en una variedad de imágenes, sin embargo esta imagen se va a utilizar consistentemente para presentar los resultados obtenidos, con el objetivo de guiar al lector, que entienda el proceso y no se pierda al cambiar de imagen. Esta imagen representativa fue seleccionada por el patólogo por presentar un corte limpio, por tener una buena distribución de células marcadas moderada e intensamente y células no marcadas.

4.4.1 Preprocesamiento

En la Figura 4.7 se observa como siguiendo la metodología propuesta se logró remover exitosamente el fondo de las imágenes adquiridas, como la que se muestra de ejemplo en la Figura 4.6. El uso de CLAHE resultó efectivo para resaltar las células sin alterar el fondo y posteriormente el filtro gaussiano ayudó a difuminar la imagen facilitando la aplicación del método de umbralizado de Otsu. Cabe resaltar que la utilización del método de Otsu permitió a la herramienta remover el fondo de una gran cantidad de imágenes diferentes, incluyendo imágenes pertenecientes a diferentes zonas de tejido conectivo e imágenes tomadas bajo condiciones diferentes a las determinadas mediante este proyecto de investigación.

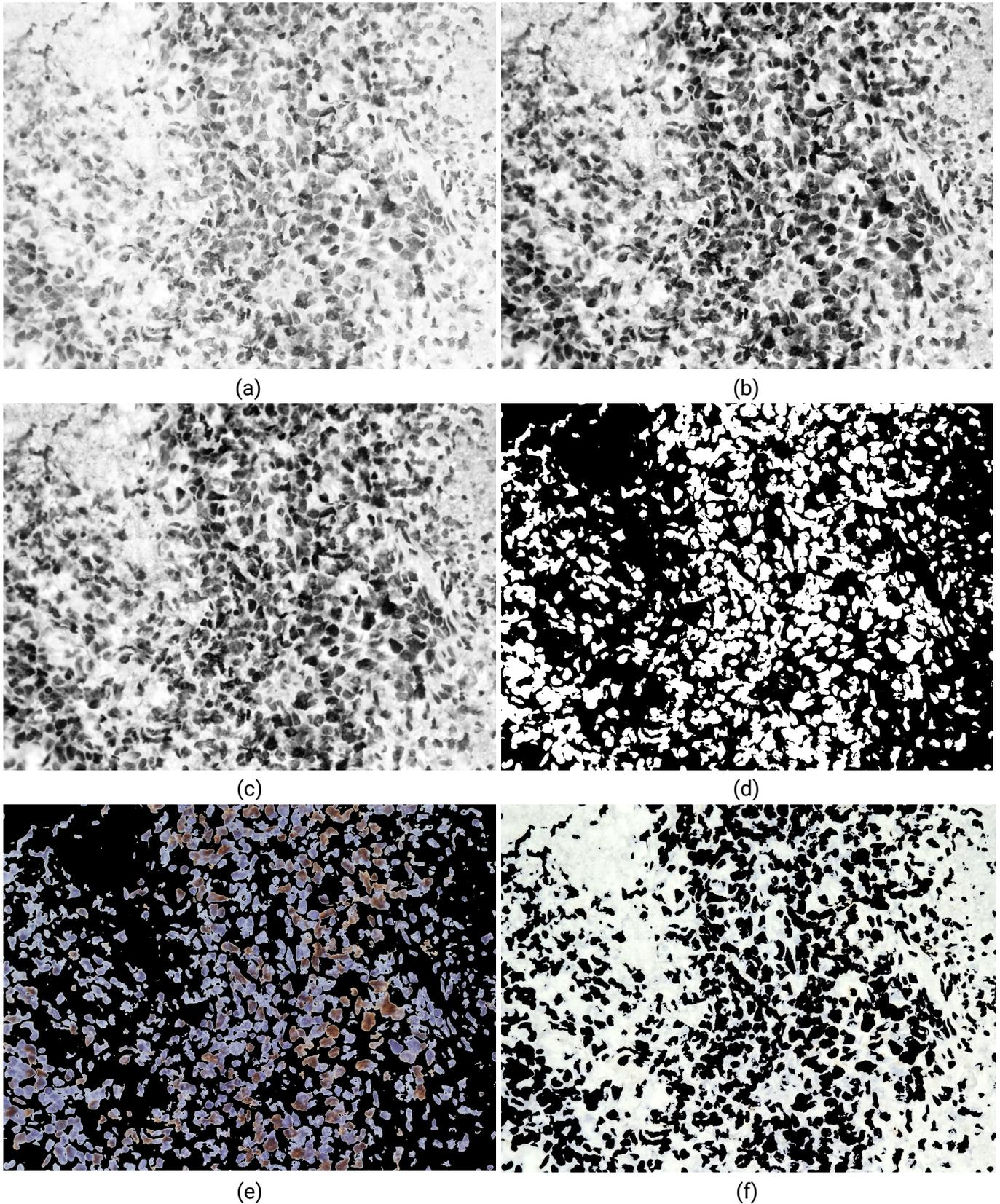
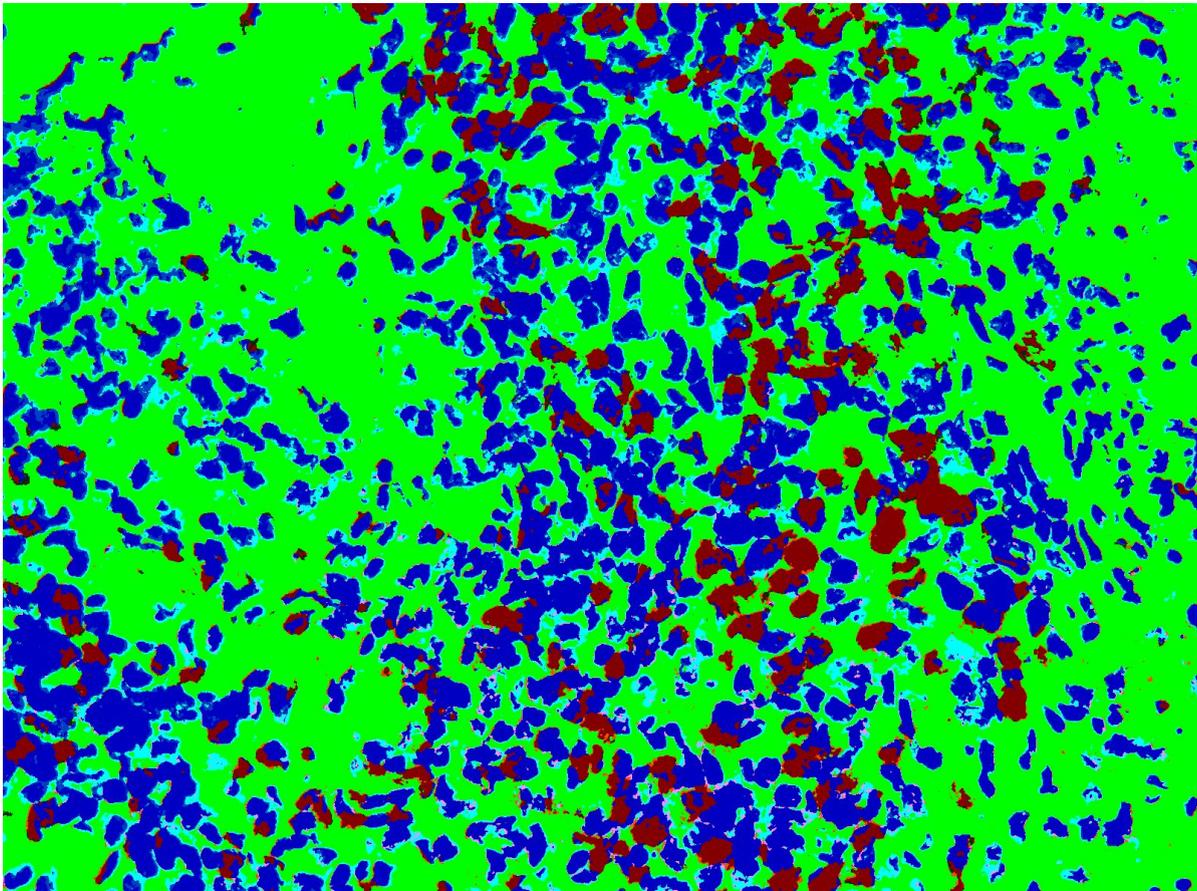
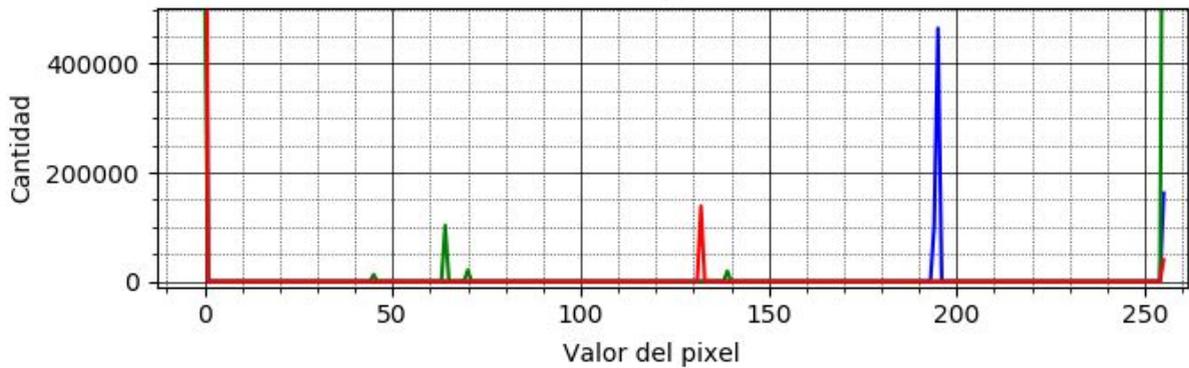


Figura 4.7. Proceso de extracción del fondo. (a) Imagen en escala de grises; (b) Aplicación de contraste adaptativo del histograma; (c) Aplicación de filtro gaussiano; (d) Aplicación del método de umbralizado de Otsu. (e) Imagen original en donde el fondo se ha reemplazado por negro; (f) Imagen original a color, tras haber aplicado inversamente la máscara d, en esta imagen solo se observa el fondo sin células. Se logra diferenciar cada vez más las células del fondo (a,c) para poder umbralizar (d) y se evalúa el resultado (e,f).



(a)

Histograma



(b)

Figura 4.8. Resultados tras aplicar el algoritmo propuesto (Algoritmo 4). (a) Imagen resultante en espacio BGR; (b) Histograma de la imagen resultante, los colores de las líneas representan el plano respectivo en el espacio de color BGR. Se observa que los tres picos de color en la imagen representan de forma aceptable: fondo (verde), marcaciones positivas (rojo) y células no marcadas (azul).

En la Figura 4.8 se muestra el resultado de aplicar el Algoritmo 4 a la imagen original y su respectivo histograma. Se logró saturar cada pixel en su componente dominante de percepción colorimétrica a través de comparaciones interplanares

en el espacio de color l^*a^*b . Este resultado es bastante sensible a los colores primarios (azules, verdes y rojos) ya que el espacio de color l^*a^*b está basado en la percepción humana, que a su vez se basa en la respuesta de células Cono del ojo de tipo S, M y L, a la longitud de onda de la luz incidente, la cual es principalmente alta en luz con longitudes de 450, 550 y 600 nm, asociadas a los colores azul, verde y rojo respectivamente [143].

Posteriormente, la aplicación de filtros pasa bandas permitió obtener máscaras para marcaciones positivas (rojos) y marcaciones negativas (azules), el histograma de la imagen en la Figura 4.8 reveló fuerte saturación de pixeles en los rangos del espacio BGR (0,0,100) a (0,0,150) para marcaciones positivas EBV y (180,0,0) a (210,0,0) para marcaciones negativas.

Tras umbralizar se obtuvieron buenos resultados de máscaras para las marcaciones positivas, negativas y fondo, por lo que por sí solo el Algoritmo 4 con umbralizado simple pasa bandas, fue capaz de extraer el fondo de la imagen sin necesidad de aplicar el método de Otsu. Sin embargo en algunas imágenes, el Algoritmo 4 presentó mucha sensibilidad a colores oscuros del fondo, marcandolos pertenecientes al rango de los rojos, e incluyó en el rango de los azules (marcaciones negativas) a secciones del fondo con predominancia azul. Estos inconvenientes no se presentaron al aplicar el método de Otsu a la imagen original, por lo tanto se combinó lo mejor de los dos métodos a través de la operación lógica AND entre la máscara de fondo (Figura 4.7.d) y las máscaras para rojo (Figura 4.9.a) y azul (Figura 4.9.b). Se comprobó que las máscaras estuviesen libres de falsos positivos (Figuras 4.9.e, 4.9.f) aplicando cada máscara en la imagen original. Además se verificó que se las máscaras cubrieran la totalidad de zonas celulares marcadas y no marcadas al aplicarlas de forma inversa (forma normal AND, forma inversa NOT-AND) en la imagen original (Figura 4.10).

A continuación en una serie de figuras, se presentan los resultados de lo explicado anteriormente, en la Figura 4.9, se observa el proceso de obtención de las máscaras definitivas para separar la imagen entre marcaciones positivas y negativas. Posteriormente en la Figura 4.10 se observa el resultado de aplicar inversamente dichas máscaras a la imagen original para verificar su efectividad y encontrar falsos negativos.

En el siguiente subcapítulo, en la Figura 4.11 se evidencian los marcadores utilizados: *semillas/sure_foreground* y *fondo/sure_background*, los cuales sirven para obtener el área desconocida, de ahí el algoritmo *watershed* identifica de qué semilla hace parte cada píxel del área desconocida, lo que se muestra en las Figuras 4.12 y 4.13 como el resultado de la segmentación.

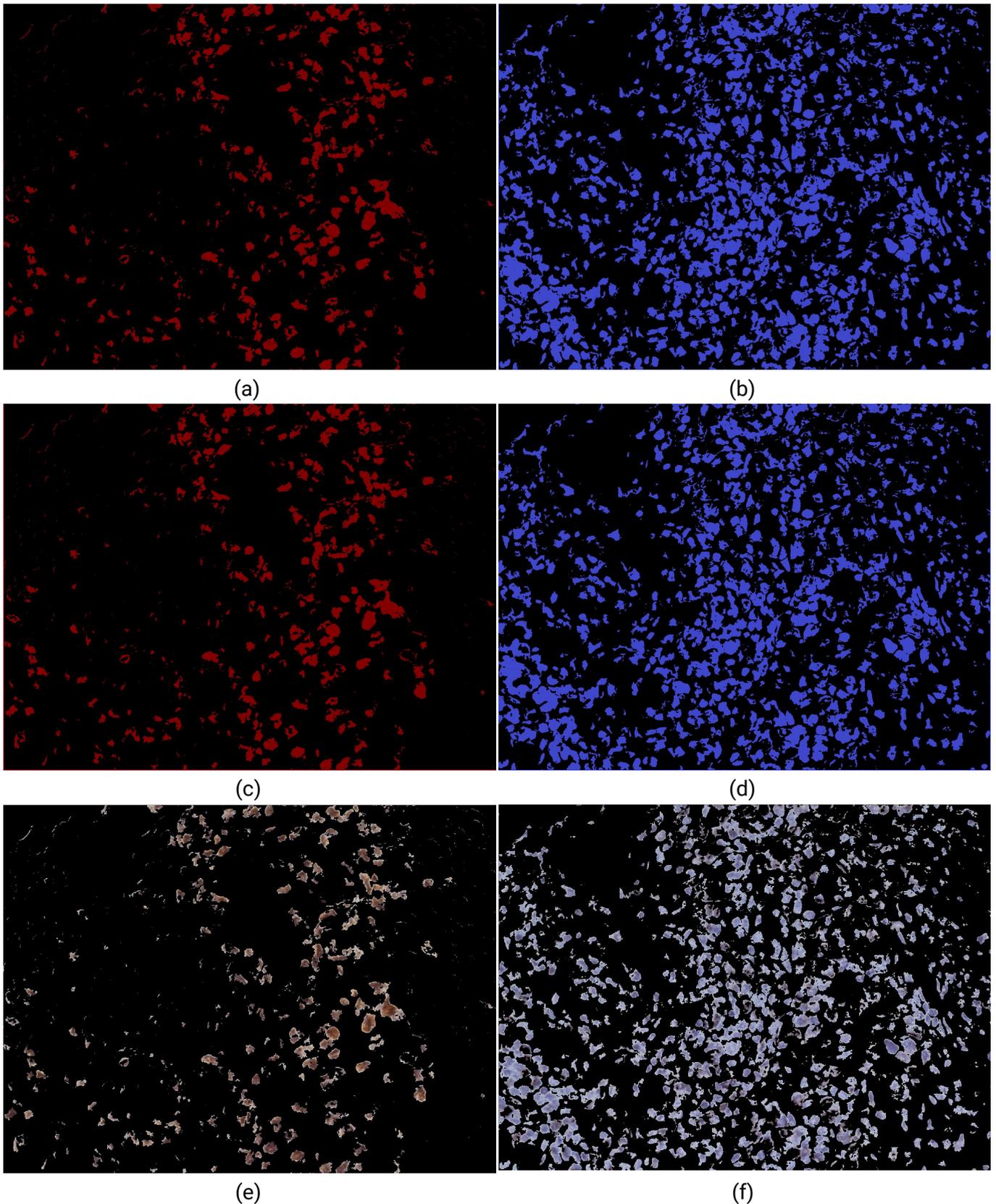
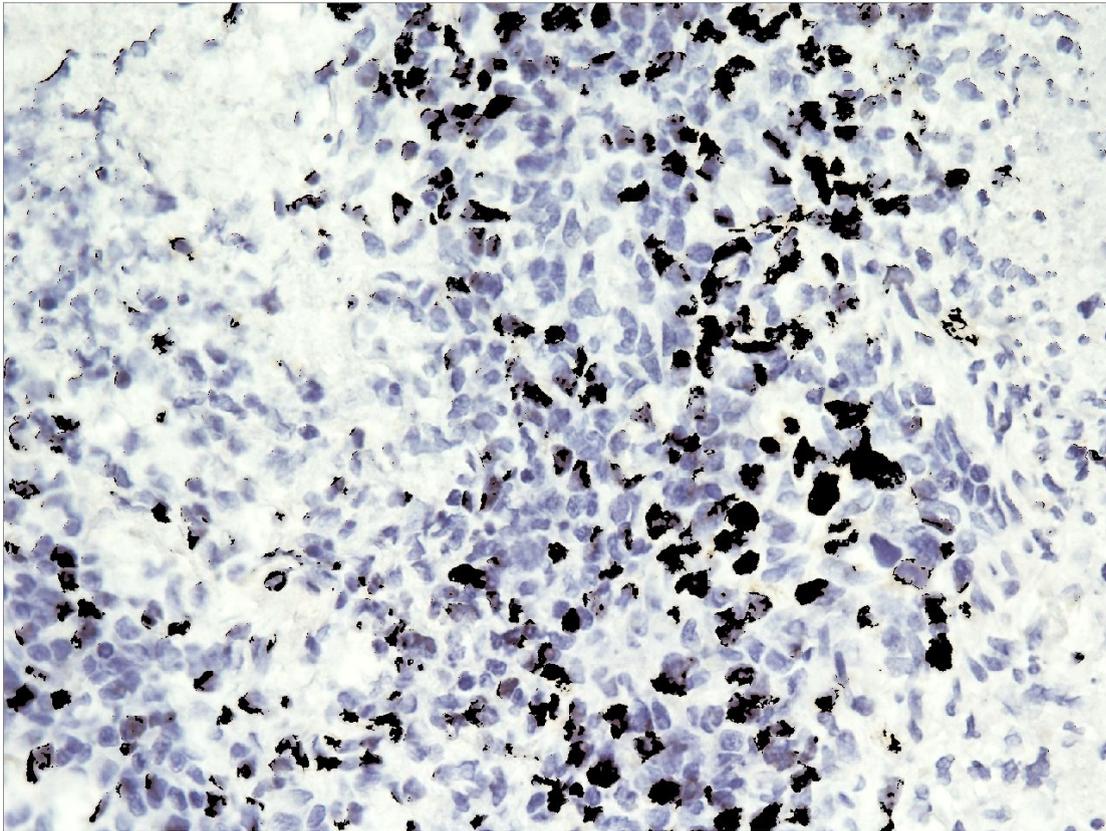
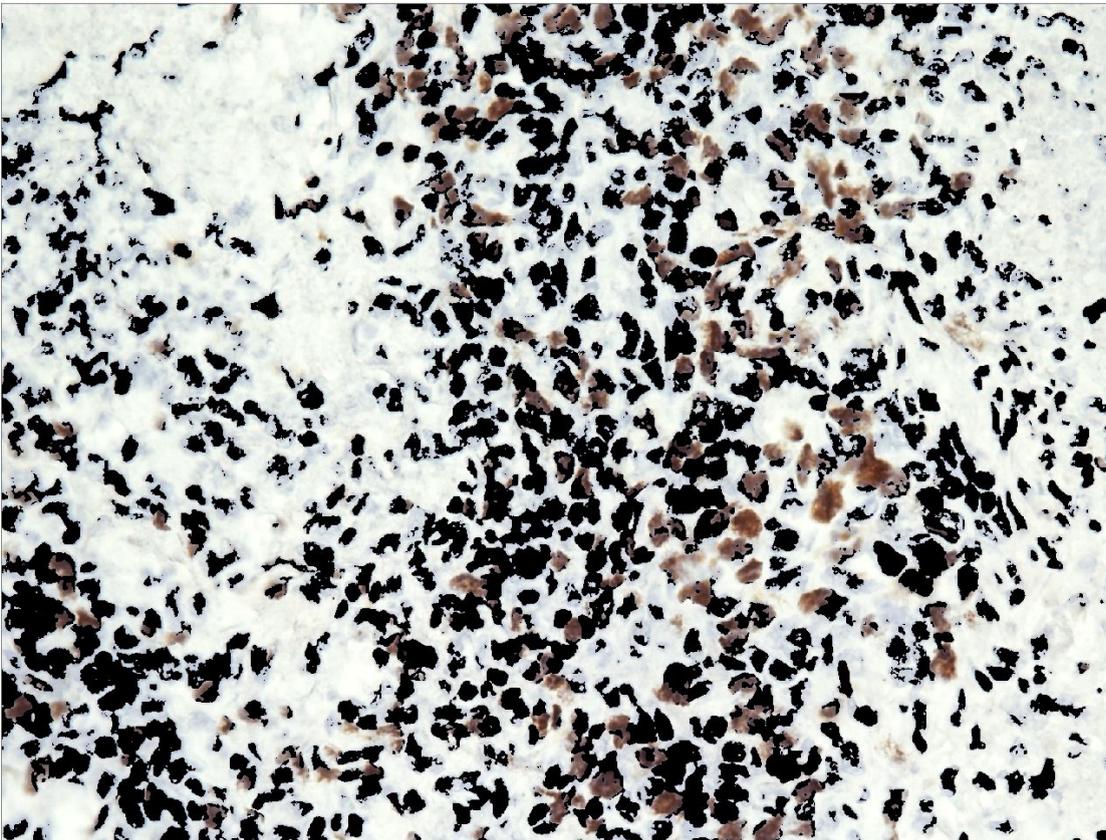


Figura 4.9. Umbralizado de la imagen en la Figura 4.8. (a) Rango de rojos; (b) Rango de azules; (c & d) Resultado de la operación AND entre la máscara de fondo (Figura 4.7.d) y las máscaras para rojo (Figura 4.9.a) y azul (Figura 4.9.b); (e & f) Aplicación de las máscaras 4.9.c y 4.9.d a la imagen original. Observe que inicialmente las máscaras cubren partes extra que no corresponden a las zonas designadas las operaciones logran disminuir este efecto al utilizar el resultado de la segmentación Otsu.



(a)



(b)

Figura 4.10. Aplicación inversa de las máscaras 4.9.c y 4.9.d respectivamente. Observe que al aplicar las máscaras no se observan células cafés en (a) y no se observan células azules en (b).

4.4.2 Segmentación

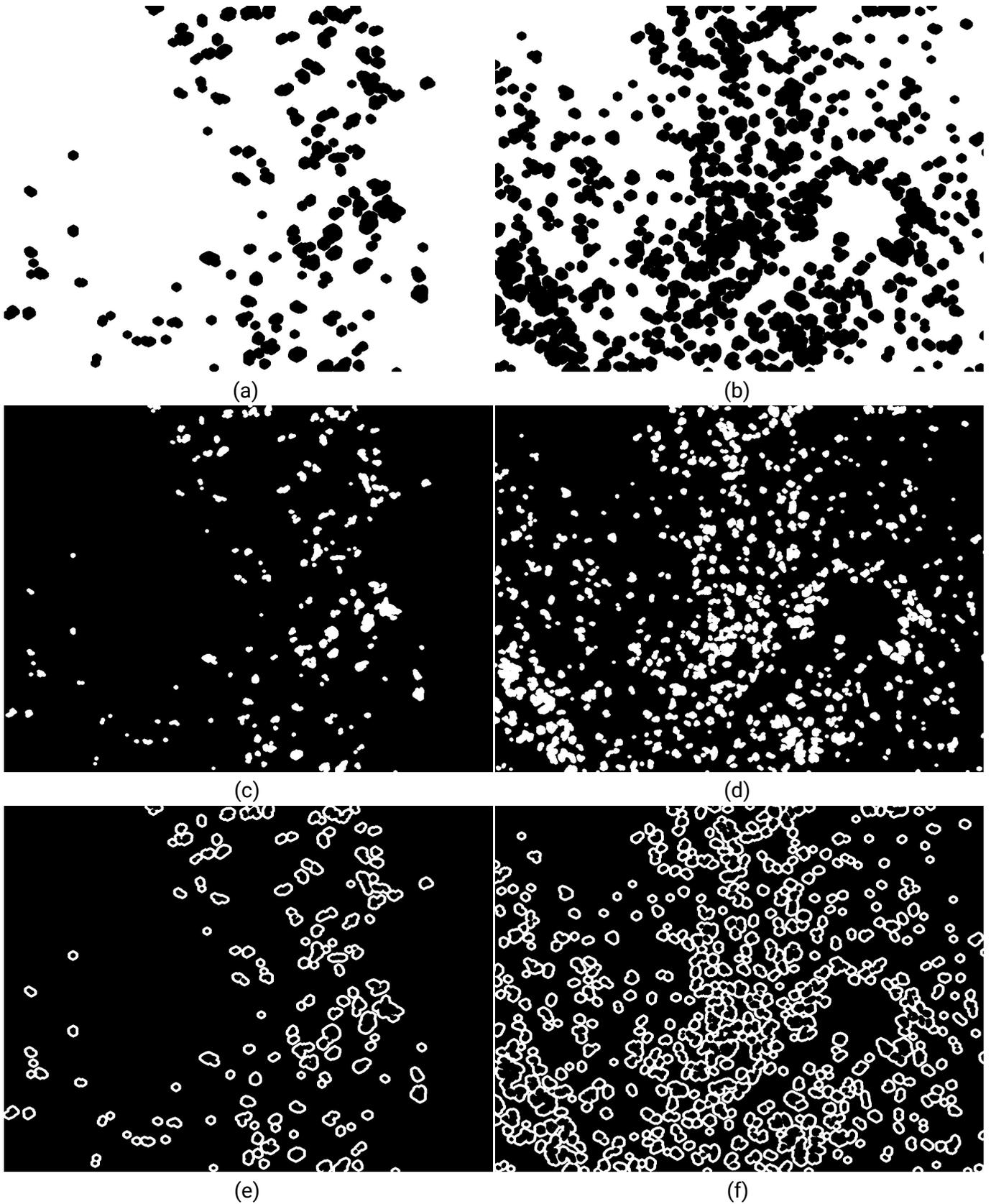
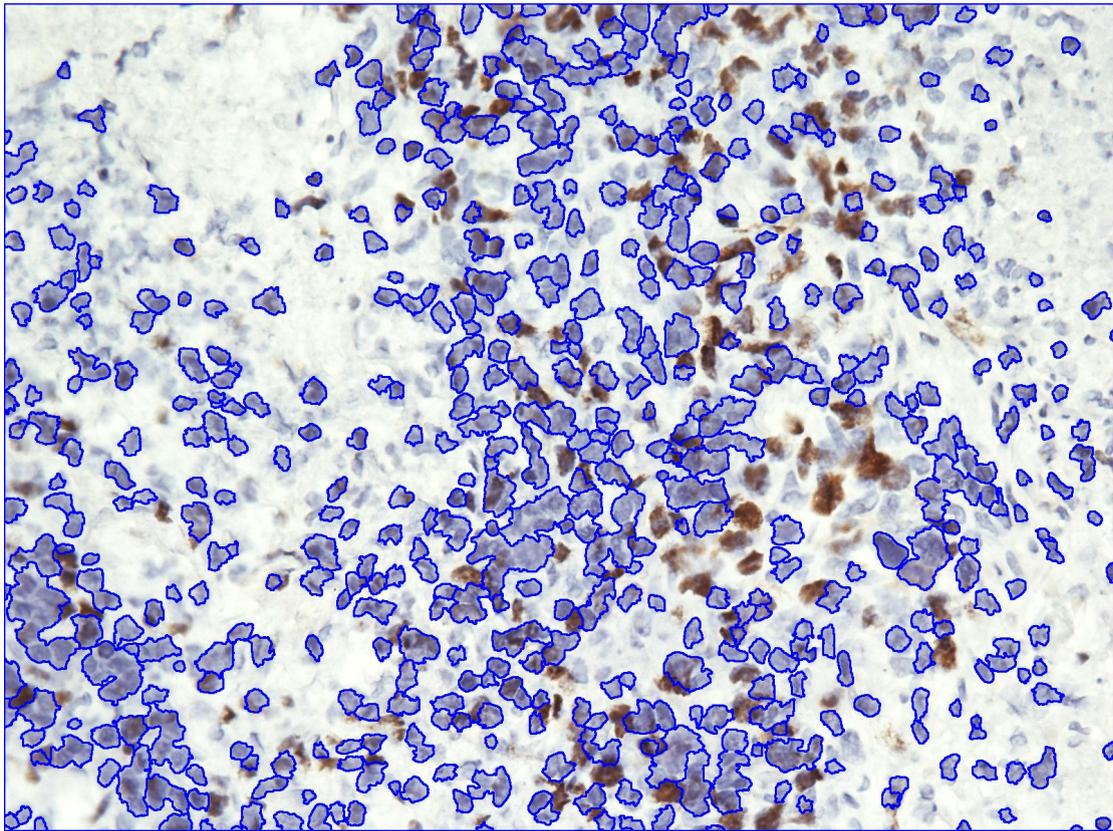
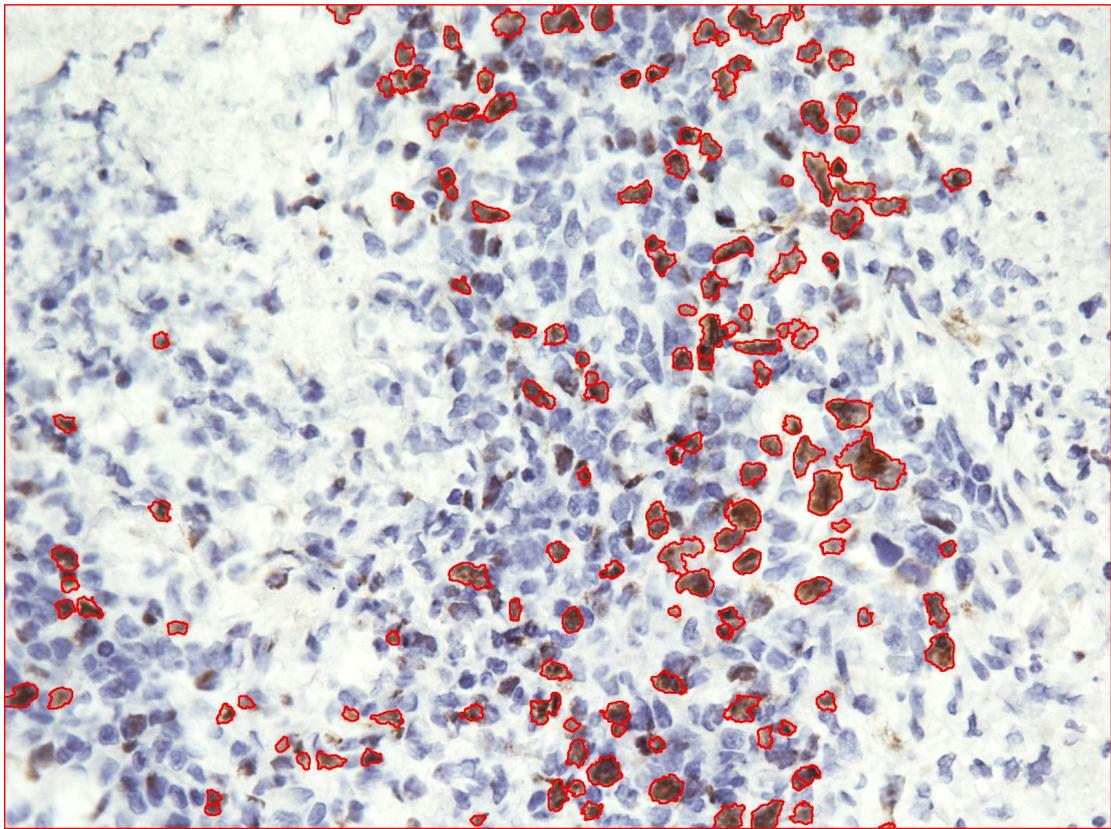


Figura 4.11. Marcadores de la imagen. (a) Sure background rojos; (b) Sure background azules; (c) Sure foreground rojos; (d) Sure background azules; (e) Área desconocida rojos; (f) Área desconocida azules.

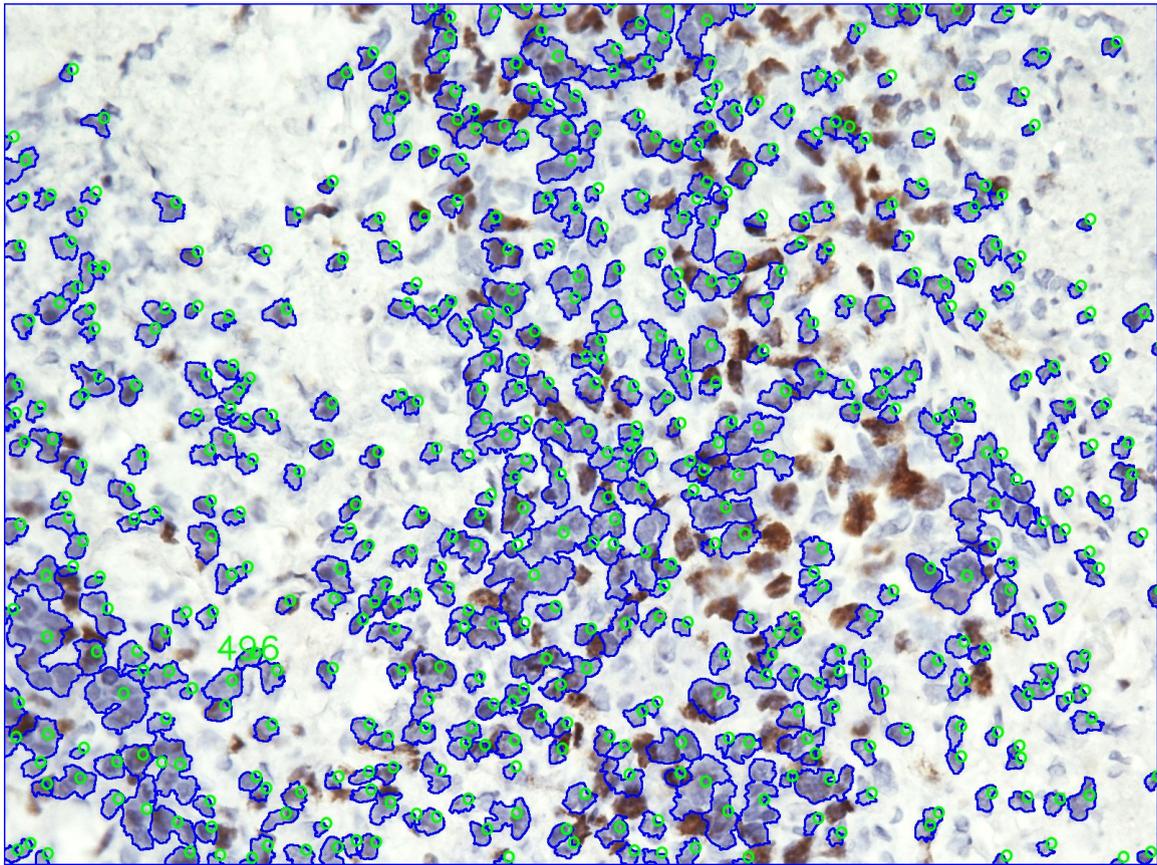


(a)

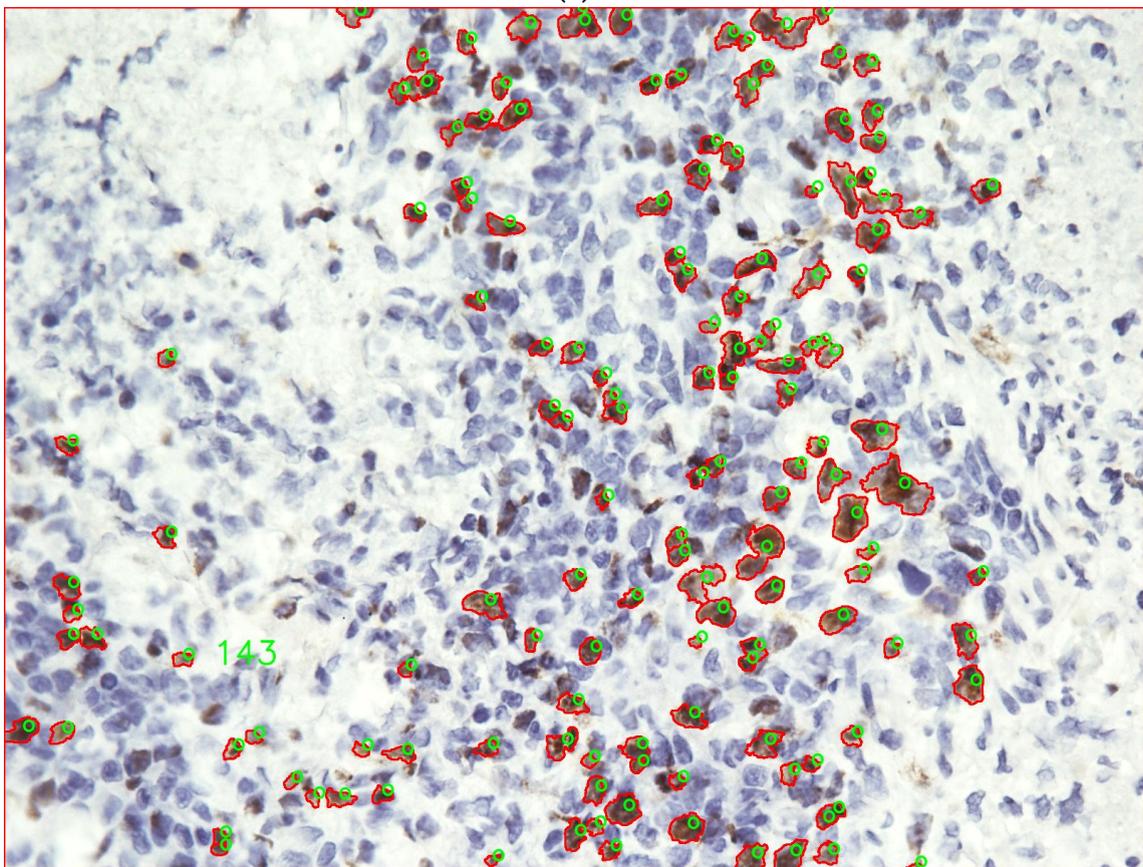


(b)

Figura 4.12. Resultado de segmentación por watershed controlado por marcadores. (a) Células no infectadas; (b) Células Infectadas (marcación positiva para EBV). Se nota sub-segmentación en las partes donde hay mayor apiñamiento, esto se debe a errores en el posicionamiento de los sure foreground.



(a)



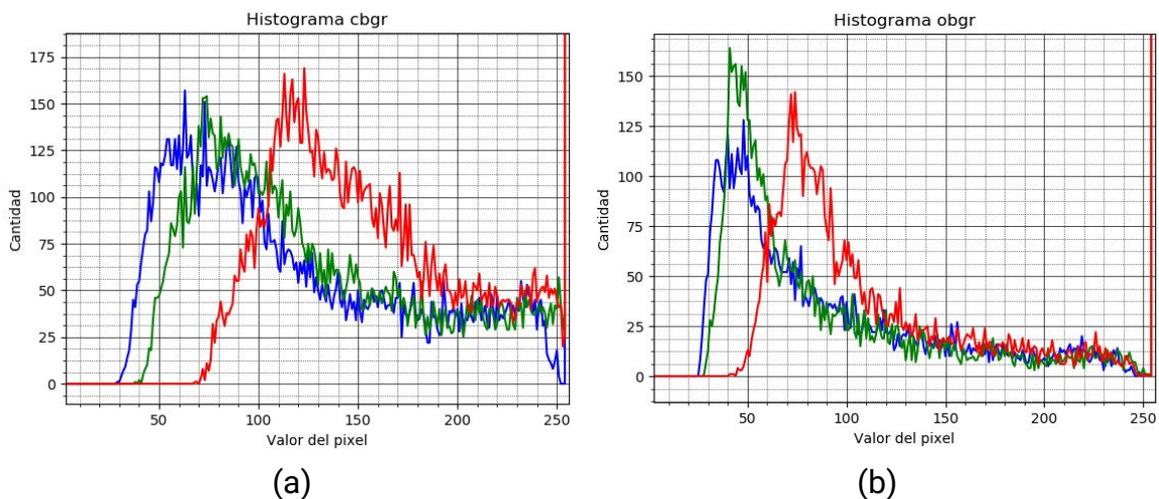
(b)

Figura 4.13. Conteo de células segmentadas. Número de Células (a) 496; (b) 143. Centroides en verde.

Se logró segmentar exitosamente las imágenes en sus componentes primarios de interés, células marcadas y no marcadas (Figura 4.12). Al utilizar marcadores para controlar el algoritmo de watershed se soluciona el problema de sobre segmentación presentado en el análisis de la literatura (Figura 4.11), sin embargo, la efectividad de este método se ve comprometida por la precisión de los marcadores, tal como se observa en la Figura 4.13.a, para imágenes con células fuertemente conglomeradas se cae en la subsegmentación, es decir que se segmentan menos células de las que realmente hay, lo que puede ser corregido mediante iteraciones del método para secciones segmentadas que superen un área determinada. Para propósitos de este trabajo de investigación no se necesitó corregir este inconveniente, ya que el análisis patológico de la imagen se enfoca en las células marcadas, que no presentan fuertes conglomeraciones, sino que se encuentran dispersas en el tejido como se muestra en la Figura 4.13.b. Cabe resaltar que haber escogido un *kernel* del tamaño y forma del núcleo de la célula promedio para la apertura morfológica, como se mencionó en el análisis de la literatura, fue la clave para obtener buenos marcadores.

4.4.3 Diferenciación de marcaciones fuertes y leves

Al analizar imágenes de células con tinción moderada e intensa y sus respectivos histogramas en diferentes espacios de color (Figura 4.14 y Figura 4.15), se identificó que el canal R del espacio de color BGR presentó la mayor diferenciación visual y cuantificable entre los dos grupos. Basándose en la lista de promedios del valor del canal R para cada célula detectada, se determinó el valor 118 como umbral óptimo para diferenciar células con tinción moderada de células con tinción intensa. De esta manera, un valor superior a 118 en el promedio de valores del canal R indica a la herramienta que es una marcación moderada y un valor inferior una marcación fuerte. El resultado se muestra en la Figura 4.16.



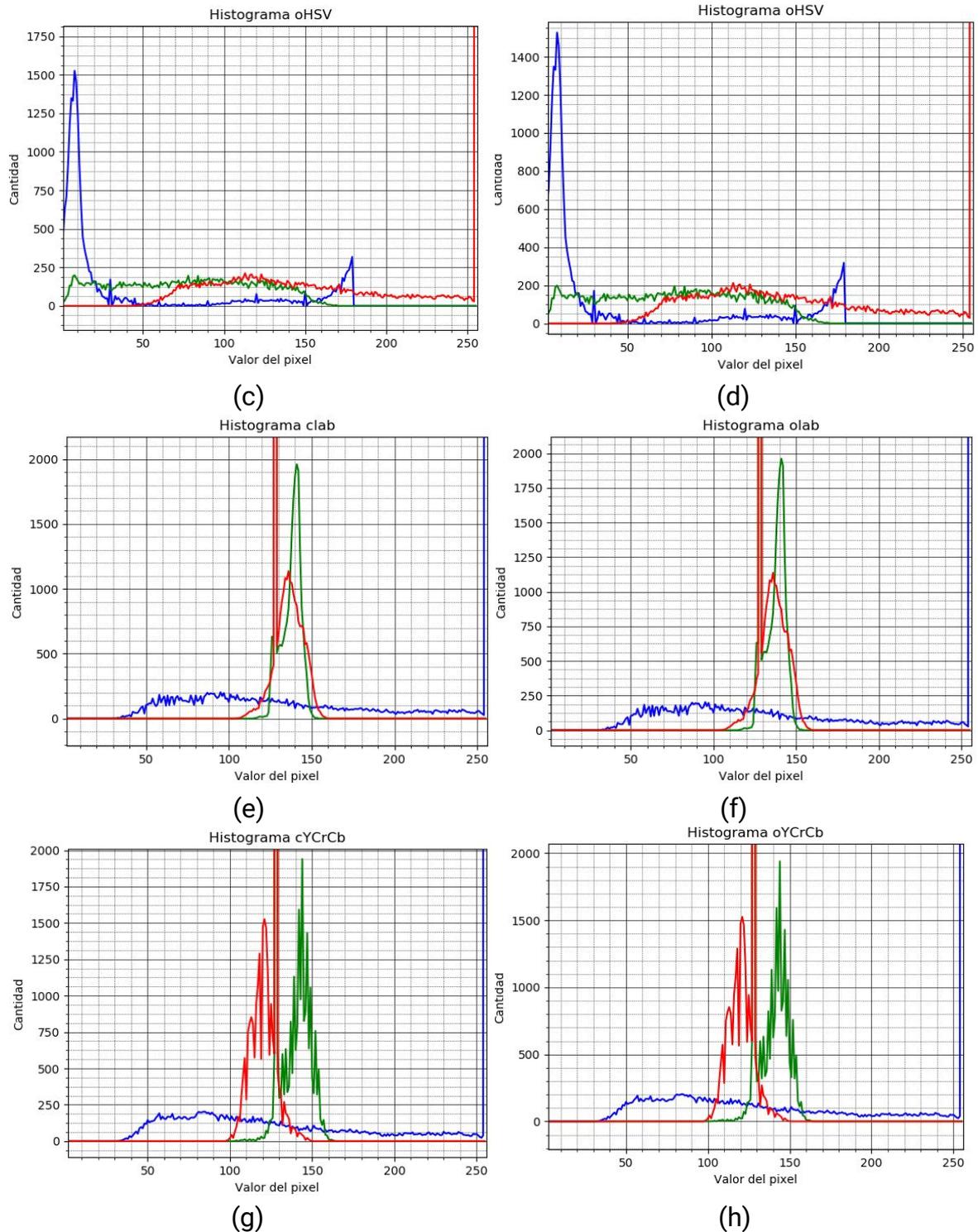


Figura 4.14. Análisis cuantitativo de células con tinción intensa y moderada en los diferentes espacios de color. La línea azul corresponde a la primera letra del espacio de color, la línea verde a la segunda y la línea roja a la tercera. Se observa que de todos el plano R del espacio BGR presenta la mayor diferencia cuantificable. Histograma para tinción: (a) Moderada BGR; (b) Fuerte BGR; (c) Moderada HSV; (d) Fuerte HSV; (e) Moderada L^*a^*b ; (f) Fuerte L^*a^*b ; (g) Moderada YCrCb; (h) Fuerte YCrCb.

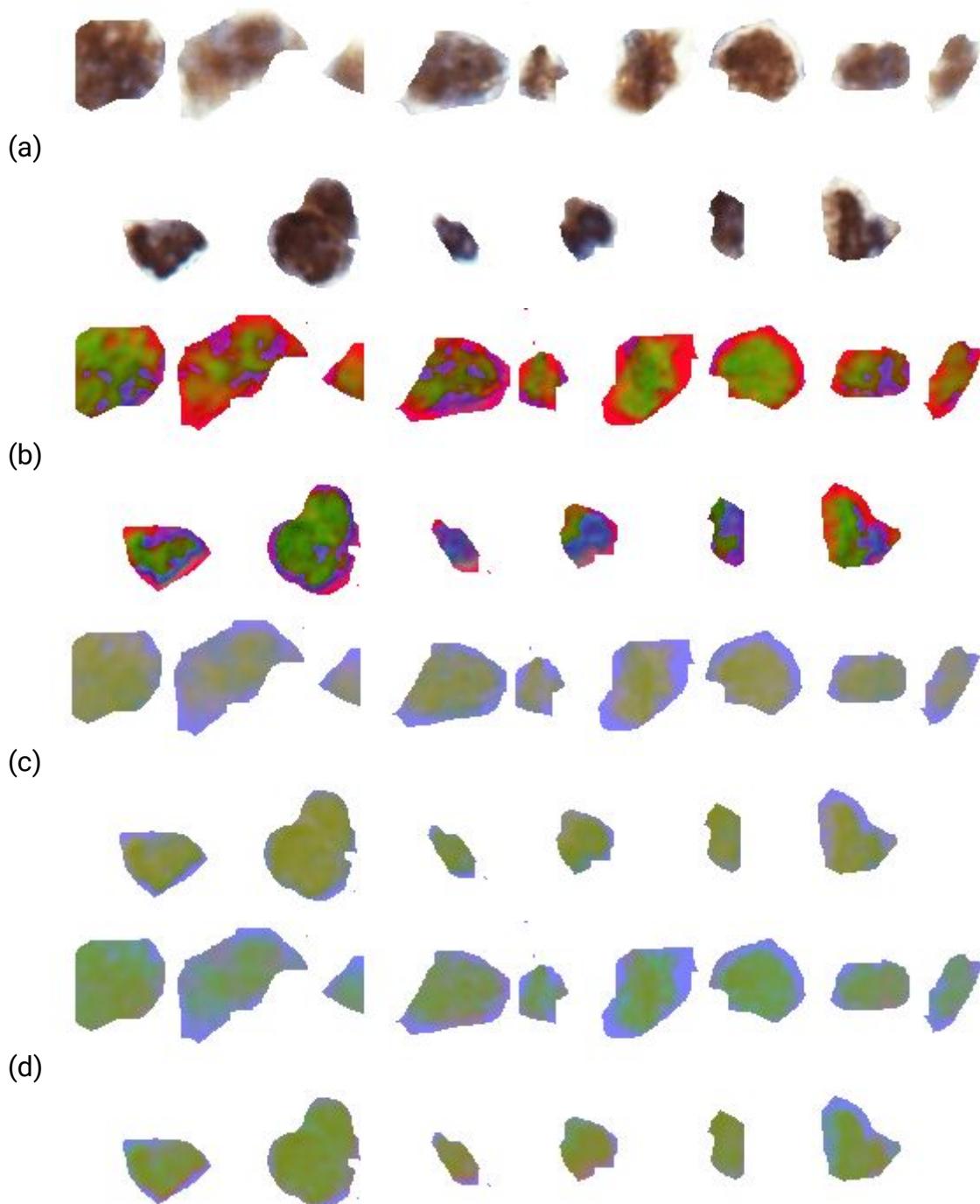


Figura 4.15 Análisis visual de células con tinción intensa y moderada en diferentes espacios de color. En las imágenes se observan dos filas de células segmentadas manualmente, las células más oscuras de la fila inferior corresponden a aquellas infectadas fuertemente y las de la fila superior a aquellas infectadas moderadamente. De todos el espacio BGR presenta la mayor diferencia visual entre los dos grupos de células. Espacio de color: (a)BGR; (b)HSV; (c)L*a*b; (d)YCbCr.

Lo que es una marcación intensa para un patólogo puede ser considerada una marcación leve para otro, por tal motivo encontrar un conjunto grande de células representativas de cada grupo que tuviera aceptación entre los patólogos no fue

posible, sin embargo el programa fue capaz de diferenciar marcaciones que evidentemente poseen un tono más oscuro.

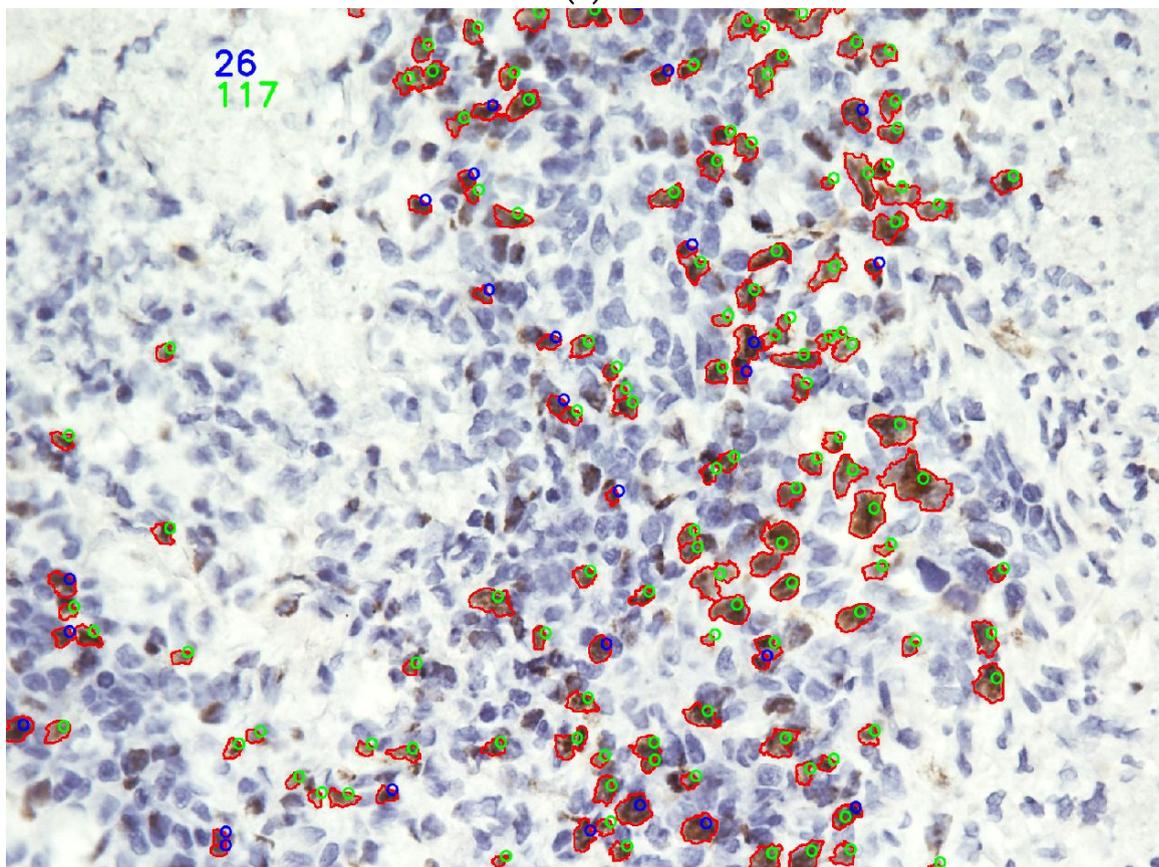
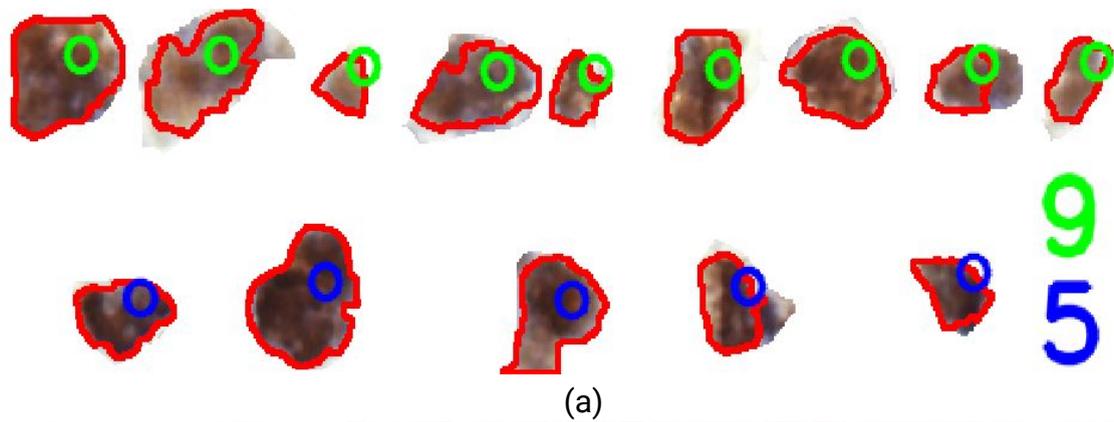


Figura 4.16. Resultado de aplicar el umbral determinado tras los análisis visual y cuantitativo, para identificar y contar células en la imagen. Los marcadores azules indican marcaciones intensas y los marcadores verdes indican marcaciones moderadas. (a)Ejemplo de imagen de calibración; (b) Ejemplo de imagen de prueba.

Se resalta de las Figuras 4.14, 4.15 y 4.16 que independientemente de los conjuntos de células que se escojan como claras y oscuras el método utilizado de verificación de los histogramas y cambio visual proporciona buenos resultados.

4.5 Digitalización de las placas

Tras realizar las pruebas respectivas, se determinó la mejor configuración para la adquisición de las imágenes, mostrada en la Tabla 4.5 y se digitalizaron las placas siguiendo la metodología. A continuación en la Tabla 4.6 se muestra la configuración ideal de la cámara y posteriormente en la Tabla 4.7 una comparación de algunos de los resultados de las pruebas sobre las imágenes adquiridas bajo diferentes configuraciones, que se muestran en la Tabla 4.6.

Tabla 4.5. Configuración de equipos para la digitalización de las placas con tejido.

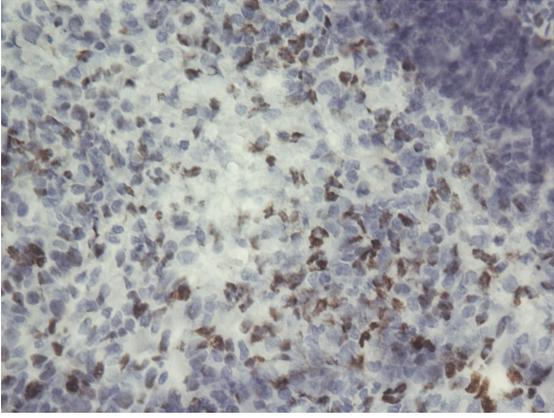
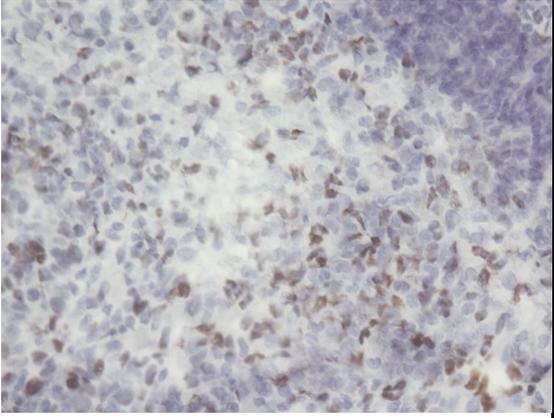
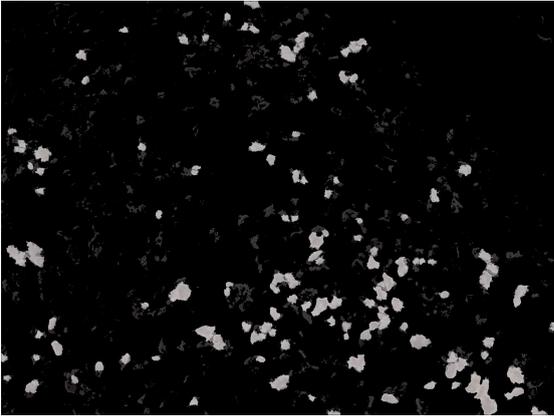
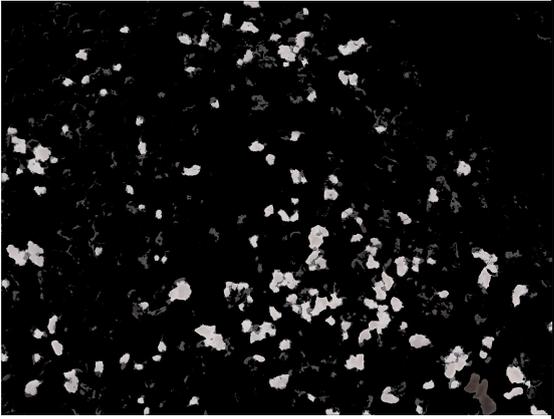
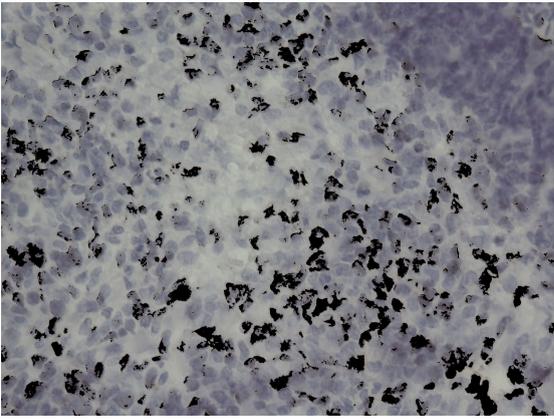
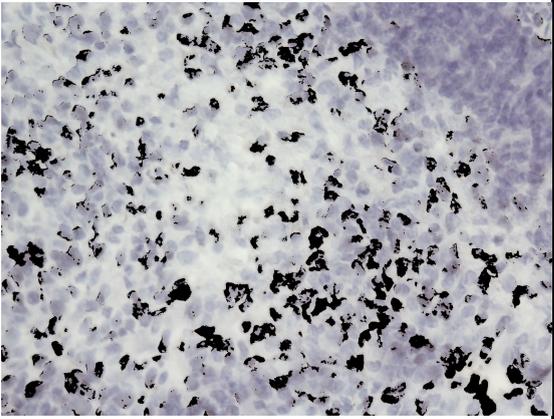
Cámara	NIKON DS-MV-U2
Velocidad Obturación	1 / 60s
Contraste	ENHANCED
Balance de blancos	SÍ
Iluminación	Seis puntos
Apertura del diafragma	Seis puntos
Ocular y Objetivo	10X y 40X

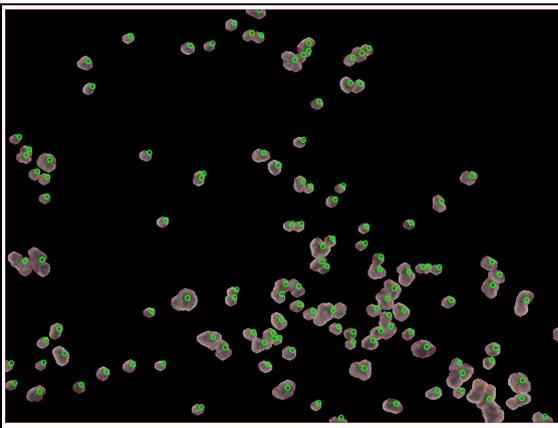
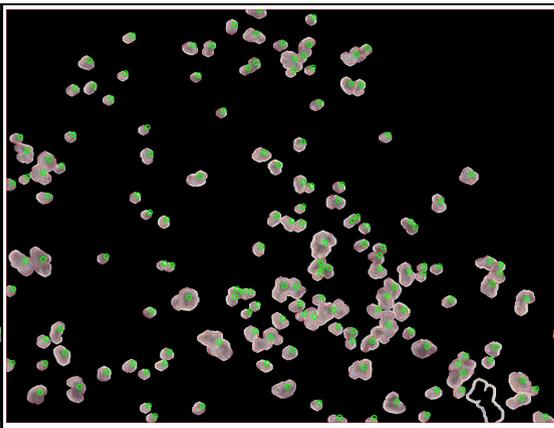
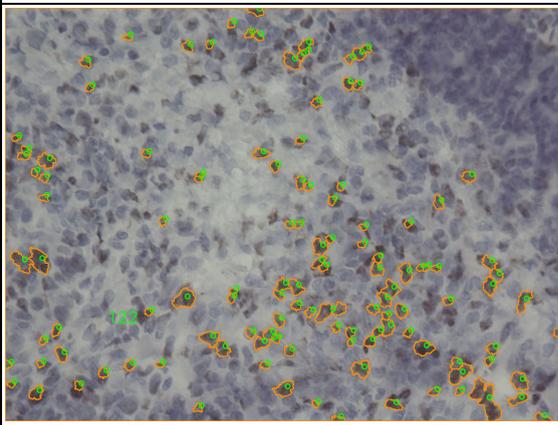
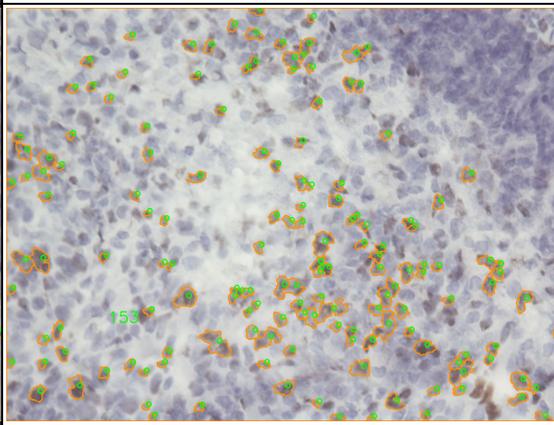
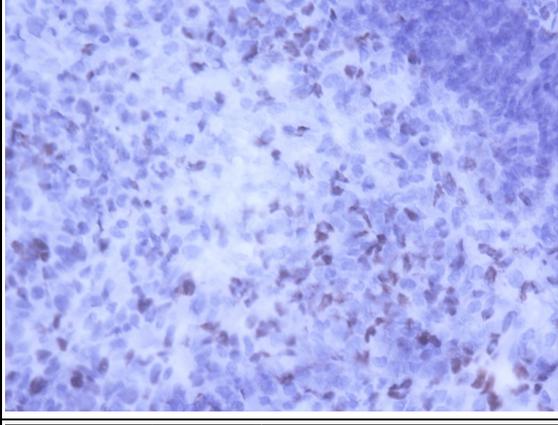
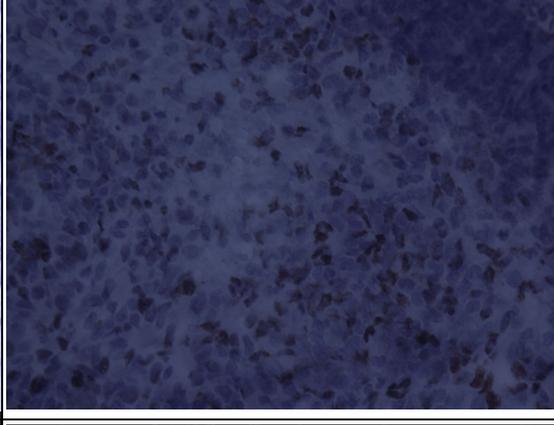
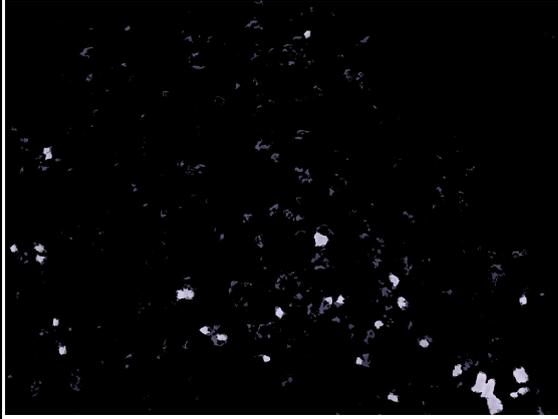
Tabla 4.6 Comparación de valores representativos de configuración (Tabla 4.4) para la adquisición de las imágenes. El # de configuración se utiliza en la Tabla 4.7.

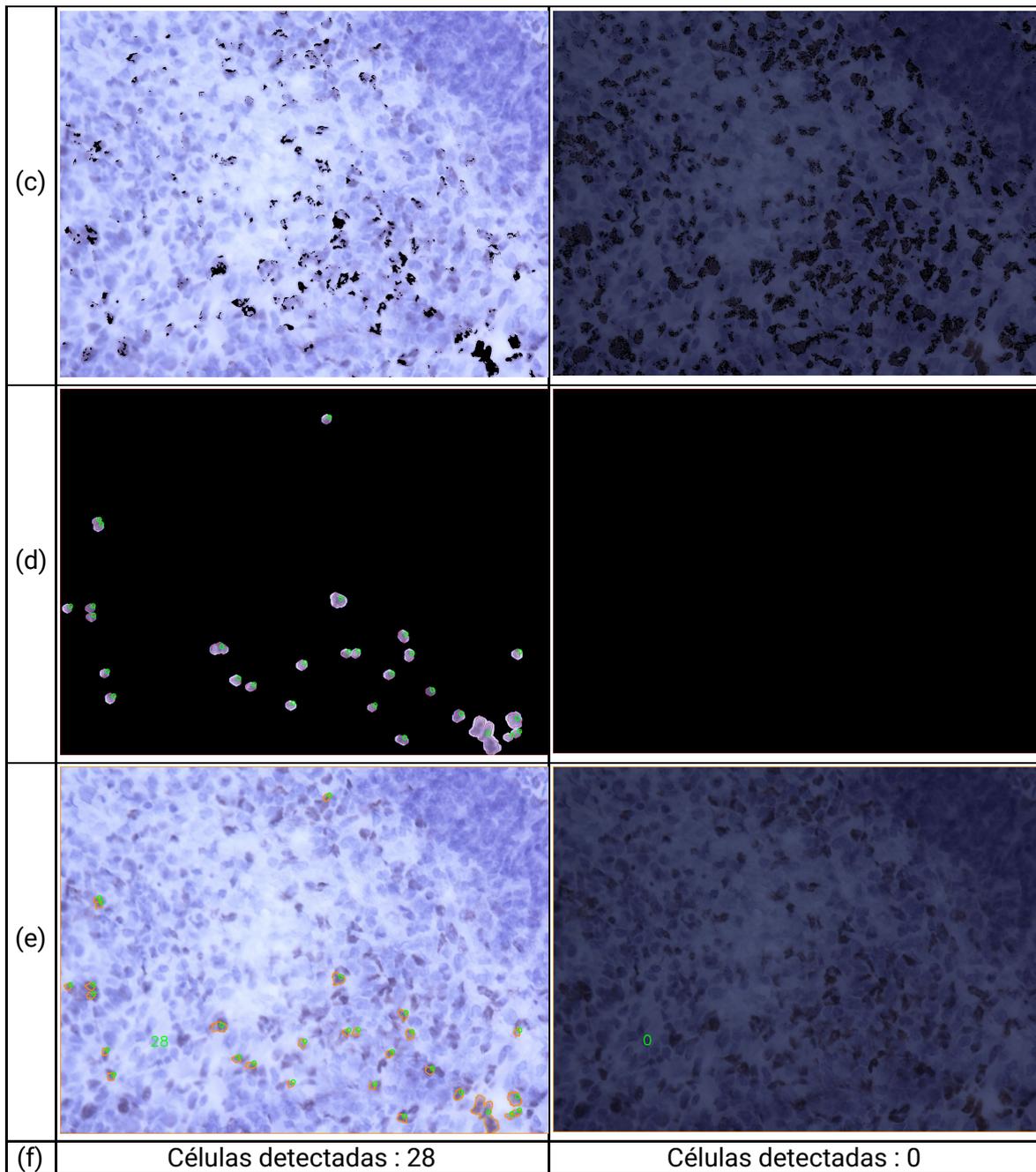
Mejor configuración	# Configuración	Velocidad obturación	Contraste	Balance de blancos
Sí	1	1/60s	ENHANCED	Sí
No	2	1/60s	HIGH	Sí
No	3	1/60s	HIGH	No
No	4	AUTO	ENHANCED	No
No	5	AUTO	ENHANCED	Si

Las cuatro pruebas subjetivas mostradas en la Tabla 4.7 revelaron aspectos importantes del proceso. La primera prueba, (a, Imagen original), reveló características sobre el color de la imagen, por ejemplo en la configuración 4 se muestra una imagen original bastante oscura que no pasa la primera prueba. La segunda prueba, (b, Marcaciones específicas en gris claro vs marcaciones inespecíficas en gris oscuro); en conjunto con la tercera y cuarta prueba (c, Imagen original sin las células marcadas; d, marcación de células sin el fondo) revelaron la eficacia de las máscaras para determinar la pertenencia de los píxeles a células sanas, infectadas o *background* y en diferenciar entre marcaciones específicas y marcaciones inespecíficas en diferentes configuraciones. Lo más importante a resaltar de la Tabla 4.7 es que estas pruebas demuestran la superioridad de ciertas configuraciones sobre otras, ejemplo las configuraciones 1 y 2 presentan mejor respuesta al algoritmo que las configuraciones 3 y 4.

Tabla 4.7 Comparación de algunos de los resultados de las pruebas sobre las imágenes adquiridas bajo diferentes configuraciones. (a) Imagen original; (b) Marcaciones específicas en gris claro vs marcaciones inespecíficas en gris oscuro; (c) Imagen original sin las células marcadas; (d) Marcación de células sin el fondo; (e) Aplicación de la herramienta sobre las células infectadas; (f) Conteo final de células infectadas reconocidas. Las diferentes configuraciones son aquellas mostradas en la Tabla 4.6.

#	Configuración 1	Configuración 2
(a)		
(b)		
(c)		

(d)		
(e)		
(f)	Células detectadas : 122	Células detectadas : 153
#	Configuración 3	Configuración 4
(a)		
(b)		



Al comparar los resultados de las pruebas para todas las imágenes iniciales, se concluyó que el balance de blancos es el parámetro de cámara más influyente en la eficacia de la herramienta, esto se explica analizando el pobre desempeño que presentó el algoritmo 4 en imágenes tomadas sin balance de blancos. Los demás parámetros afectaron el desempeño de la herramienta en menor medida, pero de todas formas, significativamente. En general a menor tiempo de obturación se obtuvieron peores resultados y entre mayor contraste mejores resultados.

4.6 Presentación de resultados de la caracterización

En la Tabla 4.8, se presentan los valores de cuantificación de identificación del EBV calculados para cada punto de los microarreglos obtenidos al someter las imágenes finales a la herramienta. Con lo que se caracteriza el microambiente tumoral de cada paciente. En esta tabla, aparece la columna de coordenadas (x, y), que se refiere a las coordenadas en la placa dadas por el microscopio para identificar cada uno de los puntos del microarreglo, como uno de los que se ve en la Figura 3.14 (sección 3.5). Además, las diez columnas marcadas con las letras (a-j) contienen los valores de los componentes cuantitativos correspondientes (Tabla 3.4, sección 3.6).

Tabla 4.8. Presentación de valores de cuantificación de identificación del EBV calculados para cada punto de los microarreglos. El índice de los cuantificadores (a-j) se presenta en la tabla 3.4 encontrada en el capítulo 3. El # indica el número marcado en la placa portaobjetos donde se encuentra el punto analizado y su posición en la placa está marcada por las coordenadas (x, y) dadas por el microscopio. La letra "C" en el listado de placas representa la placa control.

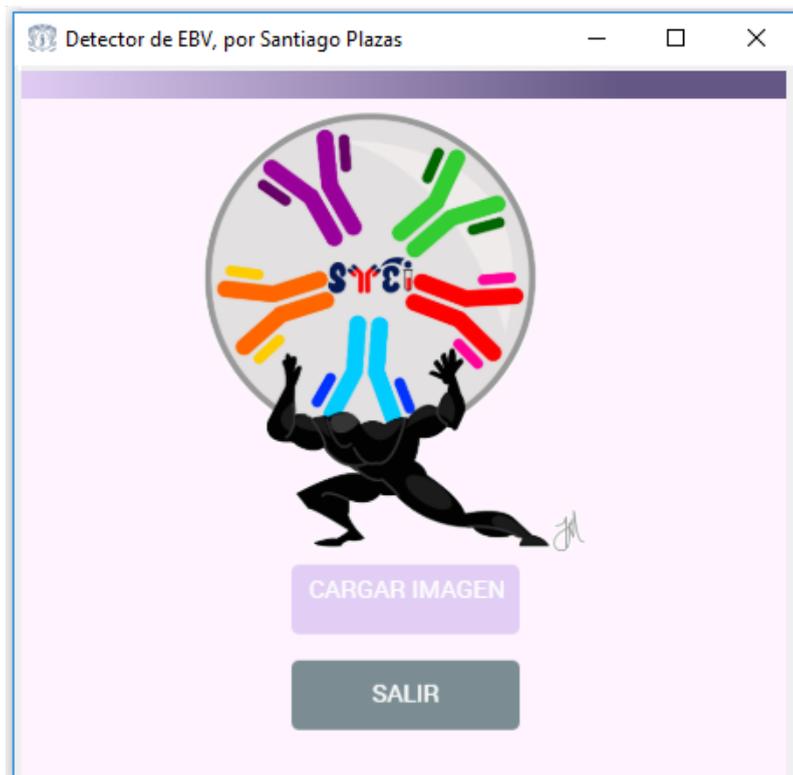
Cuantificador→		(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)
↓ Punto microarreglo	#										
#	Coordenadas (x,y)										
1	(72.8, 108.2)	11334	756	7	749	94%	6%	1%	99%	0%	6%
1	(60.5, 106.7)	10608	94	2	92	99%	1%	2%	98%	0%	1%
2	(71.0, 109.0)	11514	715	191	524	94%	6%	27%	73%	2%	4%
2	(62.8, 115.0)	2434	92	6	86	96%	4%	7%	93%	0%	3%
2	(58.2, 105.3)	9664	282	20	262	97%	3%	7%	93%	0%	3%
2	(58.8, 109.0)	11211	359	32	327	97%	3%	9%	91%	0%	3%
2	(54.0, 109.0)	12910	436	16	420	97%	3%	4%	96%	0%	3%
3	(52.9, 118.4)	12668	142	5	137	99%	1%	4%	96%	0%	1%
3	(56.3, 113.2)	8938	2939	8	2931	75%	25%	0%	1%	0%	25%
3	(51.0, 113.0)	9211	738	5	733	93%	7%	1%	99%	0%	7%
3	(47.0, 115.0)	5268	435	7	428	92%	8%	2%	98%	0%	8%
3	(50.8, 107.7)	10054	2460	3	2457	80%	20%	0%	1%	0%	20%
3	(45.7, 109.3)	11847	631	4	627	95%	5%	1%	99%	0%	5%
6	(71.0, 113.6)	5727	1171	0	1171	83%	17%	0%	1%	0%	17%
6	(66.9, 115.0)	4385	1218	9	1209	78%	22%	1%	99%	0%	22%
6	(53.5, 115.0)	6214	1428	8	1420	81%	19%	1%	99%	0%	19%
6	(70.0, 108.0)	5353	436	5	431	92%	8%	1%	99%	0%	7%
6	(60.4, 111.5)	5894	977	3	974	86%	14%	0%	1%	0%	14%
6	(54.0, 110.6)	4029	637	3	634	86%	14%	0%	1%	0%	14%
C	(52.0, 109.0)	10581	2290	881	1409	80%	20%	38%	62%	7%	11%

Los valores plasmados en la Tabla 4.8, hacen parte de una serie de resultados entregados al GYEi como parte de los compromisos adquiridos con el proyecto

4487 de la VRI. Estos resultados servirán a la investigación de la relación del EBV con el adenocarcinoma gástrico de tipo intestinal, la cual debe ser realizada por médicos calificados. Sin embargo, se observa que en la placa control, un 20% del total de células se identificaron como marcadas positivas para EBV (estadístico (f)), por lo tanto se supuso que cualquier biopsia con un valor (f) mayor a 20% se podría considerar preliminarmente positiva (esto quiere decir que se marcan dichas biopsias en particular con un alto porcentaje de ser positivas, pero sólo un patólogo puede hacer afirmaciones sobre el resultado de una prueba de CISH). En este orden de ideas se identificó que el 16% (3/19) de los puntos analizados cumplió los requerimientos para ser considerado preliminarmente positivo, un porcentaje acorde a la literatura ya que como se mencionó anteriormente entre el 10% y el 20% de los casos de CG se encuentran asociados a la infección por EBV [27].

4.7 Presentación de la herramienta

En el estudio tradicional de la identificación de EBV en biopsias de pacientes con adenocarcinoma gástrico de tipo intestinal, el patólogo expide su veredicto basándose en regiones de interés seleccionadas, descartando otras zonas. Por lo tanto en la herramienta desarrollada el patólogo debe seleccionar una imagen sobre la que sea relevante el cálculo de la cuantificación de la infección del EBV. En la Figura 4.17 se observa la interfaz de la herramienta desarrollada.



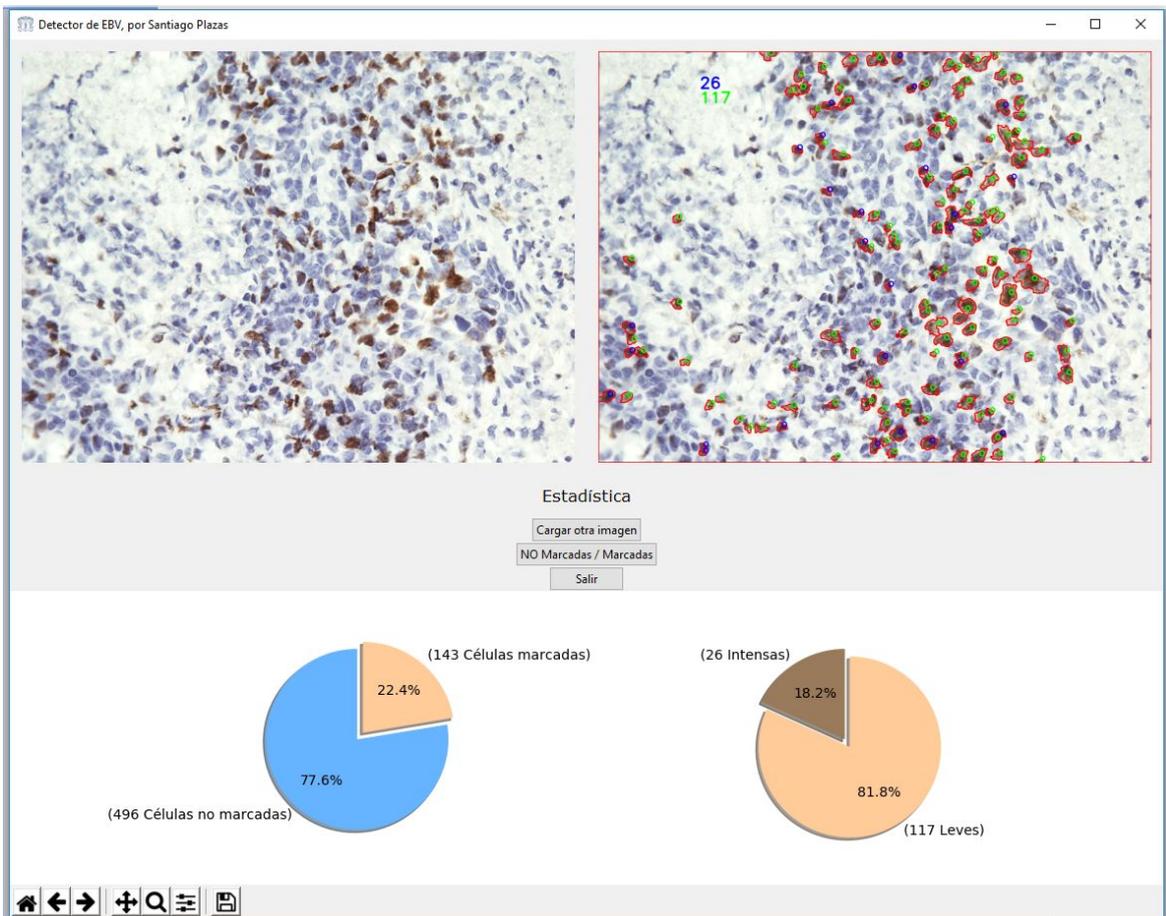


Figura 4.17. Interfaz de la herramienta. Se muestra primero en el panel de carga de imagen, dos botones, uno para cargar imagen y otro para salir, después, en el segundo panel, se observa la imagen original en la parte superior izquierda, la imagen procesada en la parte superior derecha, los porcentajes calculados para Células no marcadas / células marcadas en la parte inferior izquierda y los porcentajes calculados para marcaciones intensas/leves en la parte inferior derecha. En el centro, se encuentran tres botones, un botón para cargar nuevas imágenes, un botón para alternar las imágenes procesadas que se muestran y un botón para salir.

4.8 Evaluación

En la Figura 4.18 se observan los resultados obtenidos tras aplicar la primera encuesta (Tabla 3.5, subcapítulo 3.8) a los patólogos. Se encontró que el número total de células hallado por los patólogos varió entre 430 y 963, 124% de diferencia en los extremos y 66% en los valores medios de los cuartiles 2 y 3, por lo tanto se determinó que existe una discordancia significativa entre observadores, para el número de células marcadas, esta diferencia entre extremos se incrementa a 184%, sin embargo los datos de los cuartiles 2 y 3 se concentran en valores muy cercanos con una diferencia del 20%, por lo que se determinó que no existe una discordancia significativa entre observadores para determinar el número de células marcadas. Por otro lado se encontró gran variabilidad entre lo que los patólogos determinaron como marcaciones intensas

y marcaciones leves, ya que para lo que un patólogo es intenso para otro puede ser considerado leve. Además en la Figura 4.18, se compara directamente los valores hallados por los patólogos con el valor hallado por la herramienta y se determina que este último se encuentra dentro de los rangos aceptados.

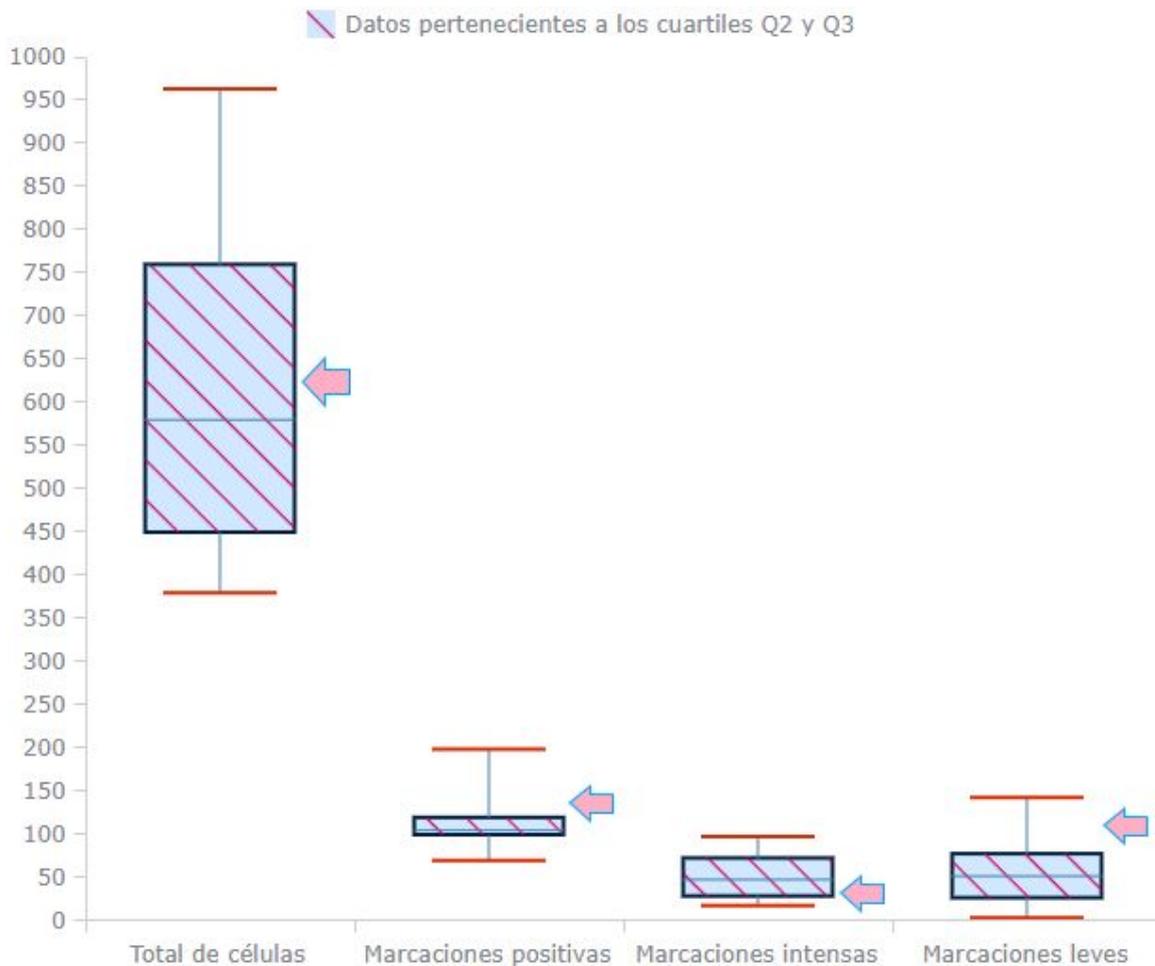


Figura 4.18. Resultados de la primera encuesta (Tabla 3.5) presentados en un diagrama de cajas y bigotes³. Se observa la diferencia inter-observacional en el conteo de células, los bigotes engloban el rango de valores hallados por los patólogos, en las cajas se encuentran los valores correspondientes a los cuartiles Q2 y Q3. Las flechas rosadas indican el número de células detectadas por la herramienta, en comparación con los valores indicados por los patólogos. En orden, la herramienta encontró 639 células totales, 143 células marcadas de café, 26 intensas y 117 leves.

³ Una gráfica de cajas y bigotes consiste en rectángulos cuyos lados más largos representan el recorrido intercuartílico. Dichos rectángulos están divididos por un segmento horizontal que indica la posición de la mediana y por lo tanto su relación con los cuartiles primero y tercero. Las líneas que sobresalen de la caja se llaman bigotes y denotan los valores máximos y mínimos. Este diagrama se utiliza para identificar a) los extremos, b) la mediana, c) los cuartiles, d) el rango intercuartil y e) los valores atípicos (si existen).

En la Figura 4.19 se presentan los resultados de la segunda encuesta (Tabla 3.6, subcapítulo 3.8), donde se observa que el gremio de patólogos en unanimidad considera que la herramienta es fácil de usar, funciona adecuadamente y la utilizaría en un futuro cercano.

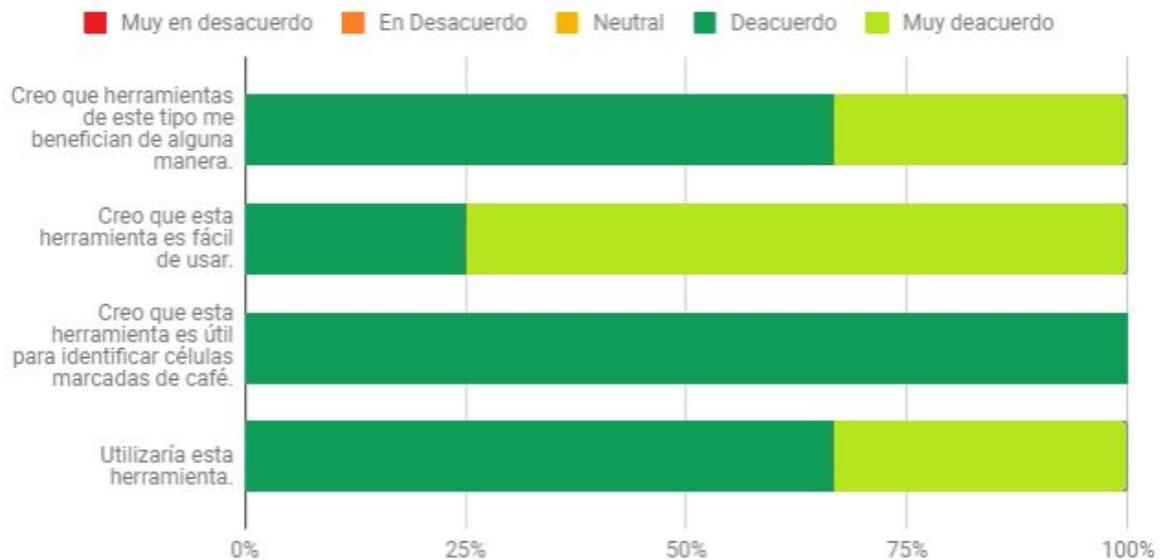


Figura 4.19. Resultado de la segunda encuesta (Tabla 3.6). *El gremio de patólogos tiene un buen recibimiento de la herramienta y está abierto a utilizar nuevas tecnologías.*

Por otro lado de los conversatorios realizados se puede recalcar que muchos patólogos no conocían de la existencia de dichas herramientas y esperan que el programa de Ingeniería física trabaje más de cerca con el departamento de patología en el desarrollo de nuevas herramientas y consideran que la información de tipo cuantitativo que aporta la herramienta para estimar la expresión de intensidad, número de células y colores se pueden utilizar para investigar más a fondo la relación que tienen condiciones externas (e.g. fumar, edad, género) con estadísticos que definen una biopsia. Además se señaló que esta herramienta podría adaptarse para identificar tinciones inmunohistoquímicas en general y cuantificar otros tipos de marcadores. Sin embargo, se enfatizó en que para realizar conclusiones y diagnósticos se requiere de la integración de los hallazgos del patólogo para cada caso.

5. CONCLUSIONES Y TRABAJOS FUTUROS

5.1 Conclusiones

Basándose en los resultados de esta investigación y considerando el análisis de la literatura se puede concluir que:

1. Se caracterizó el microambiente tumoral basándose en la experiencia del GYEi, mediante procesamiento digital de imágenes, utilizando en combinación los algoritmos de Otsu, Watershed y transferencia de color.
2. Utilizando la metodología incremental planteada, se determinó que la configuración de sensor y montaje encontrados resultaron satisfactorios.
3. Se determinó que incrementar el tiempo de fijación de la muestra, como la estandarización del procesado de las mismas es esencial para obtener resultados satisfactorios con el procesamiento digital de imágenes.
4. La herramienta desarrollada cumplió satisfactoriamente la evaluación de funcionalidad y experiencia de usuario, según la aplicación de las encuestas a los patólogos.
5. Para el correcto funcionamiento de la herramienta se debe seguir una realización de cortes estándar (4 μ m) del bloque parafinado, seguimiento estricto del protocolo CISH, adquisición de imágenes utilizando una cámara acoplada internamente a un microscopio de alta resolución Nikon eclipse 80i, con un ocular de 10X y objetivo de 40X, sin filtros añadidos, iluminación y apertura del diafragma de seis puntos, velocidad de obturación 1/60s y contraste *enhanced* de la cámara, con balance de blancos automático.
6. Se identificó una limitante importante en la falta de estandarización para realizar cortes en tejidos parafinados, ya que la variación del ancho de los cortes cambia radicalmente la apariencia visual del tejido al microscopio, pues se presenta solapamiento de capas celulares.
7. Al modificar el algoritmo de transferencia de color de Reinhard se obtuvo un potente identificador de colores que sirvió de columna vertebral a la herramienta desarrollada.

8. Al utilizar la cámara acoplada al microscopio se obtuvieron los mejores resultados, sin embargo, los microscopios con cámara acoplada se encuentran fuera del alcance de muchos debido a su elevado precio. A pesar de los resultados obtenidos la captura de la información es muy lenta, para digitalizar un punto de microarreglo se requieren entre 30 y 50 fotografías, teniendo en cuenta el tiempo que se demora en tomar dichas fotografías un patólogo pudiera estar analizando muestras manualmente.
9. Se concluye que el éxito o fracaso de un proyecto interdisciplinar como este, que involucró ingeniería y medicina, depende de poder identificar las necesidades y relaciones con el entorno, que se logra participando, desde la gestación de los proyectos en los procesos propios de las disciplinas involucradas.

5.2 Trabajos futuros

En general el procesado de los bloques parafinados es largo y repetitivo, actualmente se tiene que ser muy riguroso en el control del tiempo que permanecen las muestras en contacto con reactivos y en la medición de cantidades a aplicar, teniendo en cuenta el elevado precio de la mayoría de los kits de tinción. Por más riguroso que se sea, siempre se va a depender de la pericia humana, por lo tanto se propone diseñar y construir una estación de procesado automático de bloques parafinados, que brinde mayor control de las variables ya mencionadas y brinde cortes de tejido más uniformes.

Siguiendo el proceso de adquisición de datos, muchos laboratorios de punta en el mundo utilizan lo que se conoce como *whole slide scanner*, sin embargo esta tecnología tiene un precio muy elevado para un laboratorio ubicado en el Cauca, por lo que se propone adecuar los microscopios más utilizados, para conseguir mejores imágenes digitales y de forma rápida, posiblemente utilizando tarjetas de prototipado rápido para mover el carro del microscopio, ya que incluso con la utilización de microscopios con cámara integrada el proceso de adquisición es muy lento.

En la parte de procesamiento de imágenes, por la forma como se desarrolló, el principal desafío fue la plantación de marcadores para aplicar el algoritmo de *watershed*, por lo tanto se propone investigar en mejores formas de plantar dichos marcadores. Por otro lado la base de la herramienta es el algoritmo de transferencia de Reinhard, que incluso puede superar a Otsu, aunque no este originalmente pensado para umbralizar, por lo tanto se sugiere investigar en la

aplicación y modificación de este algoritmo en éste área pues promete facilitar procesar todo tipo de imágenes en color. Además se propone investigar en la forma de generalizar esta herramienta para su uso con diferentes tipos de tinciones diferenciales, ya que la apariencia final de las imágenes es bastante parecida.

Por otro lado, la vinculación de patólogos al proyecto, hace parte del mecanismo de obtención de indicadores (verificación) que permiten comparar a la herramienta desarrollada con aquellas nombradas en la literatura. La función del patólogo dentro de la verificación es fundamental y consiste en la clasificación manual una a una de las células identificadas por el software, en un largo y lento proceso que involucra a más de 10.000 células. Por lo tanto, se propone que en la presentación de proyectos de este tipo a entes financiadores, como Colciencias se incluyan montos para disponer del tiempo de los verificadores.

Como punto final se propone fortalecer el vínculo existente entre el programa de Ingeniería Física y el Departamento de Patología de la Universidad del Cauca, ya que en la experiencia adquirida en el trabajo con GYEi es evidente que existe una relación mutualista, donde a través de la aplicación de la física y herramientas computacionales se logra fortalecer la investigación y aplicación de la medicina.

Siguiendo el proceso de adquisición de datos, muchos laboratorios de punta en el mundo utilizan lo que se conoce como *whole slide scanner*, sin embargo esta tecnología tiene un precio muy elevado para un laboratorio ubicado en el Cauca, por lo que se propone adecuar los microscopios más utilizados, para conseguir mejores imágenes digitales y de forma rápida, posiblemente utilizando tarjetas de prototipado rápido para mover el carro del microscopio, ya que incluso con la utilización de microscopios con cámara integrada el proceso de adquisición es muy lento.

En la parte de procesamiento de imágenes, por la forma como se desarrolló, el principal desafío fue la plantación de marcadores para aplicar el algoritmo de *watershed*, por lo tanto se propone investigar en mejores formas de plantar dichos marcadores. Por otro lado la base de la herramienta es el algoritmo de transferencia de Reinhard, que incluso puede superar a otsu, aunque no este originalmente pensado para umbralizar, por lo tanto se sugiere investigar en la aplicación y modificación de este algoritmo en éste área pues promete facilitar procesar todo tipo de imágenes en color. Además se propone investigar en la forma de generalizar esta herramienta para su uso con diferentes tipos de tinciones diferenciales, ya que la apariencia final de las imágenes es bastante parecida.

Por otro lado, la vinculación de patólogos al proyecto, hace parte del mecanismo de obtención de indicadores (verificación) que permiten comparar a la herramienta desarrollada con aquellas nombradas en la literatura. La función del patólogo dentro de la verificación es fundamental y consiste en la clasificación manual una a una de las células identificadas por el software, en un largo y lento proceso que involucra a más de 10.000 células. Por lo tanto, se propone que en la presentación de proyectos de este tipo a entes financiadores, como Colciencias se incluyan montos para disponer del tiempo de los verificadores.

Como punto final se propone fortalecer el vínculo existente entre el programa de Ingeniería Física y el Departamento de Patología de la Universidad del Cauca, ya que en la experiencia adquirida en el trabajo con GYEi es evidente que existe una relación mutualista, donde a través de la aplicación de la física y herramientas computacionales se logra fortalecer la investigación y aplicación de la medicina.

6. REFERENCIAS BIBLIOGRÁFICAS

1. Irshad H, Veillard A, Roux L, Racoceanu D. Methods for nuclei detection, segmentation, and classification in digital histopathology: a review—current status and future potential. *IEEE Rev Biomed Eng* [Internet]. 2014;7:97–114. Available from: http://daniraco.free.fr/pubs/Journals/irshad2014ieee_rbme.pdf
2. Cancer [Internet]. World Health Organization. [cited 2018 Sep 13]. Available from: <http://www.who.int/news-room/fact-sheets/detail/cancer>
3. Khan G, Hashim MJ. Global burden of deaths from Epstein-Barr virus attributable malignancies 1990-2010. *Infect Agent Cancer* [Internet]. 2014 Nov 17;9(1):38. Available from: <http://dx.doi.org/10.1186/1750-9378-9-38>
4. Casa Editorial El Tiempo. Cáncer de estómago, el más mortal de Colombia [Internet]. *El Tiempo*. 2016 [cited 2018 Sep 13]. Available from: <https://www.eltiempo.com/vida/salud/cifras-de-cancer-de-estomago-en-colombia-37496>
5. De Castro Ramos FJ. Defunciones no fetales 2016 [Internet]. [cited 2018 Sep 13]. Available from: <http://www.dane.gov.co/index.php/estadisticas-por-tema/salud/nacimientos-y-defunciones/defunciones-no-fetales/defunciones-no-fetales-2016>
6. Diagnosis [Internet]. National Cancer Institute. 2015 [cited 2018 Oct 16]. Available from: <https://www.cancer.gov/about-cancer/diagnosis-staging/diagnosis>
7. Sheffield BS. Immunohistochemistry as a Practical Tool in Molecular Pathology. *Arch Pathol Lab Med* [Internet]. 2016 Aug;140(8):766–9. Available from: <http://dx.doi.org/10.5858/arpa.2015-0453-RA>
8. March J, Hand M, Truong A, Grossman D. Practical application of new technologies for melanoma diagnosis: Part II. Molecular approaches. *J Am Acad Dermatol* [Internet]. 2015 Jun;72(6):943–58; quiz 959–60. Available from: <http://dx.doi.org/10.1016/j.jaad.2015.02.1140>
9. Valasek MA, Thung I, Gollapalle E, Hodkoff AA, Kelly KJ, Baumgartner JM, et al. Overinterpretation is common in pathological diagnosis of appendix cancer during patient referral for oncologic care. *PLoS One* [Internet]. 2017 Jun 7 [cited 2018 Oct 12];12(6):e0179216. Available from: <https://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0179216&type=printable>
10. Elmore JG, Longton GM, Carney PA, Geller BM, Onega T, Tosteson ANA, et al. Diagnostic concordance among pathologists interpreting breast biopsy specimens. *JAMA* [Internet]. 2015 Mar 17;313(11):1122–32. Available from: <http://dx.doi.org/10.1001/jama.2015.1405>
11. P. Ramírez-Mendoza, A. Ángeles-Ángeles, J. Aguirre-García, R. Herrera-Goepfert, U. Ángeles-Garay, J. González-Angulo. Concordancia en el

- diagnóstico de atrofia gástrica entre patólogos según el sistema OLGA [Internet]. *Revista de Gastroenterología de México*. 2009 [cited 2018 Oct 12]. Available from: <http://revistagastroenterologiamexico.org/index.php?p=revista&tipo=pdf-simple&pii=X037509060949893X>
12. F. V. Responsabilidad Profesional del Patólogo [Internet]. Biblioteca Virtual NOBLE. 2014 [cited 2018 Oct 12]. Available from: http://www.noble-arp.com/src/img_up/10022015.0.pdf
 13. Sissons HA. Agreement and disagreement between pathologists in histological diagnosis. *Postgrad Med J* [Internet]. 1975 Oct 1 [cited 2018 Oct 12];51(600):685–9. Available from: <https://pmj.bmj.com/content/51/600/685.abstract>
 14. Elena K, Eduardo BL, Carlos BJ, Jesús A-G, Lourdes F-L, del Carmen N-VM, et al. Reproducibilidad del diagnóstico histopatológico de lesiones precursoras del carcinoma gástrico en tres países latinoamericanos. *Salud pública Méx* [Internet]. 2010 Oct;52(5):386–90. Available from: http://www.scielo.org/scielo.php?script=sci_arttext&pid=S0036-36342010000500005&lng=es&nrm=iso&tlng=es
 15. McClatchy DM, Rizzo EJ, Wells WA, Black CC, Paulsen KD, Kanick SC, et al. Light scattering measured with spatial frequency domain imaging can predict stromal versus epithelial proportions in surgically resected breast tissue. *J Biomed Opt* [Internet]. 2018 Sep;24(7):1–11. Available from: <http://dx.doi.org/10.1117/1.JBO.24.7.071605>
 16. Eraso CAA, Velasco MMP, Lopez JAV, Tellez JL, Bravo LE, Rabal H, et al. Gastric mucosa analysis using speckle patterns: a medical diagnosis alternative. In: 22nd Congress of the International Commission for Optics: Light for the Development of the World [Internet]. International Society for Optics and Photonics; 2011 [cited 2018 Oct 24]. p. 80118I. Available from: <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/8011/80118I/Gastric-mucosa-analysis-using-speckle-patterns--a-medical-diagnosis/10.1117/12.901830.short>
 17. Andrade Eraso CA, Patitno Velasco MM, Vasquez Lopez JA, Eduardo Bravo L, Leon Tellez J, Trivi M, et al. LASER SPECKLE TECHNIQUE: A NOVEL APPROACH FOR THE CLASSIFICATION AND GRADING THE GASTRIC PRECANCEROUS LESIONS. In: ANNALS OF ONCOLOGY. OXFORD UNIV PRESS GREAT CLARENDON ST, OXFORD OX2 6DP, ENGLAND; 2011. p. v45–6.
 18. Wei L, Yin C, Liu JTC. Dual-Axis Confocal Microscopy for Point-of-Care Pathology. *IEEE J Sel Top Quantum Electron* [Internet]. 2019 Jan;25(1):1–10. Available from: <http://dx.doi.org/10.1109/JSTQE.2018.2854572>
 19. Fanaian NK, Cohen C, Waldrop S, Wang J, Shehata BM. Epstein-Barr virus (EBV)-encoded RNA: automated in-situ hybridization (ISH) compared with

- manual ISH and immunohistochemistry for detection of EBV in pediatric lymphoproliferative disorders. *Pediatr Dev Pathol* [Internet]. 2009 May;12(3):195–9. Available from: <http://dx.doi.org/10.2350/07-07-0316.1>
20. Bueno G, Milagro Fernández-Carrobles M, Deniz O, García-Rojo M. New Trends of Emerging Technologies in Digital Pathology. *PAT* [Internet]. 2016 [cited 2018 Oct 14];83(2-3):61–9. Available from: <https://www.karger.com/Article/Abstract/443482>
 21. Cárdenas-Mondragón MG, Torres J, Sánchez-Zauco N, Gómez-Delgado A, Camorlinga-Ponce M, Maldonado-Bernal C, et al. Elevated Levels of Interferon- Are Associated with High Levels of Epstein-Barr Virus Reactivation in Patients with the Intestinal Type of Gastric Cancer. *J Immunol Res* [Internet]. 2017 Nov 14;2017:7069242. Available from: <http://dx.doi.org/10.1155/2017/7069242>
 22. Chang MS, Kim WH. Epstein-Barr Virus in Human Malignancy: A Special Reference to Epstein-Barr Virus associated Gastric Carcinoma. *Cancer Res Treat* [Internet]. 2005;37(5):257. Available from: <http://dx.doi.org/10.4143/crt.2005.37.5.257>
 23. Su X, Ye Z, Wang Z, Long Y, Qiu M, He C. Epstein-Barr virus infection associated with pepsinogens and Helicobacter pylori infection in patients with gastric cancer. *Virus Res* [Internet]. 2018 Sep 2;256:1–5. Available from: <http://dx.doi.org/10.1016/j.virusres.2018.07.017>
 24. Camargo MC, Sivins A, Isajevs S, Folkmanis V, Rudzīte D, Gulley ML, et al. Associations of Epstein-Barr Virus-Positive Gastric Adenocarcinoma with Circulating Mediators of Inflammation and Immune Response. *Cancers* [Internet]. 2018 Aug 23;10(9). Available from: <http://dx.doi.org/10.3390/cancers10090284>
 25. Cho J, Kang M-S, Kim K-M. Epstein-Barr Virus-Associated Gastric Carcinoma and Specific Features of the Accompanying Immune Response. *J Gastric Cancer* [Internet]. 2016;16(1):1. Available from: <http://dx.doi.org/10.5230/jgc.2016.16.1.1>
 26. Abdirad A, Ghaderi-Sohi S, Shuyama K, Koriyama C, Nadimi-Barforoosh H, Emami S, et al. Epstein-Barr virus associated gastric carcinoma: a report from Iran in the last four decades. *Diagn Pathol* [Internet]. 2007 Jul 15;2:25. Available from: <http://dx.doi.org/10.1186/1746-1596-2-25>
 27. Shinozaki-Ushiku A, Kunita A, Fukayama M. Update on Epstein-Barr virus and gastric cancer. *Int J Oncol* [Internet]. 2015;46(4):1421–34. Available from: <https://www.spandidos-publications.com/10.3892/ijo.2015.2856>
 28. Sharif JM, Miswan MF, Ngadi MA, Salam MSH, bin Abdul Jamil MM. Red blood cell segmentation using masking and watershed algorithm: A preliminary study. In: 2012 International Conference on Biomedical Engineering (ICoBE) [Internet]. 2012. p. 258–62. Available from: <http://dx.doi.org/10.1109/ICoBE.2012.6179016>

29. Cao H, Liu H, Song E. A novel algorithm for segmentation of leukocytes in peripheral blood. *Biomed Signal Process Control* [Internet]. 2018;45:10–21. Available from: <http://dx.doi.org/10.1016/j.bspc.2018.05.010>
30. Pinidiyaarachchi A, Wählby C. Seeded Watersheds for Combined Segmentation and Tracking of Cells. In: *Image Analysis and Processing – ICIAP 2005* [Internet]. Springer, Berlin, Heidelberg; 2005 [cited 2018 Oct 15]. p. 336–43. Available from: https://link.springer.com/chapter/10.1007/11553595_41
31. Holcik M, Sonenberg N. Translational control in stress and apoptosis. *Nat Rev Mol Cell Biol* [Internet]. 2005 Apr;6(4):318–27. Available from: <http://dx.doi.org/10.1038/nrm1618>
32. Jemal A, Siegel R, Xu J, Ward E. Cancer statistics, 2010. *CA Cancer J Clin* [Internet]. 2010 Sep;60(5):277–300. Available from: <http://dx.doi.org/10.3322/caac.20073>
33. Lauren P. THE TWO HISTOLOGICAL MAIN TYPES OF GASTRIC CARCINOMA: DIFFUSE AND SO-CALLED INTESTINAL-TYPE CARCINOMA. AN ATTEMPT AT A HISTO-CLINICAL CLASSIFICATION. *Acta Pathol Microbiol Scand* [Internet]. 1965;64:31–49. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/14320675>
34. Uemura N, Okamoto S, Yamamoto S, Matsumura N, Yamaguchi S, Yamakido M, et al. Helicobacter pylori infection and the development of gastric cancer. *N Engl J Med* [Internet]. 2001 Sep 13;345(11):784–9. Available from: <http://dx.doi.org/10.1056/NEJMoa001999>
35. Cancer Genome Atlas Research Network. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature* [Internet]. 2014 Sep 11;513(7517):202–9. Available from: <http://dx.doi.org/10.1038/nature13480>
36. Oliveira C, Pinheiro H, Figueiredo J, Seruca R, Carneiro F. Familial gastric cancer: genetic susceptibility, pathology, and implications for management. *Lancet Oncol* [Internet]. 2015 Feb;16(2):e60–70. Available from: [http://dx.doi.org/10.1016/S1470-2045\(14\)71016-2](http://dx.doi.org/10.1016/S1470-2045(14)71016-2)
37. Goh L-Y, Leow AHR, Goh K-L. Observations on the epidemiology of gastrointestinal and liver cancers in the Asia-Pacific region. *J Dig Dis* [Internet]. 2014 Sep;15(9):463–8. Available from: <http://dx.doi.org/10.1111/1751-2980.12164>
38. Shikata K, Kiyohara Y, Kubo M, Yonemoto K, Ninomiya T, Shirota T, et al. A prospective study of dietary salt intake and gastric cancer incidence in a defined Japanese population: The Hisayama study. *International Journal of Cancer* [Internet]. 2006;119(1):196–201. Available from: <http://dx.doi.org/10.1002/ijc.21822>
39. Sung HG, Kobayashi Y, Chang J, Ha A, Hwang IH, Ha JK. Low Ruminant pH Reduces Dietary Fiber Digestion via Reduced Microbial Attachment. *Asian-australas J Anim Sci* [Internet]. 2006;20(2):200–7. Available from:

- <http://dx.doi.org/10.5713/ajas.2007.200>
40. Jemal A, Ward E, Wu X, Martin HJ, McLaughlin CC, Thun MJ. Geographic patterns of prostate cancer mortality and variations in access to medical care in the United States. *Cancer Epidemiol Biomarkers Prev* [Internet]. 2005 Mar;14(3):590–5. Available from: <http://dx.doi.org/10.1158/1055-9965.EPI-04-0522>
 41. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* [Internet]. 2015 Mar 1;136(5):E359–86. Available from: <http://dx.doi.org/10.1002/ijc.29210>
 42. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* [Internet]. 2018 Nov;68(6):394–424. Available from: <http://dx.doi.org/10.3322/caac.21492>
 43. Website [Internet]. [cited 2018 Oct 16]. Available from: [<https://www.cancerquest.org/index.php/es/para-los-pacientes/cancer-porti-po/cancer-de-estomago>]
 44. NCI Dictionary of Cancer Terms [Internet]. National Cancer Institute. 2011 [cited 2018 Oct 24]. Available from: <https://www.cancer.gov/publications/dictionaries/cancer-terms/def/adenocarcinoma>
 45. Morgan DR, Torres J, Sexton R, Herrero R, Salazar-Martínez E, Greenberg ER, et al. Risk of recurrent *Helicobacter pylori* infection 1 year after initial eradication therapy in 7 Latin American communities. *JAMA* [Internet]. 2013 Feb 13;309(6):578–86. Available from: <http://dx.doi.org/10.1001/jama.2013.311>
 46. Medina-Ortega AP, Universidad del Cauca, López-Valencia D, Mosquera-Monje SL, Mora-Obando DL, Dueñas-Cuéllar RA, et al. Virus de Epstein-Barr y su relación con el desarrollo del cáncer. *IATREIA* [Internet]. 2017;30(2):131–45. Available from: <http://dx.doi.org/10.17533/udea.iatreia.v30n2a03>
 47. Odumade OA, Hogquist KA, Balfour HH Jr. Progress and problems in understanding and managing primary Epstein-Barr virus infections. *Clin Microbiol Rev* [Internet]. 2011 Jan;24(1):193–209. Available from: <http://dx.doi.org/10.1128/CMR.00044-10>
 48. Münz C, Moormann A. Immune escape by Epstein–Barr virus associated malignancies. *Semin Cancer Biol* [Internet]. 2008;18(6):381–7. Available from: <http://dx.doi.org/10.1016/j.semcancer.2008.10.002>
 49. Fukayama M, Hayashi Y, Iwasaki Y, Chong J, Ooba T, Takizawa T, et al. Epstein-Barr virus-associated gastric carcinoma and Epstein-Barr virus infection of the stomach. *Lab Invest* [Internet]. 1994 Jul;71(1):73–81. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/8041121>
 50. Ko Y-H. EBV and human cancer. *Exp Mol Med* [Internet].

- 2015;47(1):e130–e130. Available from:
<http://dx.doi.org/10.1038/emm.2014.109>
51. Chesnokova LS, Hutt-Fletcher LM. Epstein-Barr virus infection mechanisms. *Chin J Cancer* [Internet]. 2014; Available from:
<http://dx.doi.org/10.5732/cjc.014.10168>
 52. Hutt-Fletcher LM. Epstein-Barr virus replicating in epithelial cells. *Proc Natl Acad Sci U S A* [Internet]. 2014 Nov 18;111(46):16242–3. Available from:
<http://dx.doi.org/10.1073/pnas.1418974111>
 53. Thorley-Lawson DA, Hawkins JB, Tracy SI, Shapiro M. The pathogenesis of Epstein-Barr virus persistent infection. *Curr Opin Virol* [Internet]. 2013;3(3):227–32. Available from:
<http://dx.doi.org/10.1016/j.coviro.2013.04.005>
 54. Amon W, Farrell PJ. Reactivation of Epstein-Barr virus from latency. *Rev Med Virol* [Internet]. 2005;15(3):149–56. Available from:
<https://onlinelibrary.wiley.com/doi/abs/10.1002/rmv.456>
 55. Calderwood MA, Venkatesan K, Xing L, Chase MR, Vazquez A, Holthaus AM, et al. Epstein-Barr virus and virus human protein interaction maps. *Proc Natl Acad Sci U S A* [Internet]. 2007 May 1 [cited 2018 Oct 17];104(18):7606–11. Available from: <http://www.pnas.org/content/104/18/7606.short>
 56. Liu X, Liu J, Qiu H, Kong P, Chen S, Li W, et al. Prognostic significance of Epstein-Barr virus infection in gastric cancer: a meta-analysis. *BMC Cancer* [Internet]. 2015 Oct 24;15:782. Available from:
<http://dx.doi.org/10.1186/s12885-015-1813-9>
 57. Raab-Traub N. Novel mechanisms of EBV-induced oncogenesis. *Curr Opin Virol* [Internet]. 2012 Aug;2(4):453–8. Available from:
<http://dx.doi.org/10.1016/j.coviro.2012.07.001>
 58. Truong CD, Feng W, Li W, Khoury T, Li Q, Alrawi S, et al. Characteristics of Epstein-Barr virus-associated gastric cancer: a study of 235 cases at a comprehensive cancer center in USA. *J Exp Clin Cancer Res* [Internet]. 2009;28(1):14. Available from:
<https://jeccr.biomedcentral.com/articles/10.1186/1756-9966-28-14>
 59. Portal D, Zhou H, Zhao B, Kharchenko PV, Lowry E, Wong L, et al. Epstein-Barr virus nuclear antigen leader protein localizes to promoters and enhancers with cell transcription factors and EBNA2. *Proc Natl Acad Sci U S A* [Internet]. 2013 Nov 12;110(46):18537–42. Available from:
<http://dx.doi.org/10.1073/pnas.1317608110>
 60. Carrascal E, Tokunaga M, Akiba S, Eizuru Y, Fujiyama C, Shinkura R, et al. Adenocarcinoma gástrico asociado con el virus de Epstein-Barr en Cali. *Colombia Médica* [Internet]. 1999 [cited 2018 Nov 27];30(3):127–31. Available from:
<http://colombiamedica.univalle.edu.co/index.php/comedica/article/view/135>
 61. Ambrosio MR, Leoncini L. Epidemiology of Epstein-Barr Virus and

- Mechanisms of Carcinogenesis. In: Tropical Hemato-Oncology [Internet]. 2015. p. 127–41. Available from: http://dx.doi.org/10.1007/978-3-319-18257-5_15
62. Abe H, Kaneda A, Fukayama M. Epstein-Barr Virus-Associated Gastric Carcinoma: Use of Host Cell Machineries and Somatic Gene Mutations. Pathobiology [Internet]. 2015;82(5):212–23. Available from: <http://dx.doi.org/10.1159/000434683>
 63. Lopes LF, Miyazawa KWR, de Almeida ERD, Serafim KGG, de Almeida Gualtieri K, Costa IC, et al. Epstein–Barr Virus (EBV) MicroRNAs: Involvement in Cancer Pathogenesis and Immunopathology. Int Rev Immunol [Internet]. 2013;32(3):271–81. Available from: <http://dx.doi.org/10.3109/08830185.2012.748053>
 64. Thorley-Lawson DA. Epstein-Barr virus: exploiting the immune system. Nat Rev Immunol [Internet]. 2001 Oct;1(1):75–82. Available from: <http://dx.doi.org/10.1038/35095584>
 65. Young B, Woodford P, O’Dowd G. Wheater’s Functional Histology E-Book: A Text and Colour Atlas [Internet]. Elsevier Health Sciences; 2013. 464 p. Available from: <https://market.android.com/details?id=book-HUIYAQAAQBAJ>
 66. Corthell JT. Basic Molecular Protocols in Neuroscience: Tips, Tricks, and Pitfalls [Internet]. Academic Press; 2014. 124 p. Available from: <https://market.android.com/details?id=book-FYt7AwAAQBAJ>
 67. Grody WW, Nakamura RM. Molecular Diagnostics: Techniques and Applications for the Clinical Laboratory [Internet]. Elsevier/Academic Press; 2010. 484 p. Available from: <https://market.android.com/details?id=book-hjOmlAEACAAJ>
 68. Sawayama H, Ishimoto T, Baba H. Microenvironment in the pathogenesis of gastric cancer metastasis. JCMT [Internet]. 2018 Feb 28;4(2):10. Available from: <http://jcmtjournal.com/article/view/2407>
 69. van Beek J, zur Hausen A, Snel SN, Berkhof J, Kranenbarg EK, van de Velde CJH, et al. Morphological evidence of an activated cytotoxic T-cell infiltrate in EBV-positive gastric carcinoma preventing lymph node metastases. Am J Surg Pathol [Internet]. 2006 Jan;30(1):59–65. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/16330943>
 70. Tan GW, Visser L, Tan LP, van den Berg A, Diepstra A. The Microenvironment in Epstein-Barr Virus-Associated Malignancies. Pathogens [Internet]. 2018 Apr 13;7(2). Available from: <http://dx.doi.org/10.3390/pathogens7020040>
 71. Yoffey JM, Courtice FC, Others. Lymphatics, lymph and the lymphomyeloid complex. Lymphatics, lymph and the lymphomyeloid complex [Internet]. 1970; Available from: <https://www.cabdirect.org/cabdirect/abstract/19712203682>
 72. Ford WL, Gowans JL. The traffic of lymphocytes. Semin Hematol [Internet]. 1969 Jan;6(1):67–83. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/4886828>

73. Caligiuri M, Levi MM, Kaushansky K, Lichtman MA, Prchal J, Burns LJ, et al. Williams Hematology, 9E [Internet]. McGraw-Hill Education / Medical; 2015. 2528 p. Available from: https://books.google.com/books/about/Williams_Hematology_9E.html?hl=&id=6MuXnQAACAAJ
74. Kierszenbaum AL, Tres L. Histology and Cell Biology: An Introduction to Pathology [Internet]. Elsevier Health Sciences; 2015. 720 p. Available from: <https://market.android.com/details?id=book-i-lyCQAAQBAJ>
75. Strokotov DI, Yurkin MA, Gilev KV, van Bockstaele DR, Hoekstra AG, Rubtsov NB, et al. Is there a difference between T- and B-lymphocyte morphology? J Biomed Opt [Internet]. 2009 Nov;14(6):064036. Available from: <http://dx.doi.org/10.1117/1.3275471>
76. Glynn LE, Houck JC, Weissmann G. Handbook of inflammation [Internet]. Elsevier/North-Holland Biomedical Press; 1985. 496 p. Available from: <https://market.android.com/details?id=book-8ftJAQAIAAJ>
77. Maslak P. Mast cell leukemia. Blood [Internet]. 2003 Feb 1;101(3):789. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/12564414>
78. Yong LC. The mast cell: origin, morphology, distribution, and function. Exp Toxicol Pathol [Internet]. 1997 Dec;49(6):409–24. Available from: [http://dx.doi.org/10.1016/S0940-2993\(97\)80129-7](http://dx.doi.org/10.1016/S0940-2993(97)80129-7)
79. Haas M, Büttner M, Rau TT, Fietkau R, Grabenbauer GG, Distel LV. Inflammation in gastric adenocarcinoma of the cardia: how do EBV infection, Her2 amplification and cancer progression influence tumor-infiltrating lymphocytes? Virchows Arch [Internet]. 2011 Apr;458(4):403–11. Available from: <http://dx.doi.org/10.1007/s00428-011-1058-1>
80. Ichimura T, Abe H, Morikawa T, Yamashita H, Ishikawa S, Ushiku T, et al. Low density of CD204-positive M2-type tumor-associated macrophages in Epstein-Barr virus-associated gastric cancer: a clinicopathologic study with digital image analysis. Hum Pathol [Internet]. 2016 Oct;56:74–80. Available from: <http://dx.doi.org/10.1016/j.humpath.2016.06.002>
81. Escalera Hueso A de la. Visión por computador: fundamentos y métodos [Internet]. 2001. Available from: <http://www.sidalc.net/cgi-bin/wxis.exe/?IsisScript=AGRIUAN.xis&method=post&formato=2&cantidad=1&expresion=mfn=031674>
82. Gonzalez RC, Woods RE. Digital Image Processing, Global Edition [Internet]. Pearson Higher Education; 2017. 1024 p. Available from: https://books.google.com/books/about/Digital_Image_Processing_Global_Edition.html?hl=&id=XmZvtAEACAAJ
83. Gonzalez RC, Woods RE. Digital Image Processing [Internet]. Pearson; 2017. 1168 p. Available from: <https://market.android.com/details?id=book-0F05vgAACAAJ>
84. Chityala R, Pudipeddi S. Image processing and acquisition using Python

- [Internet]. Chapman and Hall/CRC; 2014. Available from: <https://www.taylorfrancis.com/books/9781466583764>
85. Russ J, Neal F. The Image Processing Handbook, Seventh Edition [Internet]. CRC Press; 2015. Available from: <https://www.taylorfrancis.com/books/9781498740289>
 86. Petrou M, Petrou C. Image Processing: The Fundamentals [Internet]. John Wiley & Sons; 2010. 818 p. Available from: <https://market.android.com/details?id=book-w3BpSlxN9ZYC>
 87. Méndez AJ, Tahoces PG, Lado MJ, Souto M, Vidal JJ. Computer-aided diagnosis: Automatic detection of malignant masses in digitized mammograms. *Med Phys* [Internet]. 1998;25(6):957–64. Available from: <https://aapm.onlinelibrary.wiley.com/doi/abs/10.1118/1.598274>
 88. Tang J, Rangayyan RM, Xu J, El Naqa I, Yang Y. Computer-aided detection and diagnosis of breast cancer with mammography: recent advances. *IEEE Trans Inf Technol Biomed* [Internet]. 2009 Mar;13(2):236–51. Available from: <http://dx.doi.org/10.1109/TITB.2008.2009441>
 89. Wee YTF, Alkaff SMF, Lim JCT, Loh JJH, Hilmy MH, Ong C, et al. An integrated automated multispectral imaging technique that simultaneously detects and quantitates viral RNA and immune cell protein markers in fixed sections from Epstein-Barr virus-related tumours. *Ann Diagn Pathol* [Internet]. 2018 Sep 11;37:12–9. Available from: <http://dx.doi.org/10.1016/j.anndiagpath.2018.09.002>
 90. Sudharshan PJ, Petitjean C, Spanhol F, Oliveira LE, Heutte L, Honeine P. Multiple instance learning for histopathological breast cancer image classification. *Expert Syst Appl* [Internet]. 2019 Mar 1;117:103–11. Available from: <http://www.sciencedirect.com/science/article/pii/S0957417418306262>
 91. Benzebouchi NE, Azizi N, Ayadi K. A Computer-Aided Diagnosis System for Breast Cancer Using Deep Convolutional Neural Networks. In: Behera HS, Nayak J, Naik B, Abraham A, editors. *Computational Intelligence in Data Mining* [Internet]. Singapore: Springer Singapore; 2019. p. 583–93. (Advances in Intelligent Systems and Computing; vol. 711). Available from: http://link.springer.com/10.1007/978-981-10-8055-5_52
 92. Mazo C, Bernal J, Trujillo M, Alegre E. Transfer learning for classification of cardiovascular tissues in histological images. *Comput Methods Programs Biomed* [Internet]. 2018 Oct 1;165:69–76. Available from: <http://www.sciencedirect.com/science/article/pii/S0169260718305297>
 93. Sonigo C, Jankowski S, Yoo O, Trassard O, Bousquet N, Grynberg M, et al. High-throughput ovarian follicle counting by an innovative deep learning approach. *Sci Rep* [Internet]. 2018 Sep 10;8(1):13499. Available from: <http://dx.doi.org/10.1038/s41598-018-31883-8>
 94. Huang P-W, Lai Y-H. Effective segmentation and classification for HCC biopsy

- images. *Pattern Recognit* [Internet]. 2010 Apr 1;43(4):1550–63. Available from:
<http://www.sciencedirect.com/science/article/pii/S0031320309003872>
95. Di Cataldo S, Ficarra E, Acquaviva A, Macii E. Automated segmentation of tissue images for computerized IHC analysis. *Comput Methods Programs Biomed* [Internet]. 2010 Oct;100(1):1–15. Available from:
<http://dx.doi.org/10.1016/j.cmpb.2010.02.002>
 96. Nguyen K, Jain AK, Sabata B. Prostate cancer detection: Fusion of cytological and textural features. *J Pathol Inform* [Internet]. 2011;2:S3. Available from:
<http://dx.doi.org/10.4103/2153-3539.92030>
 97. Piórkowski A, Gertych A. Color Normalization Approach to Adjust Nuclei Segmentation in Images of Hematoxylin and Eosin Stained Tissue. In: *Information Technology in Biomedicine* [Internet]. Springer International Publishing; 2019. p. 393–406. Available from:
http://dx.doi.org/10.1007/978-3-319-91211-0_35
 98. Can A, Bello M, Cline HE, Tao X, Ginty F, Sood A, et al. Multi-modal imaging of histological tissue sections. In: *2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro* [Internet]. 2008. p. 288–91. Available from: <http://dx.doi.org/10.1109/ISBI.2008.4540989>
 99. Yang L, Meer P, Foran DJ. Unsupervised segmentation based on robust estimation and color active contour models. *IEEE Trans Inf Technol Biomed* [Internet]. 2005 Sep;9(3):475–86. Available from:
<https://www.ncbi.nlm.nih.gov/pubmed/16167702>
 100. Khan AUM, Torelli A, Wolf I, Gretz N. AutoCellSeg: robust automatic colony forming unit (CFU)/cell analysis using adaptive image segmentation and easy-to-use post-editing techniques. *Sci Rep* [Internet]. 2018 May 8;8(1):7302. Available from: <http://dx.doi.org/10.1038/s41598-018-24916-9>
 101. Rodenacker K, Bengtsson E. A feature set for cytometry on digitized microscopic images. *Anal Cell Pathol* [Internet]. 2003;25(1):1–36. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/12590175>
 102. Kuncheva LI, Whitaker CJ. Measures of Diversity in Classifier Ensembles and Their Relationship with the Ensemble Accuracy. *Mach Learn* [Internet]. 2003 May 1;51(2):181–207. Available from:
<https://doi.org/10.1023/A:1022859003006>
 103. Rajpoot K, Rajpoot N. SVM Optimization for Hyperspectral Colon Tissue Cell Classification. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2004* [Internet]. Springer Berlin Heidelberg; 2004. p. 829–37. Available from: http://dx.doi.org/10.1007/978-3-540-30136-3_101
 104. Gurcan MN, Boucheron L, Can A, Madabhushi A, Rajpoot N, Yener B. Histopathological image analysis: A review. *IEEE Rev Biomed Eng* [Internet]. 2009;2:147. Available from:
<https://www.ncbi.nlm.nih.gov/pmc/articles/pmc2910932/>

105. Weind KL, Maier CF, Rutt BK, Moussa M. Invasive carcinomas and fibroadenomas of the breast: comparison of microvessel distributions—implications for imaging modalities. *Radiology* [Internet]. 1998 Aug;208(2):477–83. Available from: <http://dx.doi.org/10.1148/radiology.208.2.9680579>
106. Hamilton PW, Anderson N, Bartels PH, Thompson D. Expert system support using Bayesian belief networks in the diagnosis of fine needle aspiration biopsy specimens of the breast. *J Clin Pathol* [Internet]. 1994 Apr;47(4):329–36. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/8027370>
107. Bartels PH, Thompson D, Bibbo M, Weber JE. Bayesian belief networks in quantitative histopathology. *Anal Quant Cytol Histol* [Internet]. 1992 Dec;14(6):459–73. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/1292445>
108. Reinhard E, Adhikhmin M, Gooch B, Shirley P. Color transfer between images. *IEEE Comput Graph Appl* [Internet]. 2001;21(4):34–41. Available from: <http://dx.doi.org/10.1109/38.946629>
109. Kowal M, Korbicz J. Marked Point Process for Nuclei Detection in Breast Cancer Microscopic Images. In: *Advances in Intelligent Systems and Computing* [Internet]. 2017. p. 230–41. Available from: http://dx.doi.org/10.1007/978-3-319-66905-2_20
110. Skobel M, Kowal M, Korbicz J, Obuchowicz A. Cell Nuclei Segmentation Using Marker-Controlled Watershed and Bayesian Object Recognition. In: *Information Technology in Biomedicine* [Internet]. Springer International Publishing; 2019. p. 407–18. Available from: http://dx.doi.org/10.1007/978-3-319-91211-0_36
111. Höfener H, Homeyer A, Weiss N, Molin J, Lundström CF, Hahn HK. Deep learning nuclei detection: A simple approach can deliver state-of-the-art results. *Comput Med Imaging Graph* [Internet]. 2018 Sep 17;70:43–52. Available from: <http://dx.doi.org/10.1016/j.compmedimag.2018.08.010>
112. Singh R, Sharma M, Bhattacharya M. Segmentation of Heavily Clustered Cell Nuclei in Histopathological Images. In: Tavares JMRS, Natal Jorge RM, editors. *VipIMAGE 2017* [Internet]. Cham: Springer International Publishing; 2018. p. 244–54. (Lecture Notes in Computational Vision and Biomechanics; vol. 27). Available from: http://link.springer.com/10.1007/978-3-319-68195-5_27
113. Al-Kofahi Y, Lassoued W, Lee W, Roysam B. Improved automatic detection and segmentation of cell nuclei in histopathology images. *IEEE Trans Biomed Eng* [Internet]. 2010 Apr;57(4):841–52. Available from: <http://dx.doi.org/10.1109/TBME.2009.2035102>
114. Soille P. *Morphological Image Analysis* [Internet]. 2004. Available from: <http://dx.doi.org/10.1007/978-3-662-05088-0>

115. Romo-Bucheli D, Corredor G, García-Arteaga JD, Arias V, Romero E. Nuclei graph local features for basal cell carcinoma classification in whole slide images. In: 12th International Symposium on Medical Information Processing and Analysis [Internet]. 2017. Available from: <http://dx.doi.org/10.1117/12.2257386>
116. Arevalo J, Cruz-Roa A, Arias V, Romero E, González FA. An unsupervised feature learning framework for basal cell carcinoma image analysis. *Artif Intell Med* [Internet]. 2015 Jun;64(2):131–45. Available from: <http://dx.doi.org/10.1016/j.artmed.2015.04.004>
117. Romo-Bucheli D, Moncayo R, Cruz-Roa A, Romero E. Identifying histological concepts on basal cell carcinoma images using nuclei based sampling and multi-scale descriptors. In: 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI) [Internet]. 2015. Available from: <http://dx.doi.org/10.1109/isbi.2015.7164041>
118. Arevalo J, Cruz-Roa A, González FA. Hybrid image representation learning model with invariant features for basal cell carcinoma detection. In: IX International Seminar on Medical Information Processing and Analysis [Internet]. 2013. Available from: <http://dx.doi.org/10.1117/12.2035530>
119. Cruz-Roa AA, Arevalo Ovalle JE, Madabhushi A, González Osorio FA. A deep learning architecture for image representation, visual interpretability and automated basal-cell carcinoma cancer detection. *Med Image Comput Comput Assist Interv* [Internet]. 2013;16(Pt 2):403–10. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/24579166>
120. Veta M, van Diest PJ, Willems SM, Wang H, Madabhushi A, Cruz-Roa A, et al. Assessment of algorithms for mitosis detection in breast cancer histopathology images. *Med Image Anal* [Internet]. 2015 Feb;20(1):237–48. Available from: <http://dx.doi.org/10.1016/j.media.2014.11.010>
121. Mosquera-Lopez C, Agaian S, Velez-Hoyos A. The development of a multi-stage learning scheme using new tissue descriptors for automatic grading of prostatic carcinoma. In: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) [Internet]. 2014. Available from: <http://dx.doi.org/10.1109/icassp.2014.6854269>
122. Mosquera-Lopez C, Agaian S, Velez-Hoyos A, Thompson I. Computer-Aided Prostate Cancer Diagnosis From Digitized Histopathology: A Review on Texture-Based Systems. *IEEE Rev Biomed Eng* [Internet]. 2015;8:98–113. Available from: <http://dx.doi.org/10.1109/RBME.2014.2340401>
123. Otálora S, Perdomo O, Atzori M, Andersson M, Jacobsson L, Hedlund M, et al. Determining the scale of image patches using a deep learning approach. In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018) [Internet]. 2018. p. 843–6. Available from: <http://dx.doi.org/10.1109/ISBI.2018.8363703>
124. Lu C, Wang X, Prasanna P, Corredor G, Sedor G, Bera K, et al. Feature Driven

- Local Cell Graph (FeDeG): Predicting Overall Survival in Early Stage Lung Cancer. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2018 [Internet]. Springer International Publishing; 2018. p. 407–16. Available from: http://dx.doi.org/10.1007/978-3-030-00934-2_46
125. Moncayo R, Romo-Bucheli D, Romero E. A Grading Strategy for Nuclear Pleomorphism in Histopathological Breast Cancer Images Using a Bag of Features (BOF). In: Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications [Internet]. Springer International Publishing; 2015. p. 75–82. Available from: http://dx.doi.org/10.1007/978-3-319-25751-8_10
 126. Cruz-Roa A, Caicedo JC, González FA. Visual pattern mining in histology image collections using bag of features. *Artif Intell Med* [Internet]. 2011 Jun;52(2):91–106. Available from: <http://dx.doi.org/10.1016/j.artmed.2011.04.010>
 127. Puerto M, Vargas T, Cruz-Roa A. A Digital Pathology application for whole-slide histopathology image analysis based on genetic algorithm and Convolutional Networks. In: 2016 IEEE Latin American Conference on Computational Intelligence (LA-CCI) [Internet]. 2016. p. 1–7. Available from: <http://dx.doi.org/10.1109/LA-CCI.2016.7885738>
 128. Dorado PA, Celis R, Romero E. Color separation of H&E stained samples by linearly projecting the RGB representation onto a custom discriminant surface. In: 12th International Symposium on Medical Information Processing and Analysis [Internet]. International Society for Optics and Photonics; 2017 [cited 2018 Oct 25]. p. 101600P. Available from: <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/10160/1/Color-separation-of-HE-stained-samples-by-linearly-projecting-the/10.1117/12.2256966.short>
 129. Corredor G, Romero E. Learning histopathological regions of interest by fusing bottom-up and top-down information. In: 2015 IEEE International Conference on Image Processing (ICIP) [Internet]. 2015. p. 3200–4. Available from: <http://dx.doi.org/10.1109/ICIP.2015.7351394>
 130. Celis R, Romero E. Unsupervised color normalisation for H and E stained histopathology image analysis. In: 11th International Symposium on Medical Information Processing and Analysis [Internet]. International Society for Optics and Photonics; 2015 [cited 2018 Oct 25]. p. 968104. Available from: <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/9681/1/Unsupervised-color-normalisation-for-H-and-E-stained-histopathology-image/10.1117/12.2211536.short>
 131. Corredor G, Wang X, Lu C, Velcheti V, Romero E, Madabhushi A. A watershed and feature-based approach for automated detection of lymphocytes on lung cancer images. In: Medical Imaging 2018: Digital Pathology [Internet]. International Society for Optics and Photonics; 2018

- [cited 2018 Nov 2]. p. 105810R. Available from: <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/10581/105810R/A-watershed-and-feature-based-approach-for-automated-detection-of-10.1117/12.2293147.short>
132. Mejia M, Rubiano A, Alzate M. Segmentación de imágenes de células cervicales y evaluación de características para detección de lesiones neoplásicas. *Ingeniare Revista chilena de ingeniería* [Internet]. 2016;24(2):274–89. Available from: https://scielo.conicyt.cl/scielo.php?pid=S0718-33052016000200010&script=sci_arttext&tlng=en
 133. Dorado P, Escobar P. Reconocimiento visual artificial para determinar reactividad hormonal en tejidos con cáncer de mama [BSc]. Tellez J, editor. Universidad del Cauca; 2008.
 134. Paredes LF. Sistema asistido de visión artificial para determinar el grado gleason en biopsias diagnosticadas como adenocarcinoma de la próstata en microscopía óptica [Bsc]. Bolaños Bravo HJ, editor. Universidad del Cauca; 2013.
 135. Simayijiang Z, Grimm S. Segmentation with Graph Cuts. Matematikcentrum Lunds Universitet [Online] Available: <http://www.maths.lth.se/matematiklth/personal/petter/rappoter/graph.pdf> [Diakses 8 Mei 2017] [Internet]. Available from: <http://www.maths.lth.se/matematiklth/personal/petter/rappoter/graph.pdf>
 136. Xing F, Yang L. Robust Nucleus/Cell Detection and Segmentation in Digital Pathology and Microscopy Images: A Comprehensive Review. *IEEE Rev Biomed Eng* [Internet]. 2016 Jan 6;9:234–63. Available from: <http://dx.doi.org/10.1109/RBME.2016.2515127>
 137. Kaizaki Y, Sakurai S, Chong J-M, Fukayama M. Atrophic gastritis, Epstein–Barr virus infection, and Epstein–Barr virus-associated gastric carcinoma. *Gastric Cancer* [Internet]. 1999 Aug 1;2(2):101–8. Available from: <https://doi.org/10.1007/s101200050031>
 138. OpenCV: OpenCV modules [Internet]. [cited 2018 Nov 10]. Available from: <https://docs.opencv.org/3.3.1>
 139. Bonilla C, Hurtado Prieto J, Jaramillo Herrera C, Others. La investigación: aproximaciones a la construcción del conocimiento científico [Internet]. 2009. Available from: <http://www.sidalc.net/cgi-bin/wxis.exe/?IsisScript=SUV.xis&method=post&formato=2&cantidad=1&expresion=mfn=013978>
 140. Likert R. A technique for the measurement of attitudes. *Archives of psychology* [Internet]. 1932; Available from: <http://psycnet.apa.org/record/1933-01885-001>
 141. Norman G. Likert scales, levels of measurement and the ‘laws’ of statistics. *Adv Health Sci Educ Theory Pract* [Internet]. 2010 Dec 1;15(5):625–32.

Available from: <https://doi.org/10.1007/s10459-010-9222-y>

142. Abbas AK, Lichtman AHH, Pillai S. Cellular and Molecular Immunology, Updated Edition [Internet]. Elsevier Health Sciences; 1994. 576 p. Available from: <https://market.android.com/details?id=book-lcLJevp2Kx0C>
143. Cuthill IC. Color perception. Bird coloration [Internet]. 2006;1:3–40. Available from: [https://books.google.com/books?hl=en&lr=&id=WnJ8cSoXRMIC&oi=fnd&pg=PA3&dq=Cuthill+Innes+C+%22Color+perception+%22+Bird+coloration+1+\(2006\)+3-40&ots=VVIG5Tvp8f&sig=B-UZpmngv0ZU-jvuqOID0DdTZH0](https://books.google.com/books?hl=en&lr=&id=WnJ8cSoXRMIC&oi=fnd&pg=PA3&dq=Cuthill+Innes+C+%22Color+perception+%22+Bird+coloration+1+(2006)+3-40&ots=VVIG5Tvp8f&sig=B-UZpmngv0ZU-jvuqOID0DdTZH0)

ANEXO

Procedimiento de calibración de instrumentos

1.1 Calibración de calentador Dako Pascal

En primera instancia se realizaron medidas de la temperatura interna del *coplin* con *RISH RETRIEVAL 1x* utilizando una sonda termocupla tipo T, previamente calibrada de fábrica con el objetivo de identificar settings apropiados para mantener una temperatura estable de 90°C durante 15 minutos. En segunda instancia se realizaron medidas de la temperatura interna del *coplin* con buffer de lavado TBS para mantener una temperatura estable de 55°C por 5 minutos. Se compararon las mediciones propias del instrumento minuto a minuto con las mediciones de la termocupla, como se muestra en la Figura A.1

1. Para el proceso de digestión se determinaron los siguientes ajustes del instrumento: Temperatura máxima: 96°C y 53 minutos.

Se deja el *coplin* con *buffer RISH RETRIEVAL 1x* desde el principio, cuando falten 15 minutos para terminar se saca el *coplin*, se insertan las muestras y se deja en el calentador Dako Pascal hasta que se terminen los 15 minutos.

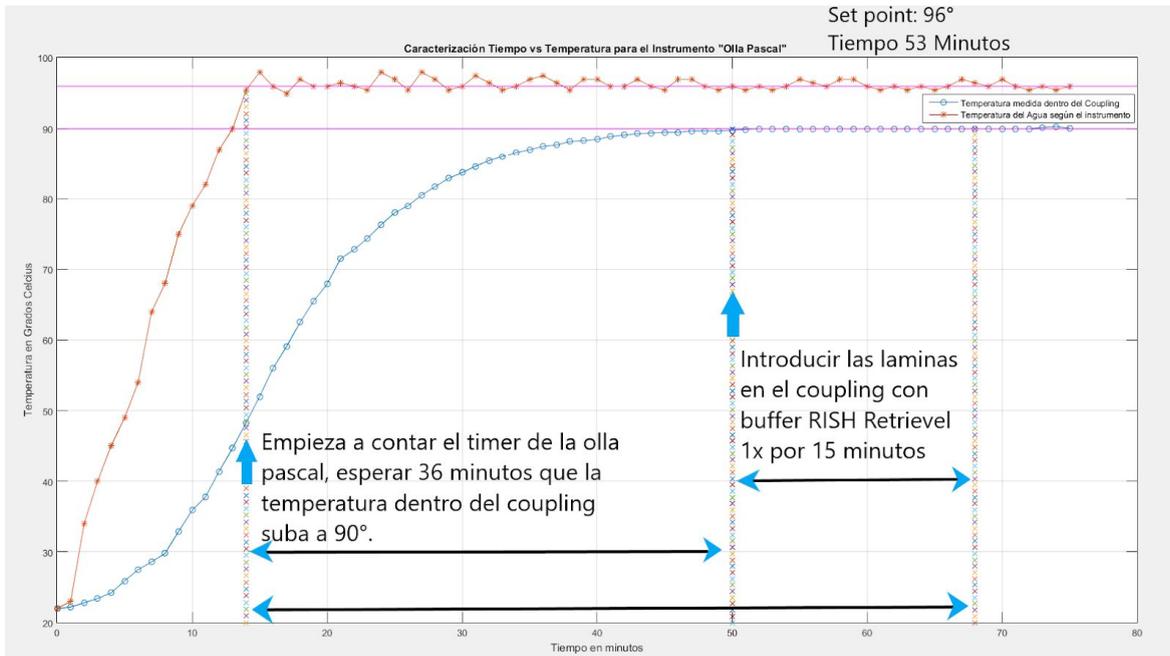
2. Para el proceso de lavado se determinaron los siguientes ajustes: Temperatura máxima 63°C y 25 minutos.

Se deja el *coplin* con *buffer de lavado TBS* desde el principio, cuando falten 5 minutos para terminar se saca el *coplin*, se insertan las muestras y se deja en el calentador Dako Pascal hasta que se terminen los 5 minutos.

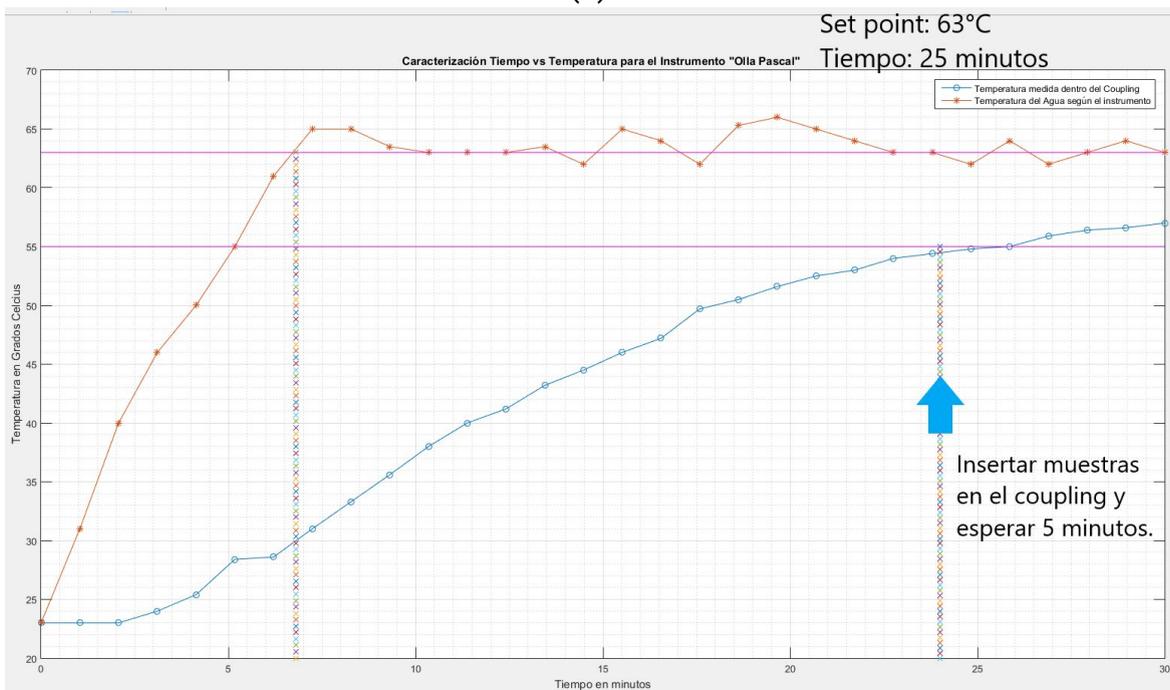
1.2 Calibración del hibridizador Dako

Se midió la temperatura de la plancha del hibridizador Dako utilizando una sonda termocupla, de forma similar al procedimiento anterior y no se encontraron diferencias significativas en las medidas de temperatura y tiempo registradas por el hibridizador y la termocupla.

1. Para el proceso de hibridación se configuró el hibridizador para hibridar las placas durante 60 minutos a 55°C



(a)



(b)

Figura A.1. Resultados de las mediciones de temperatura y tiempo para estandarizar procedimientos. La línea roja es la medición entregada por el instrumento, la línea azul es la medición hecha con la sonda termocupla. Eje Y : Temperatura en °C. Eje X: Tiempo en minutos. (a)Estandarización de recuperación de antigenicidad; (b) Estandarización de lavado después de hibridación.