

AUTOMATIZACIÓN DE LOS PROCESOS DE IDENTIFICACIÓN Y CONTEO
DE CÉLULAS EN FLUIDOS ANIMALES EN LA CLÍNICA VETERINARIA
CARLOS MARTÍNEZ HOYOS DE LA UNIVERSIDAD DE NARIÑO



JOHN JAIRO MUÑOZ ARÉVALO
JOHNNY MAURICIO PANTOJA FLOREZ

UNIVERSIDAD DE NARIÑO
FACULTAD DE INGENIERÍA
DEPARTAMENTO DE INGENIERÍA ELECTRÓNICA
SAN JUAN DE PASTO
2017

**AUTOMATIZACIÓN DE LOS PROCESOS DE IDENTIFICACIÓN Y CONTEO
DE CÉLULAS EN FLUIDOS ANIMALES EN LA CLÍNICA VETERINARIA
CARLOS MARTÍNEZ HOYOS DE LA UNIVERSIDAD DE NARIÑO**

**JOHN JAIRO MUÑOZ ARÉVALO
JOHNNY MAURICIO PANTOJA FLOREZ**

**TRABAJO DE GRADO PARA OPTAR POR EL TITULO DE INGENIERO
ELECTRÓNICO**

Director

PhD. Wilson O. Achicanoy M.

Ingeniero en Electrónica y Telecomunicaciones

**UNIVERSIDAD DE NARIÑO
FACULTAD DE INGENIERÍA
DEPARTAMENTO DE INGENIERÍA ELECTRÓNICA
SAN JUAN DE PASTO
2017**

NOTA DE RESPONSABILIDAD

“La Universidad de Nariño no se hace responsable por las opiniones o resultados obtenidos en el presente trabajo y para su publicación priman las normas sobre el derecho de autor.”

Acuerdo 1. Artículo 324. Octubre 11 de 1966, emanado del honorable Consejo Directivo de la Universidad de Nariño.

NOTA DE ACEPTACIÓN:

Firma del presidente del jurado

Firma del jurado

Firma del jurado

San Juan de Pasto, 15 de noviembre de 2017

DEDICATORIA

“A mis padres, a mis hermanos y amigos, por ser la inspiración y apoyo durante toda mi carrera y mi vida.”

JOHN JAIRO MUÑOZ ARÉVALO

“A toda mi familia, por ser la base fundamental de mis proyectos de vida, especialmente a mi padre Jesus Pantoja, a mi madre Angela Florez y a mi hermana Sonia Florez, por ser un apoyo incondicional en este proceso y por acompañarme especialmente en los momentos más difíciles. A mi pareja Carmen Pantoja, por cuidarme, ayudarme y apoyarme a crecer como persona.”

JOHNNY MAURICIO PANTOJA FLOREZ

AGRADECIMIENTO

“Agradecemos a nuestras familias por ayudarnos a cumplir nuestros objetivos y metas, a nuestros amigos por acompañarnos y ser un gran apoyo en todo este proceso académico, a nuestros profesores por compartir sus conocimientos y experiencias. Especialmente queremos agradecer a nuestro asesor, el profesor Wilson O. Achicanoy M., por brindarnos su confianza, amistad y acompañamiento para hacer este nuevo logro posible, además su paciencia y motivación han sido fundamentales para nuestra formación como investigadores, esto de verdad fue un aporte muy importante para el desarrollo de este trabajo”.

“Agradecemos también al Departamento de Electrónica, al grupo de investigación GIIEE, a la Clínica Veterinaria Carlos Martínez Hoyos de la Universidad de Nariño y al proyecto Desarrollo y promoción de competencias científicas y tecnológicas en Robótica e Informática en la Universidad de Nariño y establecimientos educativos del municipio de Pasto, Departamento de Nariño, por el apoyo académico y financiero para la realización exitosa de este trabajo de tesis.”

JOHN JAIRO MUÑOZ ARÉVALO

JOHNNY MAURICIO PANTOJA FLOREZ

RESUMEN

En esta investigación se propone un método de automatización para la identificación y conteo de células presentes en muestras de sangre de animales. Surge como respuesta a la necesidad del laboratorio de la Clínica Veterinaria Carlos Martínez Hoyos, de la Universidad de Nariño, de implementar un procedimiento confiable para el conteo automático de células en un hemograma. En este trabajo se ha propuesto un algoritmo de procesamiento y análisis de las imágenes digitales de la cámara de Neubauer y en tres pasos secuenciales: restauración, segmentación y conteo. En la restauración se propone filtros óptimos y locales para la remoción del ruido; en la segmentación se prueba el algoritmo de la transformación watershed, controlada por marcadores y mejorada a través de la proposición de un filtro óptimo local de acentuación de puntos oscuros, y finalmente, el conteo se realiza con el uso de una plantilla, obtenida también por medio del filtro local de acentuación de puntos oscuros, en conjunto con los resultados de los algoritmos de segmentación. Se obtuvieron resultados de conteo con porcentaje de error inferior al 11% comparado con el valor que obtiene el experto. También se propuso una interfaz de conteo que implementa el método propuesto y dos algoritmos más de conteo simplificado, a partir de un método de umbralización mejorado y de la técnica de agrupamiento *k-means clustering*.

ABSTRACT

We propose and implement an automatic method for identification and counting of blood cells in animal samples, which responds to a direct requirement of laboratories of Veterinarian Clinic Carlos Martínez Hoyos, from Universidad de Nariño. An algorithm of digital image processing over the Neubauer chamber is developed in three sequential steps: restoration, segmentation, and counting. In the restoration step, local and optimal filters are applied to remove the noise present in the image; in the segmentation stage, watershed transform is improved by a proposed local optimum filter of darkest point emphasis; finally, the count of blood cells is done by a matching template method, that is based on the proposed filter, and outcomes of the segmentation stage. Results with less than 11% of error are retrieved. In addition, it is developed a counting app that includes two simplified counting methods based on improved thresholding and k-means clustering.

CONTENIDO

| | |
|--|----|
| INTRODUCCIÓN | 15 |
| 1. DESCRIPCION DEL PROBLEMA..... | 17 |
| 1.1. PLANTEAMIENTO DEL PROBLEMA | 17 |
| 1.2. DEFINICION DEL PROBLEMA | 18 |
| 1.3. JUSTIFICACIÓN..... | 18 |
| 1.4. CONTRIBUCIONES DE ESTA TESIS..... | 19 |
| 1.5. ORGANIZACIÓN DEL DOCUMENTO..... | 20 |
| 2. OBJETIVOS | 21 |
| 2.1. OBJETIVO GENERAL..... | 21 |
| 2.2. OBJETIVOS ESPECÍFICOS..... | 21 |
| 3. MARCO TEÓRICO..... | 22 |
| 3.1. SANGRE Y HEMOGRAMA | 22 |
| 3.2. METODOS DE RECuento CELULAR | 26 |
| 3.2.1. Método automático..... | 26 |
| 3.2.2. Método manual: Caso laboratorio de la Clínica Veterinaria Carlos Martínez Hoyos, de la Universidad de Nariño | 27 |
| 3.3. TECNICAS DE PROCESAMIENTO Y ANALISIS DE IMÁGENES UTILIADAS EN ESTE TRABAJO | 31 |
| 3.3.1. Método de Otsu..... | 32 |
| 3.3.2. Detector de bordes de Marr-Hildreth..... | 33 |
| 3.3.3. Transformación watershed..... | 34 |
| 3.3.4. k-means clustering. | 34 |
| 4. METODOLOGÍA..... | 36 |
| 4.1. TRABAJO EN EL ÁREA | 36 |
| 4.2. AUTOMATIZACIÓN DEL PROCESO DE CONTEO DE CÉLULAS | 38 |
| 4.2.1. Sistema de adquisición de imágenes..... | 40 |
| 4.2.2. Restauración de las imágenes adquiridas..... | 43 |
| 4.2.3. Segmentación..... | 51 |
| 4.2.4. Conteo..... | 55 |
| 4.3. DISEÑO DE LA INTERFAZ DE AUTOMATIZACIÓN | 58 |

| | |
|--------------------------------|----|
| 5. RESULTADOS Y DISCUSIÓN..... | 63 |
| 5.1. MÉTODO 1 | 63 |
| 5.2. MÉTODO 2..... | 65 |
| 5.3. MÉTODO 3..... | 66 |
| 6. CONCLUSIONES..... | 70 |
| RECOMENDACIONES..... | 71 |
| BIBLIOGRAFIA..... | 72 |

LISTA DE FIGURAS

| | |
|--|----|
| FIGURA 1. COMPOSICIÓN BÁSICA DE LA SANGRE | 22 |
| FIGURA 2. ILUSTRACIÓN DE LA FORMACIÓN DE GLÓBULOS..... | 23 |
| FIGURA 3. ELEMENTOS NECESARIOS PARA REALIZAR CONTEO EN UN HEMOCITÓMETRO | 28 |
| FIGURA 4. DETALLE DE LA REJILLA DE CONTEO DE LA CÁMARA DE NEUBAUER | 28 |
| FIGURA 5. PIPETA DE DISOLUCIÓN | 29 |
| FIGURA 6. MICROSCOPIO OLYMPUS CH30RF100 Y CONTADOR MECÁNICO..... | 30 |
| FIGURA 7. PROCESO DE AUTOMATIZACIÓN DEL CONTEO DE CÉLULAS EN EL LABORATORIO DE LA CVCMH..... | 39 |
| FIGURA 8. SISTEMA DE ADQUISICIÓN DE IMÁGENES..... | 40 |
| FIGURA 9. IMAGEN DEL CUADRO ESQUINERO SUPERIOR IZQUIERDO DE LA CÁMARA DE NEUBAUER | 41 |
| FIGURA 10. MAGNITUD DE LA TRANSFORMACIÓN DE FOURIER Y PDF DE LA IMAGEN ADQUIRIDA | 42 |
| FIGURA 11. MAGNITUD DEL FILTRO IDEAL DE MUESCA..... | 45 |
| FIGURA 12. IMAGEN CON ÉNFASIS DE ALTA FRECUENCIA Y EL PATRÓN DE INTERFERENCIA CANALES RGB..... | 47 |
| FIGURA 13. REMOCIÓN DEL PATRÓN DE INTERFERENCIA, IMAGEN LON, IMAGEN LM Y PDF. CANALES RGB | 49 |
| FIGURA 14. REMOCIÓN DEL FONDO, IMAGEN DE RUIDO, IMAGEN LM E IMAGEN D. CANALES RGB | 51 |
| FIGURA 15. SEGMENTACIÓN, IMAGEN LDPE, IMAGEN MCW CON FILTRO LDPE, IMAGEN K (K- MEANS CLUSTERING) E IMAGEN MCW SIN FILTRO LDPE. CANALES RGB | 55 |
| FIGURA 16. CONTEO MÉTODOS A, B Y C. CANALES RGB..... | 58 |
| FIGURA 17. CAPTURA SIN ENFOQUE ADECUADO | 64 |
| FIGURA 18. CAPTURA DE BAJA RESOLUCIÓN Y GRAN CANTIDAD DE RUIDO | 65 |
| FIGURA 19. EJEMPLO DE CAPTURA OPTIMA | 66 |
| FIGURA 20. MUESTRA CON CÉLULAS QUE SE TRASLAPAN..... | 67 |
| FIGURA 21. MUESTRA NO ADECUADA Y RESULTADO DE LA RESTAURACIÓN..... | 68 |

LISTA DE TABLAS

| | |
|---|----|
| TABLA 1. TABLA DE RESULTADOS DE UN HEMOGRAMA | 24 |
| TABLA 2. TABLA DE ESPECIFICACIONES DEL OPERADOR DEL LABORATORIO. | 39 |
| TABLA 3. NOMENCLATURA PARA ALMACENAMIENTO DE IMÁGENES DE GB | 43 |
| TABLA 4. RESULTADOS MÉTODO 1 PARA EL CANAL ROJO. | 63 |
| TABLA 5. RESULTADOS MÉTODO 1 PARA EL CANAL VERDE. | 63 |
| TABLA 6. RESULTADOS MÉTODO 1 PARA EL CANAL AZUL..... | 64 |
| TABLA 7. RESULTADOS MÉTODO 1..... | 64 |
| TABLA 8. RESULTADOS MÉTODO 2 | 65 |
| TABLA 9. RESULTADOS MÉTODO 3..... | 66 |

LISTA DE ANEXOS

| | |
|---|-----------|
| ANEXO 1. LISTA DE ACRONIMOS | 74 |
| ANEXO 2. PONENCIA INTERNACIONAL ROPEC 2017 | 75 |
| ANEXO 3. MANUAL DE USUARIO DE LA APLICACIÓN..... | 76 |

GLOSARIO

Función de densidad de probabilidad (PDF, Probability Density Function): Histograma normalizado; es decir, dividido entre la resolución espacial de la imagen. Representa la probabilidad de ocurrencia de un nivel de gris.

Glóbulos blancos (GB): Tipo de célula sanguínea que se produce en la médula ósea y que se encuentra en la sangre y los tejidos linfáticos, son parte del sistema inmunitario del cuerpo.

Glóbulos rojos (GR): Tipo de célula sanguínea que se produce en la médula ósea y que se encuentra en la sangre. Contiene una proteína llamada hemoglobina, que transporta oxígeno desde los pulmones a todas las partes del cuerpo.

Hemograma: Examen que describe todos los elementos o componentes de la sangre; su número, proporción en el organismo y alteraciones.

Histograma: Función matemática que asigna a cada nivel de gris el número de píxeles presentes, con el mismo nivel, en la imagen.

Pixel: Unidad más pequeña de color (o nivel de gris) que compone una imagen de tipo digital.

Recuento celular: Serie de procedimientos que tienen por objeto determinar el número de cada uno de los tipos celulares por unidad de volumen de muestra.

Resolución espacial: Número de filas (el alto) por número de columnas (el ancho) de una imagen digital; es decir, el número total de píxeles que contiene.

Resolución de intensidad: Número de niveles de gris en los que se codifica una imagen digital.

RGB (red, Green, blue): Formato en el que se representa imágenes del espectro visible. Imita la forma en la que la visión humana percibe el color, a través de canales de color rojo, verde y azul.

INTRODUCCIÓN

Dada la importancia de las células sanguíneas para la mayoría de especies animales, es de gran importancia su estudio para describir y detallar sus características. El hemograma completo (HC) es un examen donde se hace el estudio de algunas características de la sangre como el número de glóbulos rojos, blancos y el estado de la hemoglobina [1]. Un recuento de glóbulos blancos o (GB) es un análisis que mide la cantidad de glóbulos blancos en el organismo y es una parte importante del (HC) [2] [3]. En la Clínica Veterinaria Carlos Martínez Hoyos de la Universidad de Nariño, para dar un diagnóstico a sus pacientes más frecuentes (caninos y felinos) se utiliza entre otras herramientas el (HC). En el presente, el personal de laboratorio hace el conteo de (GB) de forma manual, es decir, hace el conteo a través del microscopio con ayuda de objetos de conteo mecánicos [4]. Dado que este tipo de examen se debe realizar con mucha frecuencia, el personal de laboratorio está expuesto a un desgaste físico y visual lo cual puede conllevar a la existencia de un error en el conteo y a un mal diagnóstico. En la actualidad existen instrumentos electrónicos de conteo que pueden resolver el problema; lastimosamente estos dispositivos tienen un gran costo de adquisición. Esto ha motivado el desarrollo de una solución que se vea reflejada en un algoritmo, el cual puede realizar el conteo de (GB) reduciendo significativamente el esfuerzo visual y eliminando el contador mecánico del proceso, siendo más eficiente en tiempo comparado con el método actual.

Con respecto al estudio de células sanguíneas, el estudio microscópico es generalmente desarrollado a partir de imágenes en alta resolución y con el objetivo de extraer características propias de las células identificando regiones de interés (ROI) en las imágenes y a través de la combinación de técnicas básicas y avanzadas del procesamiento de imágenes. En [5], los autores proponen el uso de algoritmos de umbralización, como el método de Otsu y el algoritmo de *snake* o de contorno activo que usa el flujo del vector gradiente. Otra técnica utilizada en [6] y de uso común es la segmentación controlada *watershed* y en espacios de color, como el HSL. Esta misma técnica, junto con gradiente de textura y la transformación Wavelet, es utilizada en [7] para la extracción del núcleo de los leucocitos. Para realizar la agrupación de distintos tipos de células, los autores utilizan métodos de procesamiento y análisis de imágenes junto con técnicas de aprendizaje de máquina. En [8] por ejemplo, los autores proponen un método de agrupamiento de dos tipos de células blancas a partir de los momentos pseudoZernike junto con una máquina de soporte vectorial (Support Vector Machine - SVM).

Otra de las técnicas que se utilizan para la clasificación de células es *k-means clustering*. Por ejemplo, en [9], los autores lo utilizan para clasificar cuatro tipos de leucocitos. En este estudio se concluye que la circularidad es la característica geométrica que mayor exactitud da en la clasificación. En [8], para realizar el conteo

de las células, también se utiliza *k-means clustering* junto con un modelo *ensemble-multiclass SVM*.

Para el desarrollo de esta investigación se ha dispuesto una serie de etapas ordenadas de procesamiento digital de imágenes, la primera etapa es la restauración y mejoramiento de imágenes, en donde se dispuso de transformaciones espaciales como *local matching* y filtros frecuenciales como filtros *Notch*. Para la etapa de segmentación y extracción de las regiones de interés (ROI), se plantearon diversos métodos, entre los cuales se destacan, agrupamiento *k-means*, transformada *watershed*, método de *Otsu* y detectores de bordes como *marr-hildreth* [10]. Por último, la discriminación y el conteo de dichas regiones se hace mediante la creación y comparación de una plantilla con cada método, además en algunos casos el conteo se hace directamente de la extracción de propiedades de la imagen binaria que produce la umbralización directa del histograma de la imagen. La solución y avance a lo largo de estas etapas depende de la adecuación de las herramientas, tecnologías y técnicas ya desarrolladas, además de la creación de métodos que se ajustan y nacen de los requerimientos propios del problema que se intenta resolver. Posterior a la implementación y prueba de los métodos planteados, se desarrolló una interfaz gráfica para facilitar al operador el uso de estos y otros con menor carga computacional. La idea de la interfaz es garantizar la correcta ejecución de los algoritmos propuestos, automatizando los procesos de conteo priorizando a los GB

Para la prueba y evaluación de resultados, se crea una base de datos que recopila los datos básicos del paciente, la correcta nomenclatura de cada captura y los resultados de conteo obtenidos. Finalmente, se presenta la discusión y comparación de métodos, técnicas y resultados obtenidos en esta investigación.

1. DESCRIPCION DEL PROBLEMA

1.1. PLANTEAMIENTO DEL PROBLEMA

Un servicio de gran importancia para el diagnóstico médico veterinario prestado por la Clínica Veterinaria Carlos Martínez Hoyos (CVCMH), de la Universidad de Nariño para pacientes caninos y felinos, es la realización de exámenes de laboratorio [11]. Dentro de estos exámenes, objeto de estudio en este trabajo de investigación, se encuentra el hemograma. El hemograma es una prueba de laboratorio en donde se evalúan las características de diferentes grupos de células, como glóbulos blancos (leucocitos), glóbulos rojos (eritrocitos) y plaquetas. Brinda información de parámetros relacionados con su cantidad, forma y composición [12] [13].

Uno de los resultados obtenidos por el hemograma es el recuento de glóbulos rojos, blancos y plaquetas (por unidad de volumen) [1]. Para estimar la concentración celular en una muestra de fluido, se realiza el recuento celular para cada tipo de célula. Este recuento se realiza de manera independiente y manual para glóbulos rojos y glóbulos blancos. En este proceso, el personal del laboratorio hace uso de los respectivos manuales de procedimientos hematológicos de la clínica y se ayuda de elementos mecánicos de conteo [4]. Este método implica un esfuerzo alto, tanto físico como mental por parte del experto, ya que el conteo se realiza a través de un microscopio y conlleva a que la persona pase una cantidad considerable de tiempo en esta labor.

El manual elaborado por la Universidad de Nariño recomienda al experto no tomar descanso mientras se realiza el conteo de las células a través del microscopio, para evitar equivocaciones y no conducir a posibles errores en el diagnóstico. Además, el esfuerzo visual que hace el experto implica un desgaste progresivo en su visión [4].

Por otro lado, en la actualidad existen contadores electrónicos capaces de suplir las demandas de tiempo y exactitud en la realización de este tipo de exámenes. El margen de error que estos equipos manejan es inferior al 5%, lo que reduce considerablemente el error del procedimiento manual, cuyo margen de error puede incluso llegar al 20% [14] (este error se debe al método en sí y no al cometido por el experto). Sin embargo, el limitante en el uso de estos aparatos electrónicos sigue siendo el costo elevado para su adquisición, ya que el precio en promedio oscila alrededor de los 7.000 dólares.

De acuerdo a lo anterior, teniendo en cuenta que el recuento es una de las actividades que se realiza a diario en la clínica, y que el margen de error depende de muchas variables, entre ellas la condición del experto que hace el examen, y que además la CVCMH cuenta con una base de datos física proveniente de resultados de exámenes hematológicos de pacientes que han sido tratados en esta, se propone la automatización de los procesos de identificación y recuento de células en un hemograma por medio de un algoritmo de cómputo que utilizará una base digitalizada de las muestras de la clínica [2].

1.2. DEFINICION DEL PROBLEMA

¿Es posible realizar la identificación y recuento celular de forma automática, reduciendo el tiempo de obtención de resultados, bajo ciertos márgenes de exactitud y precisión aceptables, teniendo en cuenta la relación beneficio-costos?

1.3. JUSTIFICACIÓN

En la CVCMH existen diversos procedimientos de laboratorio para examinar a un paciente, enfocados a determinar la condición física del mismo [15]. El estudio de la concentración de las células sanguíneas es utilizado para la identificación y seguimiento de algunas enfermedades en el área de hematología e involucra dos procesos: el conteo y la inspección visual de las células [16]. La exactitud y precisión en la determinación del número de células mediante procedimientos electrónicos y computacionales cuestionan la eficacia de los métodos mecánicos (conteo en cámara); sin embargo, dado que los procedimientos mecánicos son de bajo costo, fácil utilización y uso frecuente en Colombia, se justifica conocer la eficacia relativa de ambos procedimientos a la hora de determinar el número de células sanguíneas y así conocer la concordancia que estas pruebas tienen entre sí [17].

Dentro de las dependencias de la CVCMH, especialmente en sus laboratorios, se necesitan métodos automáticos y confiables para el diagnóstico y evaluación de pacientes. Para realizar esto se hace preciso el desarrollo de una aplicación capaz de reducir el tiempo empleado por el personal de la clínica veterinaria al momento de realizar los exámenes y para garantizar un margen de error confiable en la obtención de los resultados, sobre todo a los correspondientes a la identificación y conteo celular sobre muestras de fluidos animales.

El desarrollo de esta investigación facilitará el trabajo de los expertos del laboratorio de la clínica y podría redundar en beneficios de salud para los pacientes y el

personal; además de posibles recursos económicos para la institución. Por ejemplo, se podría aumentar el número de exámenes por día (mayor cantidad de servicios) o invertir el tiempo de los funcionarios de laboratorio en otras actividades. La solución que ofrece esta propuesta permitirá realizar una cantidad mayor de análisis sobre los datos y en un menor tiempo. Esta investigación también se constituirá en un aporte tecnológico significativo que apoyará las actividades de docencia e investigación de los expertos en medicina animal de la CVCMH.

Para el problema en particular que se quiere solucionar en esta investigación, la Ingeniería Electrónica brinda los métodos teóricos y experimentales suficientes para el logro exitoso del objetivo de automatización [18]. El problema de automatización se resolverá empleando específicamente técnicas de procesamiento y análisis de imágenes digitales, que se tomarán de una base digital que se construirá también como parte de este proyecto a partir de las muestras que se encuentran almacenadas y que diariamente son tratadas en la clínica.

1.4. CONTRIBUCIONES DE ESTA TESIS

En las instalaciones de los laboratorios de la CVCMH se realizan exámenes médicos para el diagnóstico de pacientes. Para la obtención de los resultados de estos exámenes existen métodos ya implementados que hacen uso del tiempo y la pericia de los expertos. La búsqueda de nuevos métodos motiva la creación de una solución asequible para la clínica y sus funcionarios. Entre las principales contribuciones de la tesis se destaca:

- En este trabajo se propuso una técnica de conteo capaz de afrontar y mejorar criterios como margen de error, tiempo de ejecución y confiabilidad en la entrega de resultados; además, no es costosa en comparación con los métodos automáticos más básicos consultados.
- La interfaz gráfica de usuario (GUI, *Graphics User Interface*) que se logró como resultado de este trabajo de grado, está adaptada a las necesidades del laboratorio de la CVCMH y de los expertos encargados de la obtención de resultados; contribuyendo también a la ampliación y manipulación de la base de datos de los pacientes examinados.
- El desarrollo de una propuesta de automatización de los procesos de recuento celular, y el proceso de diseño e implementación de la GUI, permiten evidenciar la aplicación de investigación básica a problemas reales del entorno y relacionando distintos campos de la ingeniería y la medicina veterinaria.

- Otra contribución importante es el aporte científico y su divulgación a través de la publicación de un artículo de conferencia

1.5. ORGANIZACIÓN DEL DOCUMENTO

Este trabajo está dividido en 7 secciones principales nombradas de la siguiente manera: Introducción, descripción del problema, objetivos, marco teórico, metodología, resultados y conclusiones.

En la sección 1 se presenta el planteamiento del problema, la justificación de este trabajo y las contribuciones científicas y prácticas de esta investigación.

En la sección 2 se presenta el objetivo general y los objetivos específicos, que fueron planteados al inicio de este trabajo de grado.

En la sección 3 se presenta una revisión bibliográfica que incluyen conceptos básicos sobre procesamiento digital de imágenes, hematología, física sanguínea animal, procedimientos y manuales de laboratorio usados en la clínica veterinaria Carlos Martínez Hoyos de la Universidad de Nariño.

En la sección 4 se describe algunos antecedentes en el área de este trabajo la metodología de identificación y conteo celular propuesta, así como su implementación en los procesos de laboratorio de la clínica Carlos Martínez Hoyos a través de una interfaz gráfica también propuesta en este trabajo.

Los resultados de los experimentos, pruebas y su respectiva evaluación se discuten en la sección 5.

Finalmente, en la sección 6 y 7 se presentan las conclusiones que se obtuvieron a partir de este trabajo, así como los trabajos futuros que pueden mejorar la metodología propuesta para la identificación y conteo celular y definir nuevos trabajos de investigación.

2. OBJETIVOS

En esta sección se dan a conocer los objetivos planteados en el presente trabajo de grado.

2.1. OBJETIVO GENERAL

Desarrollar un algoritmo que sea capaz de automatizar los procesos de detección y conteo de las células presentes en muestras de fluidos animales, a partir de imágenes microscópicas tomadas a los fluidos, y que sea confiable bajo un margen de error menor al 20%.

2.2. OBJETIVOS ESPECÍFICOS

- Recopilar y estudiar métodos teóricos y experimentales sobre el procesamiento y análisis de imágenes de microscopía, especialmente aquellos relacionados con los fluidos de distintos tipos de células animales.
- Crear la base de datos de imágenes digitales de la Clínica Veterinaria Carlos Martínez Hoyos, a partir de las muestras de fluido de algunas especies de animales que la clínica recogerá en el tiempo de duración del proyecto.
- Implementar el método de identificación de células en los fluidos sanguíneos, a partir de la base de datos de imágenes digitales de la clínica.
- Desarrollar un método de conteo de células de los fluidos sanguíneos, a partir de la base de datos de imágenes digitales de la clínica.

3. MARCO TEÓRICO

3.1. SANGRE Y HEMOGRAMA

La sangre es el medio líquido donde se encuentran dispersas las células sanguíneas [17] (**Ver Figura 1**). Gracias a la sangre, todas las células del organismo reciben un aporte continuo de nutrientes y hacen la eliminación de sustancias de desecho originadas en su metabolismo [3].

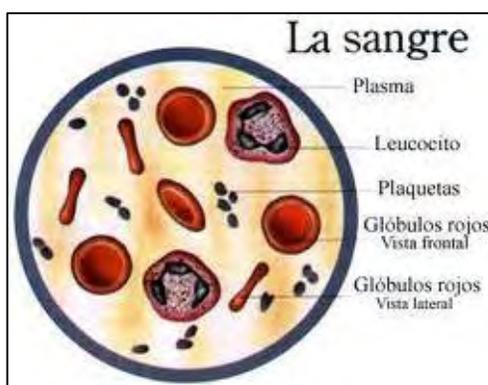


Figura 1. Composición básica de la sangre. Fuente: [2]

Entre las células que se encuentran en la sangre, se identifica principalmente los **glóbulos blancos** (GB), o leucocitos, descubiertos por Hewson en 1770 y que aparecen en los frotis sanguíneos como células de morfología un poco irregular, y los **glóbulos rojos** (GR), con forma más regular. Los GB están presentes en menor cantidad, comparada con la población de GR. En mamíferos, por ejemplo, el número de GB oscila de 5.000 a 12.000 por mm^3 , y en aves este número se concentra alrededor de los 30.000 por mm^3 [1] [2].

Los GB no se pueden considerar como elementos no exclusivos de la sangre, ya que también se pueden encontrar por ejemplo en la linfa, LCR y tejidos [8], y se clasifican en: **granulocitos**, que se caracterizan por la presencia de gránulos específicos en el citoplasma, y **agranulocitos** (lo contrario). A su vez, los granulocitos, por la afinidad tintorial de sus gránulos por colorantes neutra, ácida o básica, se clasifican respectivamente en **neutrófilos**, **eosinófilos** y **basófilos**. Los agranulocitos, carentes de gránulos, se agrupan en **linfocitos** y **monocitos** [1] [17].

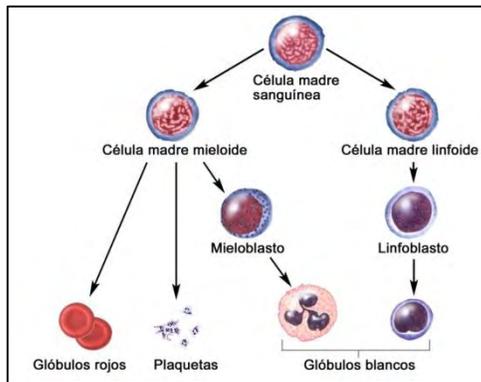


Figura 2. Ilustración de la formación de glóbulos blancos, rojos y plaquetas. **Fuente:** <https://www.cancer.gov/espanol/tipos/mieloproliferativas/paciente/tratamiento-cronicas-pdq>

Por su parte, los GR, descubiertos por Leenwehoek en 1637, se caracterizan por ser células muy especializadas y que en los mamíferos carecen de orgánulos citoplasmáticos y núcleo, y son de forma bicóncava con una depresión central (Ver Figura 2) [17] [14]. Los GR se diferencian en forma, tamaño y grosor entre especies animales y también según el estado nutricional. Por ejemplo, el diámetro del GR varía de 4 μm en la cabra a 7 μm en el perro; o el grosor es de 1,5 μm en la cabra y de 2,2 μm en la vaca [15].

La forma discoide del GR, determinada por su membrana, proporciona un aumento del área superficial, lo que facilita la plasticidad celular, la resistencia a la lisis osmótica hipotónica y el intercambio de gases [15]. Pueden darse variaciones distintas en la concentración de GR dentro del mismo animal, encontrándose diferencias entre muestras de sangre venosa y arterial. Existen además otros factores de variación entre los que se destaca: edad, sexo, raza, actividad física, altitud, alimentación, estado fisiológico y situaciones de estrés [15].

Para analizar el contenido de la sangre se utiliza el hemograma, que es uno de los exámenes de laboratorio solicitado con mayor frecuencia, y forma parte del estudio básico requerido para orientación diagnóstica y evaluación de los pacientes. El hemograma es un análisis de la sangre y conteo sanguíneo completo (CSC), donde se reflejan todos los elementos o componentes de la sangre: su número, proporción en el organismo y si sufren alteraciones (**ver Figura 3**) [12]; resultando ser muy útil a la hora de diagnosticar enfermedades o evaluar tratamientos en las especies de animales domésticos [13].

| Análisis | Valor | Referencia |
|--------------|------------|-------------------|
| Hematocrito | 40,3 % | 37 – 55 |
| Hemoglobina | 11,6 g/dL | 11,5 – 18 |
| VCM | 74,0 fL | 60 – 77 |
| CHCM | 28,9 g/dL | 28,5 – 36,5 |
| Leucocitos | 3.900/uL | 5.500 – 13.500 |
| Baciliformes | 117/uL | 0 – 500 |
| Segmentados | 3.237/uL | 4.100 – 9.300 |
| Linfocitos | 390/uL | 1.300 – 3.900 |
| Monocitos | 156/uL | 80 – 850 |
| Plaquetas | 162.000/uL | 140.000 – 550.000 |

g/dL: Gramos/decilitro; fL: Femtolitro (10-15 litro); uL: Microlitro (10-6 L)

Tabla 1. Tabla de resultados de un Hemograma realizado a una paciente canina, hembra esterilizada, mestiza y de ocho años. Se destaca en rojo los valores fuera de los rangos de referencia. **Fuente:** <http://allextruded.com/poliartritis-inmunomediada-idiopatica-tipo-i-en-un-paciente-canino/>.

Los resultados que muestra un hemograma son:

- **Hemoglobina (Hb):** es una proteína formada de hierro que se encuentra en el interior del hematíe o GR, y que es la causante del color rojo de la sangre. Cada hematíe suele contener entre 200 y 300 moléculas de hemoglobina. Como la cantidad de hemoglobina es proporcional al número de glóbulos rojos (hematíes), un descenso de esta proteína se refleja en una ineficacia de la función de los hematíes, y que se conoce como anemia. Al contrario, la elevación de esta proteína puede contribuir a la aparición de poliglobulia; es decir, un aumento de la cantidad de hematíes que puede provocar coágulos en venas o arterias. También pueden aparecer niveles elevados en pacientes con cardiopatías y problemas pulmonares crónicos [11] [3].
- **Hematocrito (Hto):** es el volumen de hematíes en la sangre expresado como un porcentaje sobre el volumen sanguíneo total. Debido a que en realidad este parámetro indica el número de hematíes, la causa principal de un descenso del hematocrito es la anemia. Otros motivos del descenso pueden ser: hemorragias, embarazo, problemas en la médula ósea, leucemia e hipertiroidismo. Un incremento del nivel de hematocrito puede producirse por problemas cardiacos, falta de hidratación y enfermedades pulmonares crónicas [2] [3].
- **Volumen corpuscular medio (VCM):** es el índice que determina el tamaño medio de los hematíes. De este modo, se pueden clasificar las anemias en: macrocíticas o microcíticas, dependiendo de si el tamaño del hematíe es mayor o menor de lo habitual. Un VCM alto (glóbulos rojos grandes) puede tener su origen en un déficit de vitamina B12 o de ácido fólico, trastornos del hígado o consumo de alcohol. Este índice no permanece constante a lo largo de toda la vida y en los recién nacidos es más elevado. Los niveles bajos

pueden estar originados por anemias o incluso talasemias (alteración de la hemoglobina) [12].

- **Hemoglobina corpuscular media (HCM):** indica la cantidad media de hemoglobina que contiene cada hematíe o glóbulo rojo. Gracias a este parámetro se pueden clasificar las anemias de otra forma: las hipocrómicas (por déficit de hierro), con un bajo nivel de HCM, y las hiperocrómicas, que son un caso raro de anemia con un alto nivel de HCM y que alertan de un déficit de vitamina B12 o ácido fólico [12].
- **Plaquetas:** son los elementos más pequeños de la sangre y resultan esenciales para la coagulación sanguínea, ya que se encargan de cerrar los vasos sanguíneos, haciendo que la sangre se coagule para taponar las lesiones cuando se produce una herida. El recuento plaquetario inferior a lo normal es lo que se conoce como trombocitopenia. Tiene como consecuencia la mala coagulación sanguínea, lo que dará lugar a hemorragias (por ejemplo: nasales, de encías, hematomas en la piel y sangre en orina). El número elevado de plaquetas en sangre, se denomina trombocitosis, y puede originar la formación de trombos o coágulos dentro de las arterias [15].
- **Velocidad de segmentación globular (VSG):** es el parámetro que mide la velocidad a la que sedimentan los glóbulos rojos de la sangre en un tiempo determinado (1-2 horas). Son muchos los procesos que se originan con un aumento del VSG, como el mieloma, los linfomas, las leucemias y los procesos inflamatorios crónicos, como la artritis reumatoide o el lupus. Pero su elevación no indica necesariamente una patología, ya que también se eleva en algunos procesos fisiológicos como el embarazo, la menstruación o en las personas de la tercera edad. La VSG raramente suele estar disminuida. Sin embargo, es un dato útil en el seguimiento de pacientes, ya que si la VSG se reduce sugiere que el tratamiento impuesto está siendo eficaz [12].
- Los **leucocitos** desempeñan una función de defensa del organismo frente a las agresiones externas. Debido a esta función, un nivel bajo de leucocitos predispone al paciente a padecer infecciones. Los linfocitos son un tipo de leucocito que carece de gránulos. Son células muy importantes en el sistema inmune, ya que son capaces de responder ante agentes desconocidos para el organismo. Los niveles normales rondan entre los 1.300 y 4.000 /ml, por otra parte, el aumento del número de linfocitos se denomina linfocitosis. Aparece en procesos infecciosos agudos, crónicos, alergias farmacológicas y procesos linfoproliferativos como la leucemia. La disminución del número de linfocitos se llama linfopenia, y es habitual encontrarla en personas cuyo sistema inmune es defectuoso o están siguiendo un tratamiento inmunosupresor (quimioterapia) [17].

3.2. METODOS DE RECUESTO CELULAR

Los recuentos celulares son una serie de procedimientos que tienen por objeto determinar el número de cada uno de los tipos celulares que están comprendidos en una unidad de volumen de sangre. Los recuentos celulares son parte fundamental del hemograma y constan de 3 fases consecutivas:

- Dilución de la sangre.
- Conteo del número de células.
- Cálculo matemático del número de células presentes en volumen dado (generalmente, en 1 mm^3).

Los recuentos celulares pueden realizarse mediante métodos automáticos, que utilizan contadores electrónicos, o métodos manuales, conocidos como recuentos en cámara [12].

3.2.1. Método automático.

En los laboratorios modernos de clínicas veterinarias es común encontrar dispositivos tecnológicos avanzados, como los auto-analizadores hematológicos o contadores electrónicos para realizar el conteo automático de células presentes en las muestras sanguíneas de pacientes. Un ejemplo son los micro-contadores desarrollados por la empresa Celeromics [19], con rango de detección de células entre 1 y 100 micrones y en soluciones de densidad de 10^3 hasta 10^8 células por unidad de volumen. La precisión registrada por estos métodos automáticos tiene un error no superior al 1% y el conteo lo realizan sobre una gama de clases de células posiblemente presentes en las muestras y sobre todos los parámetros requeridos por el hemograma. Sin embargo, el costo continúa siendo uno de los principales obstáculos a la hora ser utilizados en clínicas de bajo presupuesto; la versión más económica está alrededor de 2.000 €.

Estos sistemas automáticos se fundamentan en el siguiente esquema de medición: (i) Preparación del **solvente** para disminuir la concentración de la sangre hasta el nivel adecuado requerido por el contador; (ii) uso de un **compresor – aspirador** que aporta la presión y el vacío necesarios para transportar la sangre, convenientemente diluida, al dispositivo de medida, y (iii) uso del **dispositivo de medida**, que es la cámara donde realmente se cuentan las células sanguíneas (se denomina también cámara de medida o de lectura). El diseño de esta cámara depende del método utilizado por cada modelo de contador automático; por ejemplo, este puede basarse en un sistema de detección celular por medida de la impedancia, o en sistemas ópticos de detección celular [12] [17].

En los contadores electrónicos más avanzados, una válvula de cerámica reparte la muestra hacia distintas cámaras de medida. Cada cámara de medida está especialmente diseñada para el recuento de un tipo celular (glóbulos rojos y leucocitos o plaquetas) o para la determinación de hemoglobina. El proceso de recuento se efectúa de la siguiente forma: Un **transductor** transforma las señales procedentes del dispositivo de medida en impulsos eléctricos (se conoce como fotomultiplicador); luego el **discriminador** diferencia los impulsos eléctricos generados por cada uno de los tipos de células sanguíneas; que se pasan al **procesador**, que recoge los datos obtenidos, los procesa y los proyecta en una pantalla. Finalmente, el **registrador** imprime en papel los resultados calculados.

3.2.2. Método manual: Caso laboratorio de la Clínica Veterinaria Carlos Martínez Hoyos, de la Universidad de Nariño

El método manual sigue un procedimiento similar al automático, pero las herramientas para registrar el número de células son visuales y utilizan elementos mecánicos de ayuda de cuenta. Un ejemplo de método manual es el que se lleva a cabo en la CVCMH de la Universidad de Nariño, que a continuación se explica en mayor detalle.

El laboratorio de la CVCMH ofrece actualmente los estudios de hematología, análisis coprológico, uroanálisis, citología, manejo de autovacunas, pruebas de química sanguínea y diagnóstico de hemoparásitos, entre otros. Para los exámenes del área de hematología, específicamente el hemograma, el laboratorio no cuenta en el momento con contadores de células electrónicos automáticos. En su lugar utiliza el hemocitómetro o cámara de Neubauer [13], que un método visual y manual para la estimación de la población de células en las muestras de los pacientes (continúa siendo el método manual más extendido desde el siglo XIX) [15] [19].

Para el área de hematología y otros procedimientos, el encargado del laboratorio realiza las funciones a partir de la guía de procedimientos de la clínica [4]. Específicamente, para realizar el conteo de células en la cámara de Neubauer se requiere el siguiente conjunto de elementos: (1) Dilución celular a medir, (2) hemocitómetro o cámara de Neubauer, (3) cubierta de vidrio, (4) microscopio óptico, (5) pipeta o micropipeta y (6) PBS [19] [1] (**Ver Figura 4**).



Figura 3. Elementos necesarios para realizar conteo en un hemocitómetro. **Fuente:** [19].

La cámara de Neubauer es un cristal delgado de dimensiones 30 x 70 x 4 mm (alto, ancho y profundidad) y en la que distingue una parte central donde se realiza el conteo de las células (otras tienen dos cámaras de conteo). La Figura 5 muestra la rejilla de conteo que viene en el interior de la cámara. Esta rejilla tiene un área de 3 x 3 mm y se divide en 9 rectángulos. En muestras de sangre, los cuadrados de las esquinas se utilizan para realizar el conteo de células blancas (baja concentración), mientras que el cuadrado del centro es usado para el conteo de plaquetas y células rojas. Este cuadro central está dividido en 25 subcuadros, de tamaño 0,2 x 0,2 mm; a su vez, cada uno de estos subcuadros está dividido en 16 más pequeños. Luego, el cuadro central está conformado en total por 400 cuadrados pequeños.

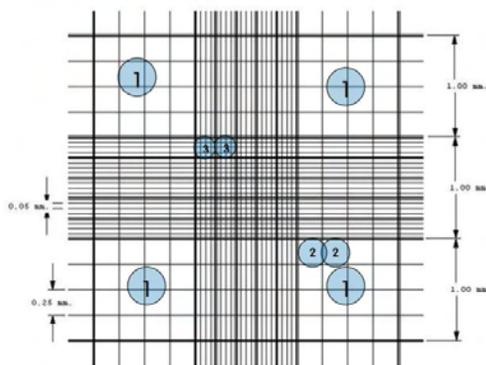


Figura 4. Detalle de la rejilla de conteo de la cámara de Neubauer. **Fuente:** [19].

El proceso completo de conteo se puede consultar en [4] y [12], que se resume en los siguientes pasos:

- **Preparación de la muestra.** Se recomienda una concentración de 10^6 células/ml, de otro modo el error puede aumentar. Para concentraciones por

debajo de $2,5 \times 10^5$ células/ml la estimación no es buena; y para concentraciones por encima de $2,5 \times 10^6$ se debe diluir hasta alcanzar un aproximado al recomendado. El reactivo utilizado para la disolución también depende de las células que se van a contar. Los líquidos diluyentes más utilizados son el de Hayem, para el recuento de hematíes, y el de Turck, para el recuento de leucocitos [12] [4].

- **Introducción de la muestra en la cámara.** Colocar la cubierta de vidrio sobre la cámara y tomar una cantidad determinada de dilución (aproximadamente $10 \mu\text{l}$) con la micro-pipeta y colocarla en la cámara (en el centro) de forma uniforme (absorción capilar). Repetir el proceso si hay burbujas o el vidrio se ha movido. En la Figura 6 se muestra un ejemplo de pipeta especial de cristal constituida por un tubo capilar largo y graduado, que termina en punta en uno de sus extremos y en un bulbo en el otro. El bulbo contiene una perla de vidrio para facilitar la mezcla de la sangre con el líquido de dilución. La perla es roja en las pipetas empleadas para el recuento de hematíes y blanca en las usadas para el recuento de leucocitos [4] [12].

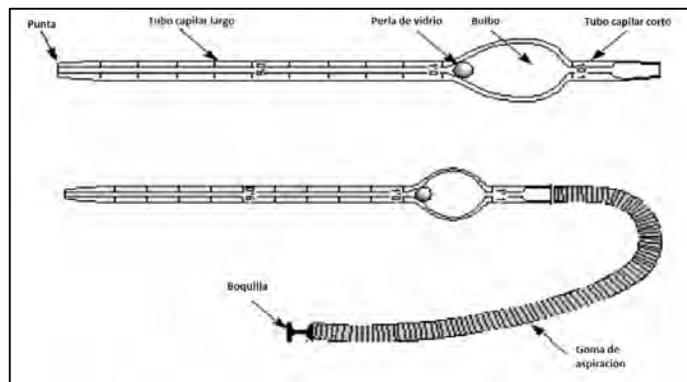


Figura 5. Pipeta de disolución. Fuente: [4].

- **Configuración del microscopio.** Colocar la cámara de Neubauer bajo el microscopio (sujetarla si es el caso) y encender la luz. Luego enfocar la imagen a una región de conteo, por ejemplo, uno de los cuadros grandes para el conteo de los glóbulos blancos, o uno de los cuadros pequeños para el conteo de los glóbulos rojos. En la CVCMH se utiliza el microscopio óptico Olympus CH30RF100 [4] (Ver Figura 7).
- **Conteo.** Realizar el conteo de las células utilizando como referencia la rejilla, por ejemplo, de izquierda a derecha (este procedimiento puede ser distinto y depende del protocolo del laboratorio). Se registra el número de células contadas y se procede a realizar el mismo proceso con otro cuadro hasta completar la información de la fórmula de estimación. En el caso de la CVCMH, el conteo se hace con la ayuda de un contador mecánico como el que se muestra en la Figura 6.



Figura 6. Microscopio Olympus CH30RF100 y contador mecánico utilizados en el laboratorio de la CVCMH de la Universidad de Nariño.

- Con la información del paso anterior se aplica la fórmula de estimación de la concentración:

$$C = \frac{N}{V} \quad (1)$$

Donde C es la concentración [células/ml], N es el número de células contadas y V es el volumen de los cuadros [ml]. Para glóbulos blancos se reduce a:

$$C = 10^4 \frac{N}{NS} \quad (2)$$

Donde NS es el número de cuadros utilizados en el conteo. Y si se ha utilizado dilución, será:

$$C = 10^4 \frac{N}{NS * D} \quad (3)$$

Donde D es la dilución [19].

El método de conteo manual tiene un error máximo alto, del 20% al 30%. A pesar de esto, sigue siendo utilizado por su bajo costo. Cabe aclarar que el porcentaje de error al que se introduce es debido al método en sí (error en el volumen real de la cámara, error en la medida de la pipeta y/o error estadístico). Entonces, si además de este error introducimos el error debido al operario y al proceso de medida como tal, su incremento podría ser alto y podría alterar significativamente el resultado.

3.3. TÉCNICAS DE PROCESAMIENTO Y ANÁLISIS DE IMÁGENES UTILIZADAS EN ESTE TRABAJO

El procesamiento y análisis de imágenes (PAI) se refiere al conjunto de teorías, procedimientos y técnicas que se han desarrollado, desde inicios del siglo pasado hasta nuestros días, para la manipulación e interpretación de la información gráfica proveniente de distintas fuentes o procesos. Su evolución ha ido desde lo analógico hasta lo digital, y ha logrado subir a un máximo en las últimas décadas debido al avance de la ingeniería, en especial la electrónica y la computación.

Este avance está íntimamente ligado con el desarrollo de la computación digital; puesto que el hecho de trabajar con imágenes digitales requiere de capacidad computacional y almacenamiento altos. Por lo tanto, los avances van a la par con el desarrollo de nuevos dispositivos electrónicos de procesamiento paralelo, como las unidades de procesamiento gráfico (GPU, *Graphics Processor Unit*) [10], nuevas arquitecturas de computadores digitales y tecnologías de soporte para el almacenamiento de datos, visualización y transmisión [10]. Ahora es posible manejar grandes volúmenes de información de entrada (imágenes en alta resolución) y ejecutar algoritmos del PAI en tiempos más cortos y que permiten pensar en el diseño de sistemas que responden en tiempo real a las exigencias de problemáticas actuales.

Por ejemplo, en automóviles los sistemas de navegación cuentan con opciones de navegación asistida y automática, parqueo automático, detección de obstáculos (peatones, por ejemplo), estimación de riesgo en vía y ya se piensa en un futuro contar con autos no tripulados. Otro ejemplo lo constituye el avance del aprendizaje de máquina, que a partir del año 2012 ha logrado ser funcionalmente eficaz con el desarrollo del procesamiento en paralelo y con el desarrollo de nuevos algoritmos conocidos como de aprendizaje profundo. Las redes neuronales convolucionales son un claro ejemplo de estos avances [20].

En esta investigación, y debido a los algoritmos del PAI que se proponen y utilizan, el tiempo de procesamiento no es un factor crítico a ser tenido en consideración como restricción, así como tampoco la capacidad de almacenamiento de las imágenes que se procesan, puesto que el tiempo empleado por el experto es siempre superior al tiempo empleado por los algoritmos propuestos y la capacidad de almacenamiento depende de la capacidad del ordenador. Los conceptos y métodos del PAI que se han propuesto y utilizado están clasificados dentro de niveles de complejidad medio y avanzado; y para un mejor entendimiento del trabajo

realizado en esta tesis, a continuación, se presenta un resumen corto con las características principales de cada uno de ellos.

3.3.1. Método de Otsu.

Es un método que se utiliza para la segmentación y extracción de regiones de interés (ROI, *Region Of Interest*) en una imagen. La importancia de este método radica en que hace una operación de umbralización o binarización automática y adaptativa; es decir, no necesita supervisión humana ni información previa de la imagen antes de su procesamiento, y se adapta a las propiedades de cada imagen.

Este método se emplea cuando hay una clara diferencia de los objetos a extraer respecto al fondo de la escena, y selecciona un umbral óptimo, maximizando la varianza entre las posibles clases de la escena, mediante una búsqueda exhaustiva. A medida que el número de clases de una imagen aumenta, el método Otsu necesita mucho más tiempo para seleccionar un umbral multinivel adecuado [5].

El método de Otsu mejora la binarización clásica, debido a que aumenta la precisión en la umbralización y se describe como:

Si se aplica un umbral $U \in \mathbb{R}$ a la imagen digital en escala de grises $f(x, y) \in \mathbb{R}^{m \times n}$, donde $[x, y]^T \in \mathbb{R}^2$ son las coordenadas espaciales, $m \in \mathbb{R}$ es el número de filas y $n \in \mathbb{R}$ es el número de columnas; esta quedará binarizada, etiquetando el resultado con 1, si los píxeles corresponden al objeto o ROI, y con 0 si no corresponden o son del fondo. Si los objetos son claros con respecto al fondo, se aplica la siguiente fórmula:

$$g(x, y) = \begin{cases} 1 & \leftrightarrow f(x, y) > U \\ 0 & \leftrightarrow f(x, y) \leq U \end{cases} \quad (4)$$

La umbralización utiliza como referencia el histograma de la imagen, o su función de densidad de probabilidad (probability density function - PDF) $P(L)$, donde L es el nivel de grises de la imagen. El número de píxeles con el nivel de gris “ i ” se denota como “ f_i ”, y la probabilidad de ocurrencia del nivel de gris “ i ” en la imagen está dada por:

$$p_i = \frac{f_i}{N * M} \quad (5)$$

3.3.2. Detector de bordes de Marr-Hildreth.

Propuesto por Marr y Hildreth en 1980, es un método de detección de bordes que se utiliza para resaltar detalles de alta frecuencia y detección de ROI o segmentación. Los bordes son curvas continuas caracterizadas por variaciones fuertes y rápidas de nivel de intensidad [10].

Para obtener los bordes de una imagen $f(x, y)$, primero se realiza la convolución de la imagen con una matriz G , que modela una función Gaussiana bidimensional (filtro de suavizado o pasa-bajas):

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (6)$$

Donde σ representa la desviación estándar y se asume valor medio igual a cero. La convolución de una imagen con esta función reduce el ruido, y el grado de suavización dependerá de la varianza de la función Gaussiana.

Luego de este suavizado se aplica el operador Laplaciano (filtro pasa-altas), que junto con el operador Gaussiano resulta en el operador

$$\nabla^2 G = \frac{\partial^2 G}{\partial x^2} + \frac{\partial^2 G}{\partial y^2} \quad (7)$$

ó

$$\nabla^2 G = \left(\frac{x^2+y^2-2\sigma^2}{\sigma^4} \right) e^{-\frac{x^2+y^2}{\sigma^2}} \quad (8)$$

Al resultado obtenido con este operador se le hace una detección de los pixeles que cruzan por cero (bordes). Estos pixeles son aquellos tal que el Laplaciano de dos de sus vecinos en posiciones opuestas tienen distinto signo. Normalmente se considera un valor umbral para el valor absoluto de la diferencia numérica entre posiciones opuestas para considerar que un píxel es de paso por cero. Para mejorar la detección se pueden definir dos umbrales que distinguen entre bordes suaves y fuertes.

3.3.3. Transformación *watershed*.

Una imagen en escala de grises puede interpretarse desde otro punto de vista como una imagen topográfica de un relieve terrestre. En este sentido, se puede decir que las intensidades de gris de mayor amplitud pueden corresponder a llanuras o montañas, mientras que las intensidades de menor valor se corresponden con valles y ríos.

A partir de esta similitud, la transformación *watershed* se define como una técnica del procesamiento digital de imágenes, que mediante la saturación de los valles es capaz de reconocer los contornos de zonas topográficas similares en toda la imagen. Esta transformada es un método de segmentación basado en regiones, que clasifica los píxeles de acuerdo a su proximidad espacial, el gradiente de sus niveles de gris y la homogeneidad de sus texturas [10].

La versión más avanzada de esta técnica es la que se conoce como controlada por marcadores, y es la que se utilizará más adelante. Su principio es fundamentalmente identificar marcadores de objetos o ROI en el frente de la imagen (mayores elevaciones) y marcadores de objetos o ROI en el fondo. Entonces el algoritmo de *watershed* se aplica solo a aquellas regiones que se hayan identificado previamente. El objetivo de esta mejora es evitar sobre-segmentación o aparición de muchas ROI en el resultado final.

3.3.4. *K-means clustering*.

Es uno de los algoritmos de aprendizaje no supervisado más eficientes para resolver el problema del agrupamiento (*clustering*) de objetos o ROI. Este método clasifica los objetos en un cierto número prefijado de grupos o *clusters*, haciendo uso de la distancia de los objetos al centroide del grupo [9] [8].

El algoritmo se ejecuta en los siguientes pasos:

- Se sitúan k puntos aleatoriamente en el espacio en los que viven los objetos que se quieren clasificar. Estos puntos representan los centroides iniciales de los grupos.
- Se asigna cada objeto al grupo con el centroide más cercano.
- Tras haber asignado todos los objetos, se recalculan las posiciones de los k centroides [8] [9].

- Se repite el segundo y tercer paso hasta que los centroides se mantengan estables. Esto produce una clasificación de los objetos en grupos que se pueden distinguir con respecto a una métrica.

Aunque se puede probar que este algoritmo siempre termina (converge), no siempre la distribución que se alcanza es la más óptima, ya que es muy sensible a la variación de condiciones iniciales [9].

4. METODOLOGÍA

Este proyecto se desarrolló en cuatro fases relacionadas con las actividades que cada objetivo específico establece para la consecución del objetivo principal. La metodología no fue secuencial, de tal forma que las actividades de cada fase se traslaparon en el tiempo que duró esta investigación.

Para poder cumplir con los objetivos, se desarrolló inicialmente el estudio de algunas de las técnicas actuales utilizadas en el procesamiento y análisis de imágenes digitales de muestras microscópicas. Al mismo tiempo, se implementó en la clínica el proceso de adquisición de imágenes de muestras para la conformación de la base de imágenes digitales.

Teniendo en cuenta las características de las muestras que se recopilan en la clínica y el proceso manual de conteo que se lleva a cabo en el laboratorio, se propuso el proceso de automatización del mismo. Este proceso consta de tres etapas: restauración, segmentación y conteo.

Finalmente, con los resultados encontrados en el proceso de automatización y con las muestras recopiladas hasta el momento, se procede a depurar la aplicación que será utilizada en la clínica. En las siguientes secciones se describen detalladamente los resultados del proceso de la automatización.

4.1. TRABAJO EN EL ÁREA

En la actualidad, los servicios de cuidado de pacientes en clínicas veterinarias van de la mano con los avances tecnológicos. Los datos por paciente que se adquieren en los distintos procesos de atención y en los exámenes de laboratorio requieren estar disponibles en los instantes normales de tiempo en los que deberían estar. A través de aplicaciones desarrolladas en hardware y software, y aprovechando los sistemas de comunicación de datos alámbricos e inalámbricos, ahora es posible el desarrollo de sistemas de gestión de servicios médicos en los que los profesionales de la salud acceden en línea a la información de los pacientes. Dichos sistemas de gestión son comunes en las denominadas Ciudades Inteligentes [21].

El propósito, al abordar la automatización del proceso de conteo de células de muestras de pacientes en la Clínica Veterinaria Carlos Martínez Hoyos de la Universidad de Nariño, es precisamente el de contribuir al mejoramiento de los procesos que en ella se llevan a cabo, de la mano del avance tecnológico y en especial en lo que se refiere a la aplicación del procesamiento y análisis digital de imágenes.

En la CVCMH, el conteo de células en muestras de pacientes se hace por medio de tecnología de bajo costo. La cámara de Neubauer es todavía el método utilizado para realizar el conteo manual de células, y como tal posee algunos inconvenientes en el proceso de realización y en el de tiempo de ejecución, que afecta tanto al profesional encargado del laboratorio como al paciente que espera por los resultados de los exámenes.

Dado que la propuesta de automatización consiste en el procesamiento y análisis de imágenes digitales de las muestras, en esta sección describe los principales trabajos consultados que realizan no solo el conteo sino la detección y análisis de las principales células presentes en muestras de sangre de animales.

- Con respecto al estudio de células blancas, el estudio microscópico es generalmente desarrollado a partir de imágenes en alta resolución y con el objetivo de extraer características propias de las células. Estas características se extraen cuando se identifican regiones de interés (ROI) en las imágenes y a través de la combinación de técnicas básicas y avanzadas del procesamiento de imágenes. Por ejemplo en [22], los autores proponen el uso de algoritmos de umbralización, como el método de Otsu y el algoritmo de *snake* o de contorno activo que usa el flujo del vector gradiente. Otra técnica utilizada en [7] y de uso común es la segmentación controlada *watershed* y en espacios de color, como el HSL. Esta misma técnica, junto con gradiente de textura y la transformación Wavelet, es utilizada en [23] para la extracción del núcleo de los leucocitos.
- Para realizar la agrupación de distintos tipos de células, los autores utilizan métodos de procesamiento y análisis de imágenes junto con técnicas de aprendizaje de máquina. En [18] por ejemplo, los autores proponen un método de agrupamiento de dos tipos de células blancas a partir de los momentos pseudoZernike junto con una máquina de soporte vectorial (Support Vector Machine - SVM). Estos momentos se eligen porque son independientes de la rotación de la imagen. Dada la calidad de las imágenes que se emplean en este estudio, es posible recuperar características físicas e indirectamente químicas en el espacio de color *Lab*. En [6], para un problema similar de clasificación, los autores emplean el operador *self-dual*

multiscale morphological toggle junto con una segmentación *watershed* mejorada. Además, se utilizan métodos denominados *Level-Set* para la detección del núcleo, y transformación *Top-Hat* y apertura de área, para detección del citoplasma.

- Otra de las técnicas que se utilizan para la clasificación de células es *k-means clustering*. Por ejemplo, en [24], los autores lo utilizan para clasificar cuatro tipos de leucocitos. En este estudio se concluye que la circularidad es la característica geométrica que mayor exactitud da en la clasificación. En [23], para realizar el conteo de las células, también se utiliza *k-means clustering* junto con un modelo *ensemble-multiclass SVM*.
- Para el conteo automático de células, la etapa de procesamiento de mayor interés es la segmentación. En el estudio realizado en [23], los autores presentan una evaluación de los principales métodos utilizados en esta etapa: Umbralización Otsu, *k-means clustering*, *watershed*, y el detector *canny edge*. Ellos concluyen que la función de distribución acumulativa del error de falla de clasificación es suficiente para elegir el mejor de los métodos; resultando que *watershed* y *k-means clustering* son los métodos con mayor exactitud.

De acuerdo a lo analizado en esta revisión bibliográfica, se propuso finalmente la utilización de los algoritmos de *watershed* y *k-means clustering* para realizar la segmentación de las imágenes y mejorar el conteo de las células bajo estudio. Junto con estas técnicas se utilizó también operaciones morfológicas como erosión, dilatación y apertura y cierre de áreas. De igual forma se evidenció las ventajas de utilizar umbralización adaptativa (método de Otsu).

4.2. AUTOMATIZACIÓN DEL PROCESO DE CONTEO DE CÉLULAS

Teniendo en cuenta las metodologías consultadas para el procesamiento y análisis de imágenes digitales microscópicas, y la forma de conteo manual de células que se lleva a cabo en los laboratorios de la CVCMH de la Universidad de Nariño, se propone una metodología de automatización de este proceso. Esta automatización conserva el uso de la cámara de Neubauer (por su bajo costo) e implementa un sistema automático de conteo de células a partir del procesamiento y análisis de las imágenes digitales tomadas sobre dicha cámara.

Con el método propuesto se pretende reducir los porcentajes de error asociados al método de medición, y no al método de estimación de la concentración en la cámara. El error debido al cansancio visual del operador (cuando se tiene que observar varias muestras) o de las fuentes de ruido que afectan la imagen observada en el microscopio también se reduce con el uso del procesamiento de

las imágenes digitales. Cabe aclarar que la propuesta de automatización no incluye la toma automática de las muestras, esto se encuentra fuera del alcance del proyecto. El método propuesto se muestra en la **Figura 7**.

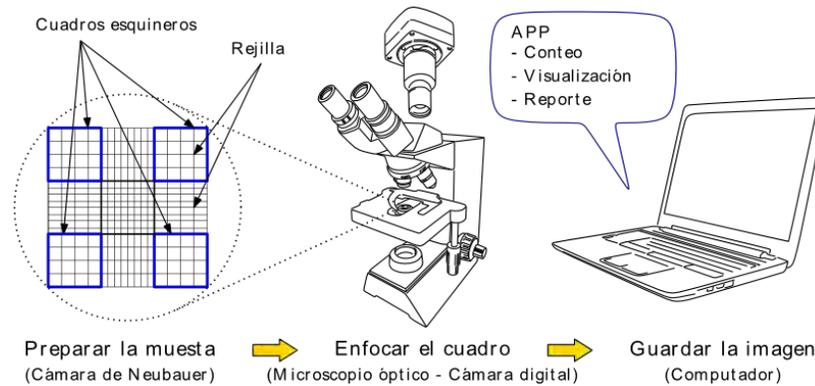


Figura 7. Proceso de automatización del conteo de células en el laboratorio de la CVCMH.

En **Figura 7**, los procesos de preparación de la muestra en la cámara hasta su enfoque en el microscopio se siguen realizando tal y como lo plantea la guía de procedimientos de la clínica [4]. Pero el registro de la imagen de microscopía se hace a través de una cámara digital y se almacena en un computador. En el computador se realiza el conteo automático de células, se visualiza los resultados más relevantes del procesamiento y se genera un reporte de acuerdo a las especificaciones del operador del laboratorio. En la **Tabla 2** se detallan los requerimientos principales exigidos por el encargado del laboratorio. Teniendo en cuenta los requerimientos planteados en la **Tabla 2**, y asegurando el cumplimiento de los objetivos de la tesis, en esta investigación se enfocó el desarrollo de la automatización tomando como ejemplo principalmente el conteo de glóbulos blancos o leucocitos. En lo que sigue se detalla el desarrollo de cada uno de los subsistemas y etapas del proceso de automatización.

| Número del requerimiento | Requerimiento |
|--------------------------|---|
| 1 | Automatizar el proceso de conteo celular |
| 2 | Priorizar el conteo de GB |
| 3 | Realizar una interfaz gráfica |
| 4 | Identificar automáticamente el área de conteo |
| 5 | Generar reporte del conteo realizado |

Tabla 2.Tabla de especificaciones del operador del laboratorio.

4.2.1. Sistema de adquisición de imágenes.

Para digitalizar las muestras observadas en el microscopio óptico del laboratorio de la CVCMH (Olympus CH30RF100 [4]), se utiliza una cámara digital (en el rango visible del espectro) diseñada especialmente para adaptarse a microscopios que carecen de esta tecnología. La cámara utilizada en el proceso de automatización es una Amscope [25] de 3,0 Mpíxeles.

La instalación al microscopio se realiza mediante lentes y adaptadores mecánicos (la versión del microscopio no permite la adaptación directa de la cámara) y se conecta directamente al computador a través de un puerto USB 2.0. En la Figura 8 se muestra la cámara con todos los accesorios e instalada en el laboratorio de la CVCMH.



Figura 8. Sistema de adquisición de imágenes.

Para la toma de las muestras se opta por utilizar el programa propio de la cámara (Amscope), previamente ajustado a las condiciones del laboratorio. Se decide dejar configurado el instrumento para tomar muestras en calidad HD, de 1536 x 2048 píxeles (alto y ancho) y en formato RGB, para obtener imágenes de alta resolución. En la Figura 9 se muestra un ejemplo de la imagen del cuadro esquinero superior izquierdo de la cámara de Neubauer para una muestra de sangre de canino. En esta figura se puede observar claramente la rejilla de la cámara y los leucocitos, que corresponden a los puntos más oscuros del frente de la imagen. Sin embargo, su identificación y conteo depende de la experiencia del encargado del laboratorio.

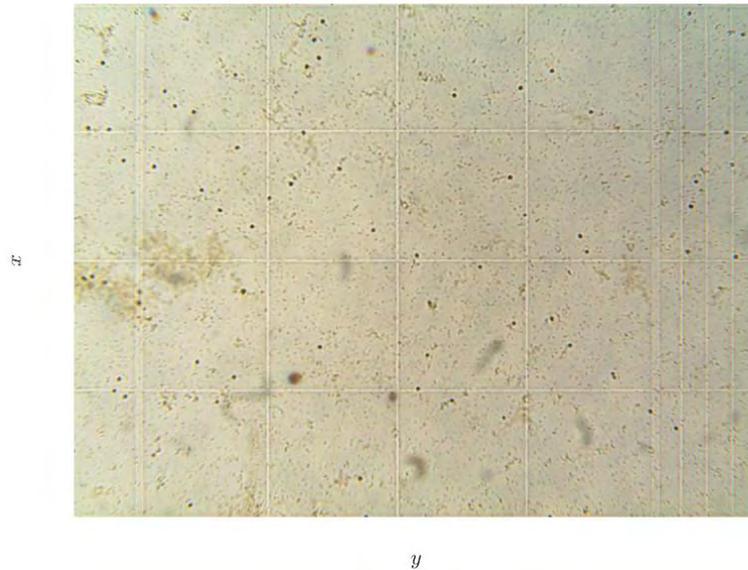


Figura 9. Imagen del cuadro esquinero superior izquierdo de la cámara de Neubauer.

La representación matemática de la imagen que se utilizará de aquí en adelante será:

$$g_{rgb}(x, y) \in \mathbb{R}^{m \times n} \quad (9)$$

donde el subíndice rgb hace referencia al espacio de color de la imagen [canales rojo (r - red), verde (g - green) y azul (b - blue)], x y y son las coordenadas espaciales a lo largo de las filas y las columnas, m es el número de filas y n es el número de columnas. Cada pixel de la imagen se codifica con 8 bits (formato entero 8 sin signo, $uint8$); por lo tanto, el número de niveles de gris por cada canal es igual a $L = 2^8 = 256$ niveles.

Nota: Como requisito esencial en el sistema de adquisición de imágenes, se exige el mantenimiento de condiciones ambientales constantes en el laboratorio donde se sitúa el montaje. Es decir, principalmente, se exige la exposición de los instrumentos a un ambiente controlado de iluminación, para no introducir más fuentes de error en el proceso de adquisición de los que se prevén o para los cuales el sistema ha sido diseñado y calibrado.

Si se mantienen las condiciones de adquisición, y teniendo en cuenta la magnitud de transformación de Fourier de la imagen digital (Ver Figura 10, izquierda), entonces estas se pueden clasificar como imágenes de baja frecuencia, en las que se concentra la mayor cantidad de la información de las células bajo estudio, y con algunas componentes de alta frecuencia, debidas a la rejilla y a la forma de las células. Por otra parte, y debido a las condiciones de iluminación, las imágenes pueden clasificarse de acuerdo a su función de densidad de probabilidad PDF, como claras y de bajo contraste (ver Figura 10, derecha).

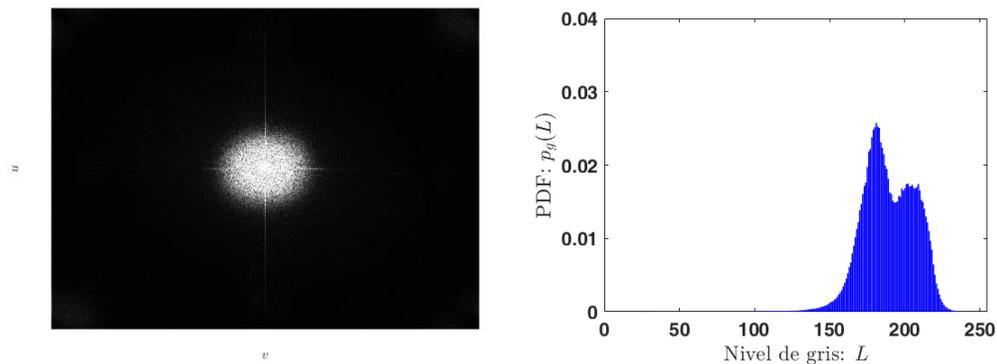


Figura 10. Izquierda: Magnitud de la transformación de Fourier (fft) de la imagen adquirida. Derecha: PDF de la imagen adquirida.

La transformada de Fourier 2-D (en dos dimensiones) de la imagen $g_{rgb}(x, y)$ se define como:

$$G_{rgb}(u, v) = \mathcal{F}[g_{rgb}(x, y)] \quad (10)$$

Donde u son las componentes de frecuencia a lo largo de las filas y v a lo largo de las columnas. En la Figura 10, el espectro (magnitud) se calculó con el algoritmo de la transformada rápida de Fourier *fft* (*Fast Fourier Transform*) y se encuentra desplazado al centro, es decir, el centro de la imagen corresponde a las frecuencias (0,0). Se aclara que los rangos de frecuencia se toman en el rango $(-\pi, \pi)$, es decir se ha tomado la separación entre píxeles como 1. Y la PDF $P_g(L)$ se calculó a partir del histograma normalizado de la imagen, previamente transformada de RGB a niveles de gris; es decir, es una PDF de todas las contribuciones de los canales de color, y no por cada color.

Las imágenes que se capturan con el sistema de adquisición se almacenan en una estación portátil con las siguientes características: procesador Intel® Core™ i5 de 2,4 GHz y 4 GB de RAM. Junto con el encargado del laboratorio de la CVCMH, se decide implementar la nomenclatura mostrada en la **Tabla 3** para la identificación de los archivos de imagen que contienen las muestras.

| GLÓBULOS BLANCOS | |
|---|--------------------------------------|
| Nombre del archivo: | IdEspecie_IdPaciente_Fecha_GB_SXX_CX |
| IdEspecie: identificador del tipo de especie a la que pertenece el paciente; por ejemplo, C: canino. | |
| IdPaciente: número de 5 dígitos (0-9) asignado como identificador único del paciente. | |
| Fecha: número en el siguiente formato: dia_mes_año. | |
| GB: identificador del tipo de célula analizada. | |
| SXX: número de la muestra tomada al paciente, de 2 dígitos (00 - 99). | |
| CX: número del cuadro de la muestra. Para glóbulos blancos: 1 - Superior izquierdo; 2 – Superior derecho; 3 – Inferior izquierdo y 4 – Inferior derecho. | |
| Ejemplo: C_10012_01_10_2017_GB_S11_Q3.jpg | |

Tabla 3. Nomenclatura para almacenamiento de imágenes de GB.

Finalmente, la estructura del lugar de almacenamiento de la base de datos de las imágenes en el computador se define como: un directorio principal que contiene las subcarpetas por especie:

Laboratorio\Caninos

Laboratorio\Felinos

Laboratorio\...

Y un subdirectorío nombrado con número del identificador único de cada paciente; por ejemplo: *Laboratorio\Caninos\10012*

4.2.2. Restauración de las imágenes adquiridas.

Los algoritmos que se proponen en esta sección tienen como fin la restauración de las imágenes capturadas por el sistema de adquisición de imágenes. Los algoritmos están propuestos de manera general para cada tipo de célula y han sido probados principalmente en células blancas de caninos.

Teniendo en cuenta los resultados mostrados en las **Figuras 9 y 10**, se logra la identificación de las siguientes fuentes de ruido que afectan las imágenes:

- Un patrón de interferencia que corresponde a la rejilla de la cámara de Neubauer.
- Un ruido en el fondo de la imagen (conocido como *background*) que se asocia a otras células, del mismo y distinto tipo, y a la iluminación no uniforme en la cámara de Neubauer.

Estas fuentes de ruido pueden influir negativamente en el proceso de conteo automático de ROI dentro de la imagen. Algoritmos básicos que hacen umbralización o reconocimiento de objetos se pueden ver afectados por los bordes de la rejilla y el fondo.

A partir de la experiencia que se posee en el campo de la restauración de imágenes digitales [10], para la eliminación del ruido presente en las imágenes adquiridas, se supone que este afecta a la imagen de acuerdo al siguiente modelo:

$$g_i(x, y) = h_i(x, y) * f_i(x, y) + n_i(x, y) + \omega_i(x, y) \quad (11)$$

donde $i = \{r, g, b\}$ indica el canal de color; $g_i(x, y)$ es la imagen degradada, $f_i(x, y) \in \mathbb{R}^{m \times n}$ es la imagen real desconocida, $h_i(x, y) \in \mathbb{R}^{m \times n}$ es la función de degradación desconocida; $\eta_i(x, y) \in \mathbb{R}^{m \times n}$ es el patrón de interferencia y $\omega_i(x, y) \in \mathbb{R}^{m \times n}$ es un ruido aditivo.

De acuerdo a este modelo, en este trabajo se plantea un proceso de reconstrucción de la imagen a partir de filtros locales adaptativos, que se basan en filtrado óptimo (en algún sentido, que se especificará más adelante).

- **Remoción del patrón de interferencia.** $n_i(x, y)$ se compone de las líneas horizontales y verticales grabadas en la cámara de Neubauer (ruido aditivo), y que se consideran no útiles en el proceso de detección y conteo automático en toda el área de la imagen, pero sí al momento de realizar el cálculo en el área considerada por la ecuación (3). Su identificación y remoción se puede realizar por medio de algoritmos de detección de bordes; sin embargo, estos mismos algoritmos podrían afectar las componentes de alta frecuencia de las formas de las células o ROI a ser contadas. Por lo tanto, se propone la extracción de las componentes de frecuencia principales de la rejilla, localizadas en los ejes vertical (líneas horizontales) y horizontal (líneas verticales) del espectro de la imagen.

El filtrado se hace a través de un filtro de muesca ideal (*Ideal Notch-Pass - INP*)

$$H_{INP}(u, v) \in R^{m \times n} \quad (12)$$

donde $|H_{INP}(u, v)| = 1$ para las frecuencias de la banda de paso (a lo largo de los ejes) y $|H_{INP}(u, v)| = 0$ para las frecuencias de la banda de rechazo. (Ver Figura 11).

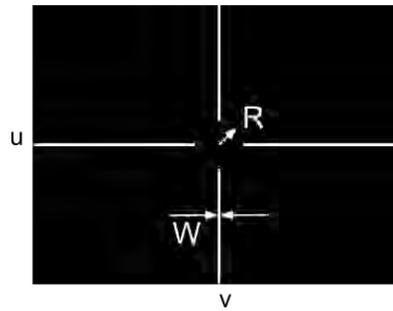


Figura 11. Magnitud $|H_{INP}(u, v)|$ del filtro ideal de muesca.

Como se puede apreciar en la **Figura 10 (izquierda)**, las componentes de frecuencia de la rejilla son impulsos a lo largo de los ejes de frecuencia; sin embargo, para evitar quitar las componentes de baja frecuencia útiles de las células, se adicionan los siguientes parámetros al filtro: R es el radio del círculo de frecuencias que no se eliminan, medido desde el centro de la imagen del espectro; y W , que es ancho de la banda removida a lo largo de los ejes. Estos parámetros se ajustan manualmente de acuerdo a su desempeño sobre el algoritmo de restauración.

Antes de aplicar el filtro descrito en la ecuación (12), y dado que se pueden presentar discontinuidades en la rejilla, se decide enfatizar los bordes de la imagen a través de un filtro de énfasis de alta frecuencia (*High-Frequency Emphasis - HFE*) [10].

$$H_{HFE} = a + b * H_{HP} \quad (13)$$

donde $H_{HP} \in R^{m \times n}$ es un filtro pasa-altas (High-Pass HP) y $\{a, b\} \in \mathbb{R}$ son parámetros de sintonización que cumplen las restricciones $a \geq 0$ y $b > a$ para mejores resultados. En nuestro caso utilizamos un filtro pasa-altas de tipo Butterworth, con frecuencia de corte $D_c = \sqrt{u_c^2 + v_c^2}$ y orden n donde $\{u, v\} \in \mathbb{R}$ son las coordenadas de la función transformada.

Luego, el patrón de ruido en el espacio, obtenido con la aplicación de los dos filtros es:

$$n_i(x, y) = \mathcal{F}^{-1}\{H_{INP}\mathcal{F}\{g_{i,HFE}(x, y)\}\} \quad (14)$$

Donde

$$g_{i,HFE}(x, y) = \mathcal{F}^{-1}\{H_{HFE}G_i(u, v)\} \quad (15)$$

Y

$$G_i(u, v) = \mathcal{F}\{g_i(x, y)\} \quad (16)$$

El operador $\mathcal{F}^{-1}\{\cdot\}$ es la transformación inversa 2-D de Fourier.

La **Figura 12** muestra un área (512 x 512 pixeles) de los resultados de la aplicación de los filtros a la imagen que se muestra en la **Figura 10**. De arriba abajo se muestran la imagen con énfasis de alta frecuencia, $g_{i,HFE}(x, y)$ y el patrón de interferencia, $n_i(x, y)$. De izquierda a derecha son los canales rojo, verde y azul. Los parámetros utilizados en la implementación de los algoritmos fueron: $a = 1.0$, $b = 1.5$, $D_c = 0.2\pi$ rad, $n = 8$, $R = 5$ pixeles y $W = 4$.

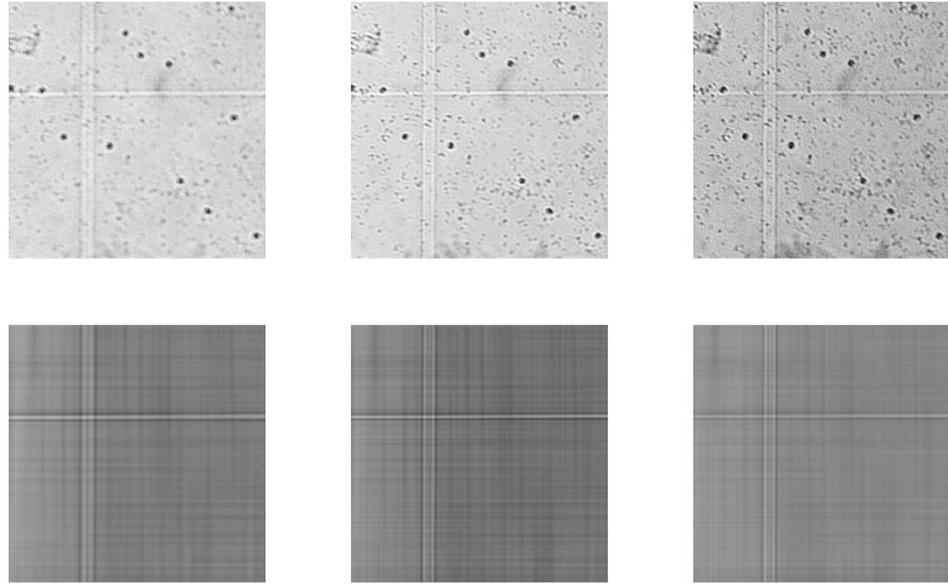


Figura 12. Arriba abajo: imagen con énfasis de alta frecuencia y el patrón de interferencia. De izquierda a derecha: canales rojo, verde y azul.

Una vez identificado el patrón de interferencia $n_i(x, y)$, su remoción de la imagen $g_{i,HFE}(x, y)$ se puede intentar con una simple resta de las imágenes en el espacio. Sin embargo, esta técnica no ofrece los mejores resultados y se opta por aplicar un filtro de muesca óptimo [10] pero local (en un vecindario) (*Local Notch Optimum* - LON). Es decir, se aprovecha las varianzas locales de la imagen para ajustar la resta ponderada del patrón de interferencias. La imagen en el espacio procesada con este filtro se puede escribir como:

$$g_{i,LON}(x, y) = g_{i,HFE}(x, y) - m_i(x, y) * n_i(x, y) \quad (17)$$

donde $m_i(x, y) \in \mathbb{R}$ se conoce como la función de ponderación o de modulación, que se calcula con la minimización de la varianza de la resta de las imágenes

$$g_{i,HFE}(x, y) - g_{i,LON}(x, y) \quad (18)$$

Su valor está dado por [10]:

$$m_i(x, y) = \frac{\overline{g_i(x,y)n_i(x,y)} - \overline{g_i(x,y)} \overline{n_i(x,y)}}{n_i(x,y)^2 - (\overline{g_i(x,y)})^2} \quad (19)$$

El cálculo de $m_i(x, y)$ se propone local porque la dilución de las células en la cámara de Neubauer no es uniforme, al igual que la iluminación interna del microscopio. El efecto de aplicar el filtro de manera local tiene la ventaja de suavizar las PDF locales; pero hace que los vecindarios en la imagen se desplacen a diferentes niveles promedio de gris (similar a un tablero de ajedrez).

Este efecto de desplazamiento, en vez de ser perjudicial en la restauración, se utiliza para hacer una transformación de ajuste o *matching* local (*Local Matching* - LM) de histograma [10], que es controlada por un histograma (o PDF) total calculado como la suma de los histogramas (o PDF) de todos los vecindarios. Esta transformación logra mantener la forma original del histograma (o PDF) de la imagen y lo centra en la escala de niveles de gris (el filtro LON es óptimo en promedio [10]). El efecto final del *matching* es el de hacer uniforme la iluminación de la imagen. Esta transformación se puede escribir como:

$$g_{i,LM}(x, y) = \mathcal{T}_{LM}\{g_{i,LON}(x, y)\} \quad (20)$$

donde $g_{i,LM}(x, y)$ es la imagen estimada y $\mathcal{T}_{LM}\{\cdot\}$ es el operador de la transformación de *matching*.

La **Figura 13** muestra los resultados del filtro LON y la transformación LM aplicados a las imágenes de la **Figura 12** (arriba). De arriba abajo se muestran la imagen $g_{i,LON}(x, y)$, la imagen $g_{i,LM}(x, y)$ y su PDF. De izquierda a derecha son los canales rojo, verde y azul. En esta imagen se aprecia el efecto del tablero de ajedrez del filtro LON, la mejora que resulta del *matching* y la forma de la PDF y su ubicación en el centro de los niveles de gris. A partir de varios experimentos, por prueba y error, se concluyó que un tamaño de vecindario apropiado (que no deforma la forma original de la PDF) es de 128 x 128 píxeles. La PDF observada en las imágenes tiene una forma prácticamente igual a una función Gaussiana.

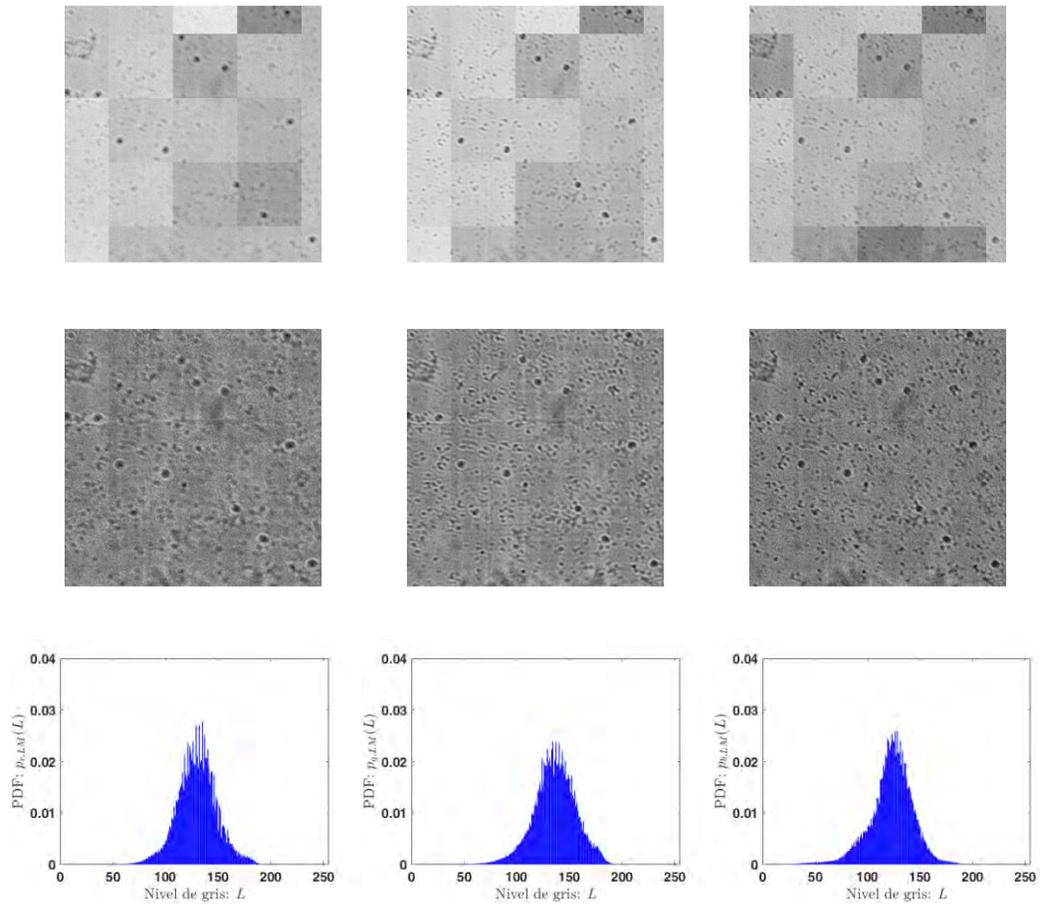


Figura 13. Remoción del patrón de interferencia. Arriba abajo: imagen LON, imagen LM y PDF. De izquierda a derecha: canales rojo, verde y azul.

- **Remoción del ruido de fondo.** Este ruido corresponde al ruido aditivo $\omega_i(x, y)$ y la función de degradación $h_i(x, y)$, cuyas fuentes de origen se asocian principalmente a iluminación no apropiada, desenfoque, desviación del haz de luz cuando pasa a través del fluido y el cristal de la cámara y poblaciones de células en el fondo de la imagen.

Teniendo en cuenta la PDF de las imágenes restauradas en la sección anterior (**ver Figura 13**), el fondo de la imagen se puede asociar a una PDF de forma Gaussiana, de acuerdo al modelamiento de esta. Luego, para obtener una versión de la imagen sin ruido de fondo, se propone un método de restauración que también utiliza un filtro LON. Para esto, primero se necesita identificar un patrón de ruido $\omega_i(x, y)$, que en esta investigación se hace a través de un filtro adaptativo local de tipo Wiener (W). Este filtro recupera una versión mejorada $g_{i,W}(x, y) \in \mathbb{R}^{m \times n}$ de la imagen $g_{i,LM}(x, y)$; y

luego, proponiendo una sustracción, resulta que el patrón de ruido se puede modelar como:

$$\omega_i(x, y) = g_{i,LM}(x, y) - g_{i,w}(x, y) \quad (21)$$

Con este ruido, la aplicación del filtro LON sobre $g_{i,LM}(x, y)$ resulta en la imagen

$$f_{i,LON}(x, y) = g_{i,LM}(x, y) - m_i(x, y)\omega_i(x, y) \quad (22)$$

donde $\omega_i(x, y)$ es nuevamente otra función de modulación, que es calculada de igual forma como se muestra en la ecuación (19).

Al igual que en la remoción de la rejilla, el filtro LON produce un efecto de tablero de ajedrez, que nuevamente es eliminado con la aplicación de un filtro de *matching* controlado con la PDF total de la suma de las PDF (de cada vecindario) de $f_{i,LON}(x, y)$. Este resultado se denomina $f_{i,LM}(x, y)$ y es el resultado final del proceso de restauración propuesto.

Para comparar los resultados (de forma visual) de los algoritmos propuestos, se compara el resultado final de la restauración con un filtro de uso común en eliminación de ruido de fondo: el filtro *deconvolution* (D). [10] La imagen obtenida por este filtro se denomina $f_{i,D}(x, y) \in \mathbb{R}^{m \times n}$.

La **Figura 15** muestra los resultados del filtro LON y la transformación LM aplicados a la imagen de la **Figura 13 (segunda fila)**. De arriba abajo se muestran la imagen del ruido $\omega_i(x, y)$, la imagen LM $f_{i,LM}(x, y)$ y la imagen $f_{i,D}(x, y)$. De izquierda a derecha son los canales rojo, verde y azul. Para el filtro Wiener se logró mejores resultados con un vecindario igual a 12 x 16 píxeles y para los filtros LON y LM de 128 x 128 píxeles. La comparación visual entre las imágenes LM y D concluye que esta propuesta logra eliminar y suavizar de manera más fuerte el fondo de la imagen. El filtro D preserva artefactos del fondo que no son convenientes para las siguientes etapas a desarrollar.

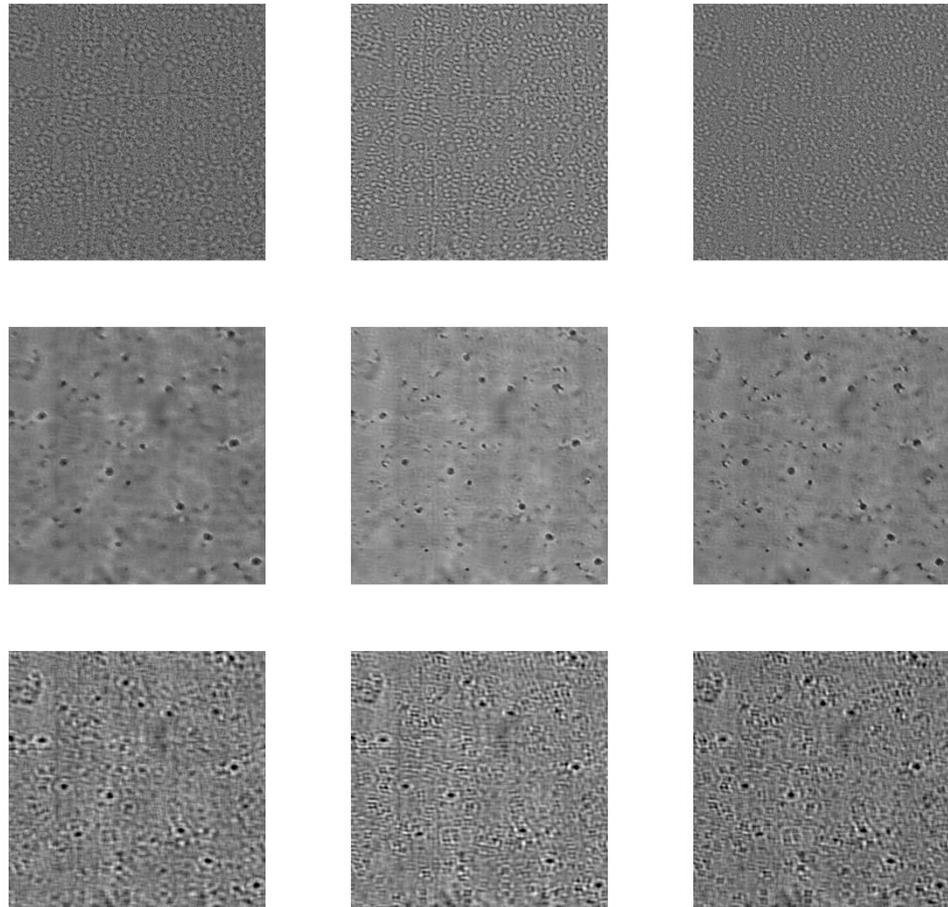


Figura 14. Remoción del fondo. Arriba abajo: imagen de ruido, imagen LM e imagen D. De izquierda a derecha: canales rojo, verde y azul.

4.2.3. Segmentación.

Para lograr el objetivo final de conteo sobre las imágenes restauradas, se propone el uso de técnicas avanzadas de extracción de las ROI donde se localizan las células. De acuerdo a lo investigado en la revisión de la bibliografía, las técnicas de mayor uso son *Marker-Controller Watershed (MCW)* y *k-means clustering* [10] [8], que son las técnicas probadas en esta investigación.

El problema asociado a ambas técnicas de segmentación es la sobresegmentación; es decir, la extracción de muchas ROI que pueden provocar error en el conteo final de las células. Para evitar este problema, se propone un filtro local que se denomina de énfasis de puntos oscuros (LDPE, *Local Darkest Point*

Emphasis), que se fundamenta en las siguientes características de la información contenida en la imagen:

- *Las células diluidas hacen que el vecindario donde se ubican cambie ligeramente los parámetros de su PDF, si se compara con el mismo vecindario sin las células.*
- *La concentración de células es baja en un vecindario (especialmente para los glóbulos blancos de tipo leucocitos) y se caracterizan porque los pixeles son oscuros (debido a las características propias de las células, por ejemplo, el núcleo y el citoplasma).*

Entonces, se asume que la probabilidad de detección de las células aumenta si se enfatiza la intensidad de los pixeles oscuros en cada vecindario (de ahí el filtro propuesto). Este énfasis se puede lograr con cualquier transformación localizada de intensidad. En este trabajo se propone el énfasis de los puntos oscuros por medio de una transformación exponencial que se adapta a las propiedades de la PDF del vecindario, y que se escribe como:

$$L_{out} = AL_{in}^B \quad (23)$$

donde $L_{in} \in L$ es el nivel de gris de entrada (de la imagen restaurada), $L = [0, \dots, 255]$, y $L_{out} \in L$ es el nivel de gris de salida. $A \in \mathbb{R}$ es un parámetro de ajuste manual que ayuda a la mejora del contraste y $B \in \mathbb{R}$ es calculado de forma óptima del siguiente problema de minimización:

$$\min_B \int_{\langle L_{out} \rangle} P_{i,N}(L_{out})P_{i,R}(L_{out})dL_{out} \quad (24)$$

s.t. $1 < B \leq B_{max}$

donde $i = \{r, g, b\}$, $P_{i,N}(L_{out})$ es la PDF de la imagen sin células a contar y $P_{i,R}(L_{out})$ es la PDF de la imagen con células a contar. La integral en la función a minimizar es una medida de la correlación entre las dos PDF, lo que significa que el problema en (24), a través de la transformación en (23), obtiene la máxima disimilitud entre las dos PDF, resultando finalmente en una imagen con ROI o células a contar resaltadas. En (24), $B_{max} \in \mathbb{R}$ restringe los valores calculados de este parámetro para evitar saturación de los niveles oscuros. Además, el filtro se aplica solo a aquellos pixeles que cumplen la restricción:

$$L_{in} \leq m_{i,R}/C \quad (25)$$

donde $m_{i,R}$ es el valor medio del vecindario de la imagen sin células y $C \in \mathbb{R}$ es un parámetro de ajuste manual.

Para resolver el problema en (24), se elige como imagen libre de células a contar la imagen $\omega_i(x, y)$, y como imagen con células a contar la imagen $f_{i,LM}(x, y)$, obtenidas en el proceso de restauración. Las PDF son calculadas con aproximaciones de forma Gaussiana. El problema es solucionado mediante integración numérica.

La imagen que resulta del filtro LDPE, denominada $f_{i,LDPE}(x, y) \in \mathbb{R}^{m \times n}$, se utiliza como entrada al algoritmo MCW, que es implementado en las siguientes etapas secuenciales [10]:

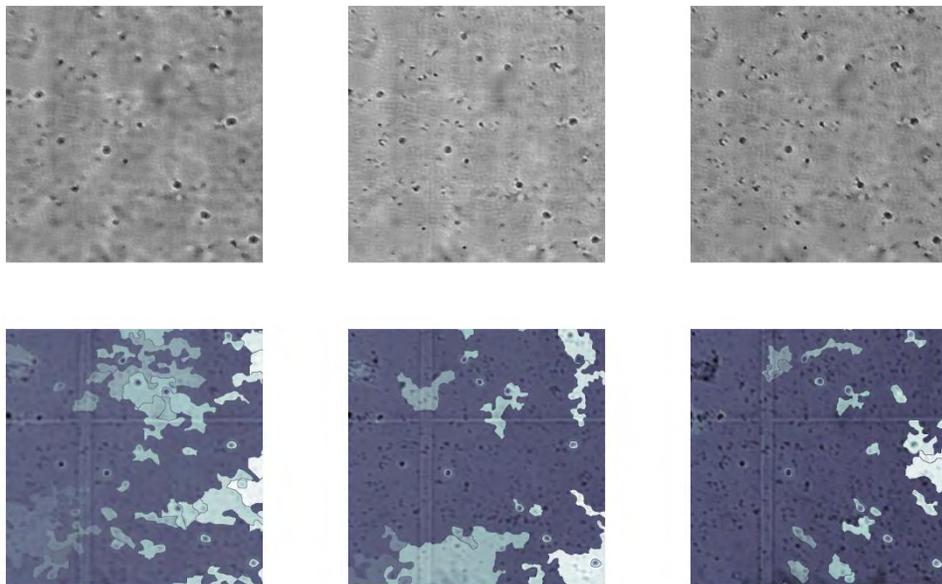
- Se escoge una función de segmentación (detector de bordes) de tipo *Marr-Hildreth edge detector* (MHED) [10], que evita la detección de bordes falsos del fondo de la imagen, por medio de un filtro de suavizado Gaussiano, y extrae los bordes más fuertes por medio de un operador Laplaciano, usando una doble umbralización en los cruces por cero. La salida de esta etapa es una imagen de segmentación denominada $M_{i,MHED}(x, y) \in \mathbb{R}^{m \times n}$.
- Los marcadores del frente de la imagen (FG, *Foreground*) se construyen mediante un filtro morfológico de apertura y cierre, que utiliza una combinación de operaciones de erosión-dilatación-reconstrucción. Este filtro se aplica a la imagen $f_{i,LDPE}(x, y)$ y su resultado es una máscara binaria que se denomina $M_{i,FG}(x, y) \in \mathbb{R}^{m \times n}$.
- El fondo de la imagen (BG, *Background*) se calcula por medio de la transformación watershed de la transformación de distancia de una máscara obtenida por medio de operaciones morfológicas de apertura, que se basan en erosión-reconstrucción. Las transformaciones se aplican a la imagen $f_{i,LDPE}(x, y)$ y el resultado de esta etapa es la imagen de *ridge lines* denominada $M_{i,BG}(x, y) \in \mathbb{R}^{m \times n}$.
- Finalmente, la transformación MCWT es aplicada sobre $f_{i,LDPE}(x, y)$ y solamente para mínimos locales señalados por las imágenes del frente $M_{i,FG}(x, y)$ y del fondo $M_{i,BG}(x, y)$ (de ahí el nombre de controlada por marcadores). La salida de esta etapa es la imagen de rótulos (que identifican las ROI identificadas) $M_{i,MCW}(x, y) \in \mathbb{R}^{m \times n}$.

Por otro lado, para comparar visualmente los resultados de la segmentación propuesta, se utiliza la segmentación calculada con *k-means clustering*. Sin embargo, su aplicación directa sobre la imagen restaurada resulta en sobre-

segmentación, para lo cual se adiciona un filtro mediana. La salida del *k-means clustering* es la imagen de rótulos $M_{i,K}(x, y) \in \mathbb{R}^{m \times n}$.

En la **Figura 15** se muestra los resultados de las segmentaciones. De arriba abajo están la imagen LDPE, $f_{i,LDPE}(x, y)$; la imagen MCW de la segmentación basada en el filtro LDPE, $M_{i,MCW}(x, y)$; la imagen $M_{i,K}(x, y)$ y una imagen de MCW sin filtro LDPE, $M_{i,MCW}^*(x, y) \in \mathbb{R}^{m \times n}$. De izquierda a derecha están los canales rojo, verde y azul. Los parámetros elegidos fueron: $A = 10^{-3}$, $B_{max} = 10$ y vecindario de 128X128 pixeles.

Para el filtro de suavizado Gaussiano del detector MHED se utilizó un valor medio de $m_{MHED} = 0$ pixeles y una desviación estándar de $\sigma_{MHED} = 8$ pixeles; con estos valores se cumple la regla del tamaño mínimo de este filtro [10]: $6\sigma_{MHED} + 1 = 49$ pixeles. Para el filtro Laplaciano, el tamaño de la máscara es de 3 x 3 pixeles y los umbrales para la detección de cruces por cero se tomaron como $u_H = 0.9L_{max}$ y $u_L = 0.01L_{max}$, L_{max} es el valor máximo del resultado del Laplaciano. Los elementos de estructuras morfológicas fueron tomados como círculos y cuadrados de tamaño 5 x 5 pixeles. Áreas o marcadores con tamaños más pequeños que 150 pixeles se eliminaron de las máscaras de fondo y frente. El filtro de mediana utilizado tuvo un tamaño de ventana de 9 x 9 pixeles.



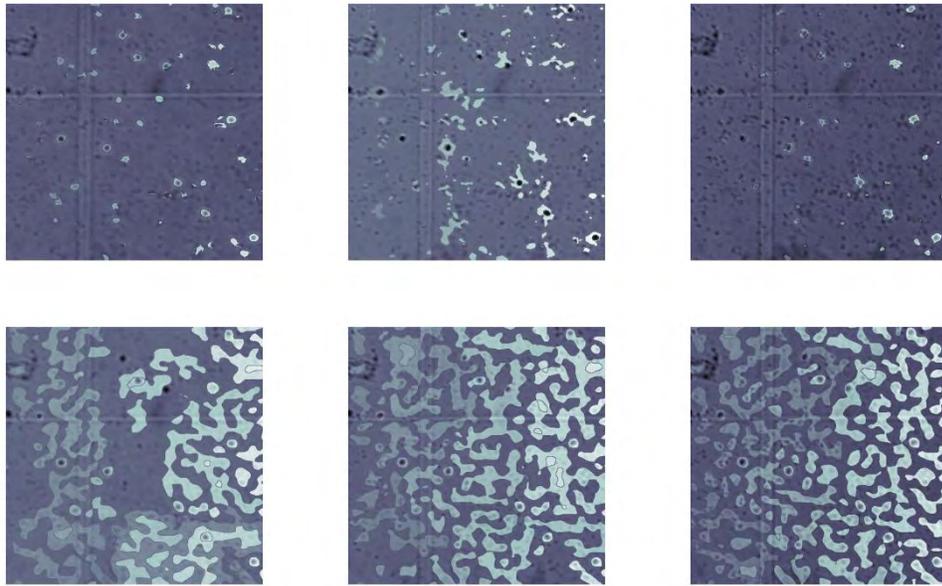


Figura 15. Segmentación. Arriba abajo: imagen LDPE, imagen MCW con filtro LDPE, imagen K (*k-means clustering*) e imagen MCW sin filtro LDPE. De izquierda a derecha: canales rojo, verde y azul.

Teniendo en cuenta los resultados observados en la **Figura 15**, se nota la mejora en la segmentación dada por los algoritmos de MCW con filtro LDPE y *k-means clustering*. Se puede decir que la propuesta de segmentación propuesta aventaja un poco al *k-means clustering*, en el sentido de que son menos las regiones encontradas. Además, la peor segmentación obtenida es la MCW sin filtro LDPE, y así se ratifica las ventajas del filtro LDPE propuesto.

4.2.4. Conteo.

Hacer el conteo de células directamente de la imagen restaurada, a través de un algoritmo de binarización, puede funcionar, pero no de manera adecuada. El error en el conteo puede incrementarse debido a que un solo umbral no puede recuperar todas las ROI donde se localizan las células. Por lo tanto, se propone hacer un conteo más exacto utilizando los resultados de la segmentación.

Al método que se propone en este trabajo se le denomina *matching template*, y significa que se utiliza una plantilla para conteo construida a partir de la imagen restaurada y con énfasis de puntos oscuros $f_{i,LDPE}(x, y)$, y en conjunto con los resultados de las imágenes segmentadas $M_{i,MCW}(x, y)$ y $M_{i,K}(x, y)$; es decir, se

restringe las áreas a contar por medio de las medidas de las ROI identificadas en la segmentación.

La operación de conteo se realiza en cada uno de los canales de color, para observar su aporte real a la exactitud de la medida. Este método se realiza de la siguiente manera:

- Se aplica una operación de umbralización a la imagen $f_{i,LDPE}(x,y)$ para obtener una máscara binaria $T_{i,1}(x,y) \in \mathbb{R}^{m \times n}$, con valores 1: células a contar, y 0: no células.
- Se asigna la intensidad equivalente al color blanco (255) a las ROI de las células a contar, utilizando $T_{i,1}(x,y)$ en $f_{i,LDPE}(x,y)$. La imagen resultante es una imagen en niveles de gris denominada $f_{i,T,1}(x,y) \in \mathbb{R}^{m \times n}$.
- Para remover regiones muy pequeñas sin células a contar, se utiliza un filtro mediana que produce una imagen en niveles de gris, que se denomina $f_{i,T,2}(x,y) \in \mathbb{R}^{m \times n}$.
- Una segunda umbralización se aplica a la imagen $f_{i,T,2}(x,y)$, para producir una segunda máscara binaria $T_{i,2}(x,y) \in \mathbb{R}^{m \times n}$, con valores 1: células a contar, y 0: no células.
- Finalmente, una operación de erosión es aplicada a la máscara $T_{i,2}(x,y)$ para limpiar y suavizar los bordes de las ROI. La plantilla final obtenida se denomina $T_i(x,y) \in \mathbb{R}^{m \times n}$.

Una vez calculada la plantilla $T_i(x,y)$, se plantea hacer un conteo, pero ajustado a algunos criterios de las ROI obtenidas en la segmentación y en la siguiente forma:

- Después de algunos experimentos, se concluye que los conteos más precisos ocurren con las ROI definidas mediante las siguientes operaciones de intersección:

$$ROI_{i,T} = T_i \cap g_i(x,y) \quad (26)$$

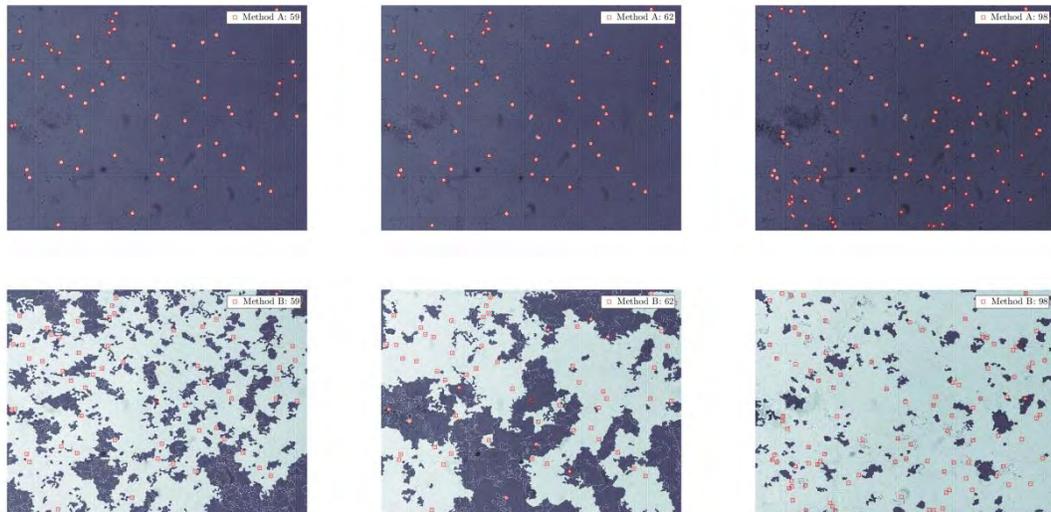
$$ROI_{i,S} = M_{i,MCW}(x,y) \cap f_{i,LM}(x,y) \quad (27)$$

$$ROI_{i,K} = M_{i,K}(x,y) \cap f_{i,LM}(x,y) \quad (28)$$

Además, solamente los pixeles más oscuros que un umbral $U_{i,int} \in \mathbb{R}$ se mantienen en la ROI y se mide su área.

- Los centroides de la plantilla $T_i(x,y)$ se utilizan para comparar las ROI obtenidas por medio de (X), y se proponen los siguientes métodos de conteo:
 - Método **A**: Si $area(ROI_{i,T}) \leq U_{i,area}$, entonces el centroide se cuenta como célula de interés. De lo contrario no.
 - Método **B**: Si $mean(ROI_{i,T}) - mean(ROI_{i,S}) \leq U_{i,mean}$, entonces el centroide se cuenta como célula de interés. De lo contrario no.
 - Método **C**: Si $mean(ROI_{i,T}) - mean(ROI_{i,K}) \leq U_{i,mean}$, entonces el centroide se cuenta como célula de interés. De lo contrario no.
- El conteo final de células se computa como el promedio de los conteos más similares de los tres canales. Se propone así porque hay un canal que introduce una cuenta exagerada y depende del tipo de célula a contar.

En la **Figura 16** se muestra el resultado de los tres métodos de conteo. De arriba abajo se observan los resultados de los métodos **A**, **B** y **C**, así como las regiones $ROI_{i,T}$, $ROI_{i,S}$ y $ROI_{i,K}$ (los centroides se dibujan en cuadrados de color rojo). De izquierda a derecha son los canales rojo, verde y azul. Los umbrales se configuraron a los valores $U_{i,int} = 70$, $U_{i,area} = 60$ pixeles y $U_{i,mean} = 20$ pixeles. Para esta muestra en particular, que contiene leucocitos de canino, sucede que la mayor precisión de conteo la proporcionan los canales rojo y verde, y el canal azul introduce muchas ROI que no corresponden a las células. El error en la medida se presenta en la sección de resultados.



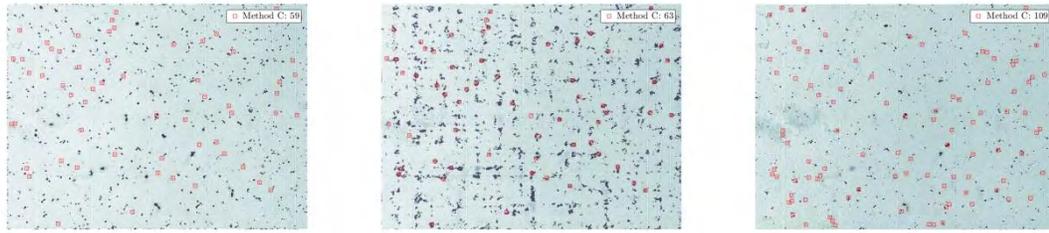


Figura 16. Conteo. Arriba abajo: método **A**, **B** y **C**. De izquierda a derecha: canales rojo, verde y azul.

4.3. DISEÑO DE LA INTERFAZ DE AUTOMATIZACIÓN

En esta sección se describen las funciones principales que se deben ejecutar en una interfaz de usuario gráfica (GUI, Graphic User Interface) para observar los resultados del proceso de conteo de células. El usuario de la interfaz es el encargado del laboratorio de la CVCMH de la Universidad de Nariño, y junto con él se llegó a un conceso acerca de las principales funciones a ser implementadas.

La interfaz se diseña para ser implementada en el entorno de MATLAB[®], y luego se exporta como aplicación App para ser instalada en el laboratorio de la clínica veterinaria. La App no requiere que el computador del laboratorio tenga instalado MATLAB[®], solo un compilador de C++ que es de libre distribución [26].

Se asume que las imágenes a procesar y analizar se encuentran ya almacenadas en la base de datos, después de seguir el protocolo exigido por la clínica. Las tareas principales a diseñar (y que son secuenciales), descritas en pseudocódigo son:

- **Cargar imagen (botón):** su función es cargar la información del archivo imagen a una variable en el entorno de MATLAB[®]. La etiqueta de este objeto para el algoritmo es: **pb_Cargar_Imagen**.
 - *Abrir un cuadro de búsqueda para cargar la(s) imagen(es) de la muestra.* Se propone hacer uso de la función **uigetfile()**, cuyos parámetros principales de entrada son: la ruta de la carpeta por defecto y el tipo de archivo. Como parámetros de salida se obtienen el nombre del archivo y su ruta, que se almacenan en las variables globales **NombreArchivo** y **RutaArchivo**.
 - *Crear una variable global para almacenamiento de los datos del archivo de imagen.* Se propone hacer uso de la función **load()**, cuyos parámetros de entrada son el nombre del archivo y la ruta, obtenidos

en el paso anterior. Los parámetros de salida se asignan a la variable global **Imagen**, tipo *uint8*, donde se almacenan los niveles de gris de cada canal: rojo, verde y azul.

- *Visualizar la variable imagen en la interfaz.* Se utiliza un eje en la imagen denominado **a_Imagen**. Se propone hacer uso de las funciones **imshow()** o **image()**.
 - *Activar controles de la interfaz.* Para seguir con el proceso de conteo se propone hacerlo mediante el cambio de la propiedad **enable**, de **off** a **on**, para aquellos objetos de la interfaz que lo requieran.
-
- **Seleccionar método (menú pop-up):** su función es la de seleccionar de un menú de opciones, relacionado con los métodos implementados en la interfaz para el conteo de células, el que el usuario desea ejecutar. Los métodos propuestos son: *Método 1* - Umbralización directa sin preprocesamiento, *Método 2* - Umbralización mediante K-means clustering y *Método 3* - Umbralización directa con preprocesamiento. La etiqueta de este objeto para el algoritmo es **pu_Metodo**.
 - *Abrir una lista de métodos disponibles para realizar el conteo.* Se propone utilizar la propiedad **string** del objeto para asignar el valor seleccionado a la variable global **Metodo**.
 - *Visualizar en la interfaz una descripción del método seleccionado.* Se propone hacerlo cambiando la propiedad **string** de una etiqueta estática **t_Descripcion_Metodo**.
 - *Activar controles de la interfaz.* Para seguir con el proceso de conteo se propone hacerlo mediante el cambio de la propiedad **enable**, de **off** a **on**, para aquellos objetos de la interfaz que lo requieran.
-
- **Descripción del método (texto estático):** Su función es la de mostrar en la interfaz una descripción pequeña y muy general del método seleccionado en el pop-up Menú. La etiqueta del objeto para el algoritmo es **t_Descripcion_Metodo**. Como es un objeto estático no tiene funciones asociadas.
-
- **Constante de calibración (barra de desplazamiento):** Su función es la seleccionar un valor para la constante que sirve para ajustar la sensibilidad del método seleccionado. Su rango es el intervalo [-5, 5]. Si se acerca a -5 (Izquierda) es más discriminante en el conteo; es decir, no cuenta las células más pequeñas. De forma contraria pasa si desliza hacia la derecha, el método incluye en la cuenta células más pequeñas. La etiqueta de este objeto para el algoritmo es **s_Constante_Calibracion**.
 - *Leer el valor seleccionado en la barra de desplazamiento.* Y almacenar su valor en la variable global **Calibracion**.

- *Visualizar el valor de la variable **calibracion** en la interfaz.* Se propone hacerlo mediante la propiedad **string** de la etiqueta estática **t_valor_Constante_Calibracion**.
- **Valor de la constante de calibración (texto estático):** Su función es la de mostrar en la interfaz el valor de la constante seleccionada en la barra de desplazamiento de la constante de calibración. La etiqueta del objeto para el algoritmo es **t_valor_Constante_Calibracion**.
- **Hacer conteo (botón):** Su función es la de ejecutar la rutina del método de conteo seleccionado. La etiqueta del botón para el algoritmo **pb_Realizar_Cuento**.
 - Leer las variables de método seleccionado, método, constante de calibración, calibración, e imagen, imagen.
 - Si método == 1, ejecutar rutina **Metodo_1**, sino
 - Si método == 2, ejecutar rutina **Metodo_2**, sino
 - Ejecutar rutina **Metodo_3**
 - Asignar el valor de salida del método de conteo a la variable global **Resultado**.
 - Visualizar en pantalla el valor del conteo. Se propone hacerlo mediante la propiedad **string** de la etiqueta estática **t_Valor_Resultado_Cuento**.
 - Activar el objeto de generar reporte. Se propone cambiar la propiedad **enable** del botón **pb_Generar_Reporte**, de **off** a **on**.
- **Valor del resultado del conteo (texto estático):** Su función es la de mostrar en la interfaz el valor del resultado del conteo. La etiqueta del objeto para el algoritmo es **t_Valor_Resultado_Cuento**.
- **Generar reporte (botón):** Su función es almacenar en la base de datos un reporte del conteo de células de la muestra. La etiqueta del botón para el algoritmo es **pb_Generar_Reporte**.
 - A partir de las variables **NombreArchivo** y **Resultado** se genera un archivo de texto con los siguientes campos: identificador del paciente, fecha, tipo de célula contada, y el número total calculado. Otros campos adicionales son el nombre del encargado de realizar el conteo y el médico que trata al paciente.
 - A partir de la variable **RutaArchivo** se guarda el archivo en la base de datos y se abre un visualizador de archivo de texto.

- **Reiniciar (botón):** Su función es reiniciar todas las funciones de la interfaz al punto de ejecución inicial; es decir, no hay imagen(es) cargada(s) y no están habilitados los controles de conteo. La etiqueta del botón para el algoritmo es **pb_Reiniciar**.
 - Elimina las variables imagen, método y calibración.
 - Limpia el eje **a_Imagen**.
 - Cambia las propiedades **string** y **enable** de los controles a sus valores por defecto.

Métodos: Se describe en pseudocódigo los métodos implementados en la interfaz. Los parámetros de entrada a cada función que los implementa son la variable que contiene la imagen (**Imagen**) y la constante de calibración (**Calibracion**). Se aclara que la imagen es una versión restaurada de la imagen original (sin la rejilla).

Método 1: Umbralización a partir del histograma.

1. Inicio.
2. Cambiar el formato de la imagen a double. Se obtiene la variable **ldouble**. Se propone utilizar la función **double()**.
3. Calcular el tamaño de la imagen **ldouble**. La variable **M** contiene el número de filas y **N** el número de columnas. Se propone utilizar la función **size()**.
4. Obtener los tres canales de color a partir de la variable **ldouble**. Se obtienen las variables **lr**, **lg** e **lb**.
5. Aplicar un filtro de suavizado de tipo Gaussiano a cada canal de color. Se obtiene las variables **lrg**, **lgg** e **lbg**. Se propone utilizar la función **imgaussfilt()**;
6. Calcular el histograma de cada canal. Se obtiene las variables **hr**, **hg** y **hb**. Se propone utilizar la función **imhist()**.
7. Visualizar los canales de color en los ejes **a_Imagen**. Se propone la función **imshow()**.
8. Calcular función de densidad de probabilidad PDF de cada canal. Se obtiene las variables **Pr**, **Pg** y **Pb**.
9. Ecuilibrar las imágenes **lrg**, **lgg** e **lbg**. Se obtiene las variables **lre**, **lge** e **lbe**.
10. Definir un umbral de niveles de gris a partir de las PDF **Ur**, **Ug** y **Ub**.
11. Realizar una umbralización de las imágenes ecualizadas **lre**, **lge** e **lbe** a partir de sus respectivos umbrales **Ur**, **Ug** y **Ub**. Se obtienen las máscaras **Mr**, **Mg** y **Mb**.
12. Mostrar las imágenes de las máscaras **Mr**, **Mg** y **Mb** en el eje **a_Imagen**. Se propone la función **imshow()**.
13. Rotular las máscaras **Mr**, **Mg** y **Mb**. Se obtienen las variables **Mrr**, **Mgr** y **Mbr**. Se propone utilizar la función **bwlabel()**.
14. Calcular áreas de las máscaras **Mrr**, **Mgr** y **Mbr**. Se obtiene las variables **Ar**, **Ag** y **Ab**. Se propone utilizar la función **regionprops()**.

15. Definir un umbral de área mínima y uno de área máxima: **Uareamin** y **Uareamax**, a partir de la variable **Calibración**.
16. Identificar áreas a contar y que estén entre los umbrales **Uareamin** y **Uareamax**. Se obtienen los identificadores de las áreas **ldr**, **ldg** e **ldb**. Se propone utilizar la función **find()**.
17. Contar áreas identificadas o número de células. Se obtiene la variable resultado a partir del promedio de cada canal.
18. Fin.

Método 2: *k-means clustering* simplificado.

1. Inicio.
2. Cambiar el formato de la imagen a *double*. Se obtiene la variable **ldouble**. Se propone utilizar la función **double()**.
3. Calcular el tamaño de la imagen **ldouble**. La variable **M** contiene el número de filas y **N** el número de columnas. Se propone utilizar la función **size()**.
4. Reformar la variable **ldouble** a vector columna. Se obtiene la variable **ldoublecol**. Se propone utilizar la función **reshape()**.
5. Definir el número de clases $K = 7$ y el de iteraciones $KMI = 7$.
6. Aplicar el algoritmo *k-means clustering*. El parámetro de entrada es **ldoublecol**. Se propone utilizar la función **kmeans()** cuyo parámetro de salida en el vector **lo**.
7. Reformar el vector **lo** a imagen de tamaño **M x N**. Se obtiene la variable de regiones **T**. Se propone utilizar la función **reshape()**.
8. Definir un umbral de área mínima y uno de área máxima: **Uareamin** y **Uareamax**, a partir de la variable **Calibración**.
9. Identificar áreas a contar y que estén entre los umbrales **Uareamin** y **Uareamax**. Se obtienen los identificadores de las áreas **ldr**, **ldg** e **ldb**. Se propone utilizar la función **find()**.
10. Contar áreas identificadas o número de células. Se obtiene la variable resultado.
11. Fin.

Método 3: *Matching Template* utilizando el filtro local de énfasis de puntos oscuros, LDPE.

Este método fue explicado de forma amplia en la sección anterior. Se recuerda que el método realiza el conteo a partir de una plantilla de conteo, construida mediante el filtro LDPE, y en conjunto con los resultados de las segmentaciones watershed, controlada por marcadores, y el k-means clustering.

5. RESULTADOS Y DISCUSIÓN

Se realizaron experimentos para medir los conteos de GB realizados a un conjunto básico de muestras de pacientes caninos, y que se han logrado recolectar hasta el momento de presentación de este documento. Para cada uno de los tres métodos descritos en las secciones anteriores, se realiza la presentación y discusión de resultados.

Para el cálculo del error porcentual se hace uso de la siguiente ecuación:

$$Error\ porcentual = \left| \frac{Resultado\ conteo\ manual - resultado\ metodo\ propuesto}{Resultado\ conteo\ manual} \right| \times 100\% \quad (29)$$

5.1. MÉTODO 1

Este método presenta un algoritmo sencillo, el cual se describe en la metodología. El margen de error para este método, depende de las condiciones iniciales de la imagen, es decir, su resolución, el ruido aditivo y la iluminación no uniforme de la cámara de Neubauer durante la captura. El tiempo de ejecución del algoritmo es muy bajo dado que no necesita realizar operaciones complejas. Este es máximo de 1 minuto, en una prueba se realizada en un computador portátil con procesador Intel® Core™ i5 de 2,4 GHz y 4 GB de RAM. Se probó en 8 imágenes de muestras diferentes y los resultados se describen en las **Tablas 4, 5, 6 y 7**.

| MÉTODO | MUESTRA | | | | | | | |
|-------------------|---------|------|-----|-----|-----|-----|-----|------|
| | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 |
| Conteo Manual | 79 | 65 | 80 | 61 | 53 | 52 | 53 | 53 |
| Conteo canal Rojo | 78 | 51 | 74 | 63 | 55 | 52 | 55 | 60 |
| Error % | 1,3 | 21,5 | 7,5 | 3,3 | 3,8 | 0,0 | 3,8 | 13,2 |

Tabla 4. Resultados método 1 para el canal Rojo.

| MÉTODO | MUESTRA | | | | | | | |
|--------------------|---------|------|-----|-----|-----|-----|-----|-----|
| | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 |
| Conteo Manual | 79 | 65 | 80 | 61 | 53 | 52 | 53 | 53 |
| Conteo canal Verde | 76 | 40 | 75 | 60 | 55 | 54 | 50 | 57 |
| Error % | 3,8 | 38,5 | 6,3 | 1,6 | 3,8 | 3,8 | 5,7 | 7,5 |

Tabla 5. Resultados método 1 para el canal Verde.

| MÉTODO | MUESTRA | | | | | | | |
|-------------------|---------|------|-----|-----|-----|-----|-----|------|
| | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 |
| Conteo Manual | 79 | 65 | 80 | 61 | 53 | 52 | 53 | 53 |
| Conteo canal Azul | 76 | 53 | 74 | 61 | 58 | 56 | 50 | 60 |
| Error % | 3,8 | 18,5 | 7,5 | 0,0 | 9,4 | 7,7 | 5,7 | 13,2 |

Tabla 6. Resultados método 1 para el canal Azul.

Para mejorar el desempeño de este método, se escogen los resultados de los dos canales que presenten el promedio de error más bajo y se los toma como referencia. Del análisis de las **Tablas 4, 5 y 6** se puede decir que los dos canales que cumplen esta condición son el canal Rojo y el canal Verde.

| MÉTODO | MUESTRA | | | | | | | |
|--------------------|---------|------|------|------|------|------|------|------|
| | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 |
| Conteo Manual | 79 | 65 | 80 | 61 | 53 | 52 | 53 | 53 |
| Conteo Canal Rojo | 78 | 51 | 74 | 63 | 55 | 52 | 55 | 60 |
| Conteo Canal Verde | 76 | 40 | 75 | 60 | 55 | 54 | 50 | 57 |
| Conteo Canal Azul | 76 | 53 | 74 | 61 | 58 | 56 | 50 | 60 |
| Promedio Conteo | 76,7 | 48,0 | 74,3 | 61,3 | 56,0 | 54,0 | 51,7 | 59,0 |
| Promedio Error | 3,0 | 26,2 | 7,1 | 1,6 | 5,7 | 3,8 | 5,0 | 11,3 |

Tabla 7. Resultados método 1.

Este método cumple en la mayoría de los casos con el margen de error que se propuso como objetivo al inicio de esta investigación, salvo en aquellos casos donde la calidad de la imagen es demasiado baja. Para el caso de la muestra **S2**, en la **Figura 17** la imagen presenta un grado de resolución bajo, lo que hace que la segmentación mediante los umbrales fijados sea imprecisa.



Figura 17. Captura sin enfoque adecuado (muestra **S2**).

De acuerdo a esto, se puede decir que este método cumple con los objetivos de error propuestos, pero es demasiado sensible a cambios y factores de la imagen de entrada.

5.2. MÉTODO 2

Este algoritmo trabaja directamente con la imagen original, la segmentación se hace mediante el algoritmo de agrupamiento K-means clustering, seguido de una etapa de umbralización. Las pruebas de este método se hicieron sobre 8 muestras distintas. El tiempo de ejecución de este método oscila entre 4 y 5 minutos dependiendo de la resolución de la imagen, que lo que hace es afectar el contraste de esta. De acuerdo a las pruebas realizadas se puede decir que a mayor resolución (mayor contraste) el algoritmo tarda menos tiempo en su ejecución. Esta prueba se condujo en un computador portátil con procesador Intel® Core™ i5 de 2,4 GHz y 4 GB de RAM. Los resultados de conteo proporcionados por este método se muestran en la **Tabla 8**.

| MÉTODO | MUESTRA | | | | | | | |
|----------------|---------|------|-----|-----|-----|-----|-----|-----|
| | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 |
| Conteo Manual | 79 | 65 | 80 | 61 | 53 | 52 | 53 | 53 |
| Conteo K-means | 76 | 53 | 75 | 62 | 54 | 53 | 51 | 56 |
| Error % | 3,8 | 18,5 | 6,3 | 1,6 | 1,9 | 1,9 | 3,8 | 5,7 |

Tabla 8. Resultados Método 2.

Dentro de este método no tiene sentido separar los canales (RGB), ya que después de la etapa de agrupamiento, los valores de intensidad para cada pixel son iguales para los tres canales.

Al analizar los resultados de conteo, conjuntamente con la imagen de la muestra respectiva se puede establecer que este método, aunque relativamente lento, presenta un alto grado de confiabilidad ya que los umbrales se definen de forma automática a partir de la etapa de agrupamiento. Al observar la **Figura 18** se puede notar que la exactitud en el conteo se ve afectada cuando las imágenes presentan una cantidad considerable de artefactos de ruido (burbujas, rayones, manchas y degradación en la cámara de Neubauer), debido a que este algoritmo puede contar tales artefactos como regiones de interés (células).



Figura 18. Captura de baja resolución y gran cantidad de artefactos de ruido (muestra **S2**).

De acuerdo a lo anterior se recomienda tener especial cuidado durante la escogencia de la cámara de Neubauer, la dilución de la muestra y la captura de las imágenes. Se recomienda que las capturas tengan una calidad óptima parecida a la captura de la muestra en la **Figura 19**.

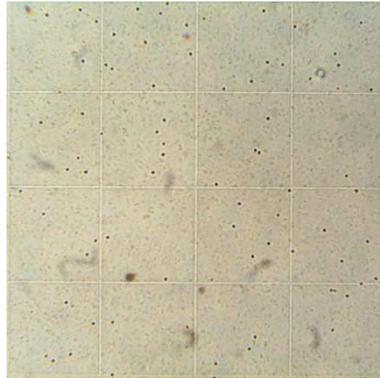


Figura 19. Ejemplo de captura óptima (muestra **S1**).

También se puede decir que el factor de tiempo no tiene mayor relevancia si se tiene en cuenta que el objetivo principal es obtener un conteo exacto.

5.3. MÉTODO 3

Este método hace uso del algoritmo más complejo que se implementó. Debido a los pasos que desarrolla su confiabilidad es alta, y el tiempo de ejecución aumenta. Por cada imagen se ha medido un tiempo de ejecución promedio de entre 5 a 6 minutos; esta prueba se condujo en un computador con procesador Intel® Core™ i3-3110M de 2,4 GHz y 8 GB de RAM. El método de conteo se probó en un conjunto de 8 imágenes, y los resultados encontrados se muestran en la **Tabla 9**.

| MÉTODO | MUESTRA | | | | | | | |
|---------------|---------|------|------|------|------|------|------|------|
| | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 |
| Manual | 61 | 57 | 50 | 62 | 47 | 19 | 32 | 25 |
| A | 60,5 | 57,5 | 49,5 | 61,0 | 35,5 | 17,0 | 32,0 | 21,5 |
| B | 61,0 | 57,0 | 50,5 | 60,0 | 38,5 | 16,0 | 32,0 | 20,5 |
| C | 61,0 | 57,0 | 50,5 | 61,5 | 39,0 | 15,0 | 32,0 | 21,5 |
| \bar{c} | 60,9 | 57,3 | 50,1 | 61,1 | 40,0 | 16,5 | 32,0 | 22,1 |
| σ | 0,2 | 0,3 | 0,4 | 0,7 | 4,3 | 1,5 | 0,0 | 1,7 |

Tabla 9. Resultados método 3.

Como se indicó en la sección anterior de conteo, este método ofrece tres opciones para calcular el número de glóbulos blancos o leucocitos presentes en las muestras. Si se considera el conteo manual como el de referencia, la desviación σ de los métodos, con respecto a la media \bar{c} , es inferior a 5 en el mayor de los casos; es decir, menos del 8% de error en el caso de la muestra que tiene el mayor número de GB (**S1**). Como conteo final, el experto del laboratorio podría tomar el promedio de los métodos **A**, **B** y **C**.

De acuerdo a estos resultados, y realizando un análisis detallado de cada una de las imágenes, se ha podido determinar que el error máximo ha ocurrido para aquellas imágenes en las que aparecen células que se traslapan y se cuentan como una sola. Ejemplo de este caso es la muestra **S5 (Ver Figura 20)**, en donde muchas células han quedado traslapadas dentro de la solución.

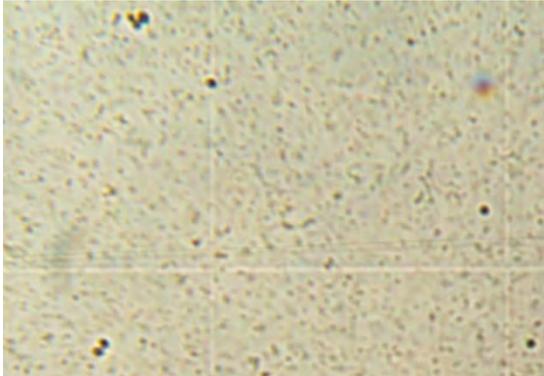


Figura 20. Muestra con células que se traslapan.

El algoritmo del watershed y el k-means clustering pueden lograr cierta robustez al momento de encontrarse con este tipo de muestras. Sin embargo, cuando las células se traslapan mucho siempre serán contadas como una sola.

Las muestras seleccionadas para la prueba del método 3 han sido tomadas con mucho cuidado por el experto del laboratorio, y deberían servir como punto de referencia para el proceso de captura de imágenes y calibración del método. Se resalta el hecho de que la iluminación fue adecuada, el enfoque muy bien logrado y la solución fue bien preparada.

En el caso de imágenes bajo condiciones no adecuadas de iluminación, o con artefactos creados en el proceso de dilución, el error del conteo se puede

incrementar y requiere nuevamente de volver a calibrar la GUI. Idealmente muestras de este tipo no deberían ser ingresadas al algoritmo. En la **Figura 21** se muestra un ejemplo de una imagen de una muestra no apropiada.

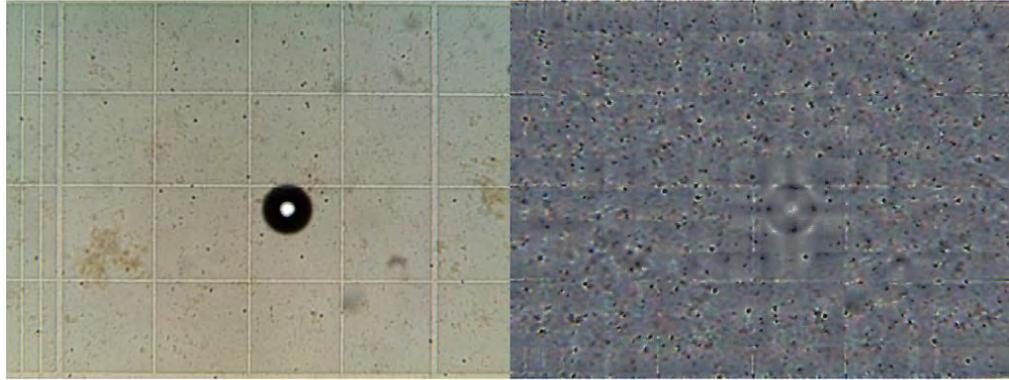


Figura 21. Muestra no adecuada. Izquierda: original con mala iluminación y un artefacto (burbuja). Derecha: resultado de la restauración.

A pesar de que el algoritmo trata de eliminar el artefacto en la imagen original (**Figura 21, izquierda**), su remoción debe hacerse con otro tipo de filtro. Los bordes oscuros que se resaltan en la imagen restaurada (**Figura 21, derecha**) del artefacto pueden no ser contados gracias a una nueva calibración de los umbrales utilizados por los métodos **A, B** y **C**. En esta imagen también se nota la eficiencia del método 3 para resaltar los puntos oscuros y hacer uniforme la mala iluminación de la muestra original.

Para esta muestra en particular, el enfoque tampoco es el apropiado. El resultado ideal debería ser de 45 GB, y los tres métodos **A, B** y **C** computan 35.5, 32 y 41.5 GB respectivamente. Esto sugiere que se debe recalibrar el método 3 o mejor, debería servir al experto del laboratorio para mejorar la captura de la imagen. Este resultado también sugiere que el método de captura de imágenes es muy importante antes de ejecutar el método 3. Se podría entonces sugerir el uso de la GUI para entrenamiento de estudiantes o pasantes

Finalmente, se observa que para este método el tiempo de ejecución no se debe tomar como punto de evaluación, ya que sirve al experto para validar el procedimiento de conteo, cuando él lo requiera o por dudas en el conteo, y tomar nota sobre una nueva calibración del sistema, solo si el experto nota un cambio

drástico en las condiciones del sistema (idealmente debería no cambiar mucho). También se sugiere hacer un nuevo proceso de calibración cuando la base de muestras contenga más datos a analizar.

6. CONCLUSIONES

La propuesta de automatización desarrollada en este trabajo de investigación, representa una forma alternativa de conteo (distinto al visual-manual) de células en muestras de sangre animal para la Clínica veterinaria Carlos Martínez Hoyos de la Universidad de Nariño. Este desarrollo se puede extrapolar a cualquier laboratorio que desee implementar el método propuesto.

Con la implementación del método de recuento automático, basado en visión por computadora y análisis de imágenes, se da pie al replanteamiento de los procedimientos que la guía del laboratorio de la clínica veterinaria exige; por ejemplo, se puede evitar el uso de contadores mecánicos y se reduce en gran medida el estrés visual y el posible error que pueda introducir el encargado del laboratorio. De igual forma, permite mantener organizada la historia clínica del paciente y facilita su acceso inmediato por parte del personal que así lo requiera.

Tres métodos de conteo diferentes se han logrado implementar y ser funcionales a través de una interfaz gráfica de usuario. Su eficacia se ve respaldada por la exactitud lograda en las pruebas realizadas con algunas muestras disponibles en la base de datos de la clínica. Por medio de esta interfaz, la complejidad de los métodos y su programación no se convierten en problema alguno para el encargado del laboratorio; además, no se requiere ningún soporte hardware o software extra, o adquisición de licencias adicionales, para funcionar.

Un beneficio que es importante resaltar es el hecho de que cuando se realiza gran número de exámenes, a veces el personal se ve afectado por el cansancio visual y así mismo la exactitud y precisión del conteo celular. Al automatizar el proceso de conteo con una máquina, garantiza fiabilidad del conteo sin importar el número de veces que se realice.

RECOMENDACIONES

El área de investigación en clasificación y conteo automático de células para un hemograma permanece abierta y su aplicación en los procesos del laboratorio de la clínica veterinaria es de alta relevancia. Por esto se recomienda complementar e incluir métodos apropiados para otro tipo de células y en otro tipo de muestras (líquido cefalorraquídeo, peritoneal u otros), y probando métodos del aprendizaje de máquina supervisada y de aprendizaje profundo.

El objetivo final del proceso de automatización debería ser el de abarcar la mayor cantidad de parámetros medidos y analizados en el hemograma, y en otros procedimientos que el laboratorio ejecuta. Para lograr este objetivo, se recomienda incrementar la base de datos de la clínica a un número lo suficientemente grande de imágenes de muestras de pacientes, y para distintas especies de animales; de tal forma que los algoritmos se puedan validar correctamente.

Por otra parte, se podría sugerir también la inclusión de mejoras adicionales al sistema de automatización; por ejemplo mejorar la adaptación de la cámara al microscopio, para que tanto el enfoque como la toma de las imágenes sean totalmente automáticas, e incluir la automatización del proceso de la dilución de la sangre. De esta forma se tendría un sistema totalmente automático y exacto bajo ciertos parámetros de confiabilidad.

BIBLIOGRAFIA

- [1] F. Fernandez, Manual de interpretación del Hemograma, Editorial medica panamericana, 1991.
- [2] D. P. Alan H. Rebar, Interpretación del Hemograma Canino y Felino, Wilmington, Delaware: Nestlé Purina PetCare Company, 2003.
- [3] A. R. Meder, L. D. Lattanzi. y L. M. Adagio, El hemograma en animales pequeños. Tomo 1: Eritrocitos, ULN PAM, 2012.
- [4] R. E. Alejandro, *Elaboración de los manuales de procedimientos hematológicos serológicos, citológicos, parasitológicos y uroanálisis, como parte integral del proceso de control de calidad interno del laboratorio de la clínica veterinaria "Carlos Martínez Hoyos" de la univ*, Pasto: Universidad de Nariño, 2009.
- [5] N. Otsu, «A threshold selection method from gray level histogram,» *IEEE Transaction on system man cybernetics*, vol. 9, 1979.
- [6] L. W. Hao, W. X. Hong y C. L. Hu, «A novel auto-segmentation scheme for colored leukocyte images,» de *First International Conference on Pervasive Computing, Signal Processing and Applications*,, 2010.
- [7] C. Hagwood, J. Bernal, M. Halter y J. Elliott, «Evaluation of segmentation algorithms on cell populations using cdf curves,» *IEEE Transactions on Medical Imaging*, vol. 31, p. 380–390, 2012.
- [8] D. Shiffman, Learning Processing, Elsevier, 2008.
- [9] C. M. Bishop, Bishop Pattern Recognition and Machine Learning., New York: Springer, 2006.
- [10] R. C. Gonzalez y R. E. Woods, Digital Image Processing, Upper Saddle River, NJ, USA: Prentice-Hall, Inc, 2006.
- [11] M. Schaer, Clinical Medicine of the dog and cat, London: Veterinary Press, 2003.
- [12] F. Rubio, B. García y M. Carrasco, Fundamentos y técnicas de análisis hematológicos y citológicos, Editorial Paraninfo, 2004.
- [13] N. B. Andrade, Manual de patología clinica en animales domesticos, Ibaguè: Universidad del Tolima, 2009.

- [14] J. L. Vives y J. L. Aguilar, *Técnicas de laboratorio en hematología*, elsevier, 2006.
- [15] C. P. Hickman, *Zoología. Principios Integrales*, Madrid: Interamericana McGraw-Hill, 1991.
- [16] C. Urroz, *Elementos de anatomía y fisiología animal*, San jose: UNED, 1991.
- [17] C. A. Villee, *Biología*, Mexico: McGraw-Hill, 2003.
- [18] O. Ryabchykov, A. Ramoji, T. Bocklitz, M. Foerster, S. Hagel, C. Kroegel, M. Bauer, U. Neugebauer y J. Popp, «Leukocyte subtypes classification by means of image processing,» de *Federated Conference on Computer Science and Information Systems (FedCSIS)*, 2016.
- [19] O. Bastidas, «Cell counting with neubauer chamber,» [En línea]. Available: <http://www.celeromics.com/en/resources/docs/Articles/Cell-counting-Neubauer-chamber.pdf>.
- [20] A. Gilat, *Matlab: una introducción con ejemplos prácticos*, Reverte, 2006.
- [21] M. Sajjad, S. Khan, Z. Jan, K. Muhammad, H. Moon, J. T. Kwak, S. Rho, S. W. Baik y I. Mehmood, «Leukocytes classification and segmentation in microscopic blood smear: A resource-aware healthcare service in smart cities,» *IEEE Access*, vol. 5, p. 3475–3489, 2017.
- [22] M. Hamghalam, M. Motameni y A. E. Kelishomi, «Leukocyte segmentatio in giemsa-stained image of peripheral blood smears based on active contour,» de *International Conference on Signal Processing Systems*, 103–106, 2009.
- [23] T. Rosyadi, A. A. Nopriadi, B. Achmad y Faridah, «Classification of leukocyte images using k-means clustering based on geometry features,» de *6th International Annual Engineering Seminar (InAES)*, 2016.
- [24] W. Gao, Y. Tang y X. Li, «Segmentation of microscopic images for counting leukocytes,» de *International Conference on Bioinformatics and Biomedical Engineering*, 2609–2612, 2008.
- [25] Amscope, «Amscope MU300,» Amscope 2015, 2017. [En línea]. Available: <http://www.amscope.com/3mp-usb2-0-microscope-digital-camera-software..> [Último acceso: 2017].
- [26] Mathworks, «deep learning,» Mathworks, 20 enero 2017. [En línea]. Available: <https://www.mathworks.com/discovery/deep-learning.html>. [Último acceso: 2017].

ANEXOS

ANEXO 1. LISTA DE ACRONIMOS

- **CHCM** Hemograma Corpuscular Medio
- **CSC** Conteo Sanguíneo Completo
- **CVCMH** Clínica Veterinaria Carlos Martínez Hoyos
- **D** Filtro de Convolución
- **FFT** Transformada Rápida de Fourier
- **GB** Glóbulo Blanco
- **GPU** Unidad Procesadora Grafica
- **GR** Glóbulo Rojo
- **GUI** Interfaz Gráfica de Usuario
- **Hb** Hemoglobina
- **HC** Hemograma Completo
- **HFE** Énfasis de Alta Frecuencia
- **HSI** Matiz, Saturación y luminosidad
- **Hto** Hematocrito
- **INP** Filtro de Muesca Ideal
- **LCR** Líquido Cefalorraquídeo
- **LM** Matching Local
- **LON** Filtro de Muesca Optimo Local
- **PAI** Procesamiento y Análisis de imágenes
- **PDF** Función de Densidad de Probabilidad
- **ROI** Regiones de Interés
- **SVM** Máquina de Vectores de Soporte
- **VCM** Volumen Corpuscular Medio
- **VSG** Velocidad de Segmentación Globular

ANEXO 2. PONENCIA INTERNACIONAL ROPEC2017

Se hizo la presentación de los principales resultados de la investigación en el “**2017 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC 2017)**”, que se realizó en la ciudad de Ixtapa (México), los días 8 al 11 de Noviembre de 2017. El evento fue organizado por Sección IEEE Centro Occidente con asistencia técnica de la Universidad Michoacana de San Nicolás de Hidalgo, el Instituto Tecnológico de Morelia, la Universidad de Colima, el Instituto Tecnológico de la Costa Grande y la División Centro Occidente de la Comisión Federal de Electricidad. Su importancia se respalda por la continuidad en las versiones del evento (este fue el número XIX) y en el respaldo de *The Library of Congress*, que concedió el ISSN ID (2573-0770) para la publicación del proceedings del evento. También será publicado en *IEEE Xplore Digital Library* y *Scopus*..

2017 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC 2017), Ixtapa, Mexico

Automatic Count of Leukocytes in Images of Animal Blood Samples

John Muñoz ¹, Ilsemy Patricia ², Mabel Ortega ³, Kaito Benavides ⁴, and Wilson D. Achicanoy M. ⁵

¹Departamento de Electrónica, Clínica Veterinaria Carlos Martínez Hoyos

Universidad de Nariño, Pasto, 320002, Colombia

Email: {john.jmuoz172, ilsemyan14}@hotmail.com, kaitobenavides@gmail.com, wdaclac@uninar.edu.co

²Programa de Pós-Graduação em Engenharia Elétrica

Pontifícia Universidade Católica do Rio de Janeiro, Rua Marquês de São Vicente, 225, Glória, Rio de Janeiro, Brasil
Email: mortega@ele.puc-rio.br

Abstract—We propose an image processing methodology to automate the leukocyte count in Neubauer chambers. The methodology works in three stages: restoration, segmentation, and automatic count. Restoration is performed by taking local properties of neighborhood into account. To improve segmentation, and since leukocyte concentration is low compared to other cell populations, we propose a filter of emphasis of darked points, which is based on maximization of dissimilarity between an estimated free-leukocyte image and the respective image with leukocytes. To make the automatic count, a template based on the proposed filter is computed and used together with marker-controlled watershed segmentation and k-means clustering. A discussion about the results is presented.

1. INTRODUCTION

New applications have recently emerged for human health-care services, and online patient data are required to be available for medical specialists. One objective is to translate these services, already present in Smart Cities [1], to animal care in veterinary clinics. The priority in clinical laboratories is the automation of leukocyte counting based on low cost Neubauer chambers. Since manual counting has drawbacks such as waste of time and operator's eye strain, we propose a method for leukocyte counting based on digital image processing and analysis.

Research in this area has been conducted using high resolution leukocyte images, which allow feature extraction by combination of basic and advanced algorithms. For example, in [2], authors propose a method for recognizing leukocyte's region of interest (ROI), which is based on Otsu's method and active contour (snake) that use gradient vector flow. In [3], it is proposed a controlled watershed segmentation in HSV color space. In [4], to extract leukocyte nucleus by an improved watershed technique, color features are obtained from textural gradient, which is computed by a non-decimated complex wavelet transform.

In [5], authors propose a method for distinction between two white blood cell subtypes. This is based on pseudo-Zernike moments (which do not depend on image rotations) together with a support vector machine (SVM). Physical and chemical cell characteristics are retrieved in Lab color space. In [6], authors apply the self-dual multiscale morphological toggle operation, which improves watershed segmentation.

Level-Set methods are used for detection of leukocyte's nucleus, and cytoplasm regions are detected by top-to bottom and area opening operation.

In [7], authors propose classification, based on k-means clustering, for four types of leukocytes. They found that circularity is the geometric feature that gives the most accuracy. Similarly, in [8] and to obtain reliable count, authors use segmentation by k-means clustering and classification by an ensemble-multiclass SVM model.

One of the most important steps in automatic cell count is accuracy of segmentation algorithms. In [9], authors present an assessment of four methods: Otsu thresholding, k-means clustering, watershed, and convex edge detector. They showed that the cumulative distribution function of misclassification errors is enough to choose the best method, yielding that watershed and k-means clustering are the ones with more accuracy.

Unlike the above methods, images of animal blood samples from Neubauer chamber do not show well-defined features of cells, like nucleus or cytoplasm. Then, in this work we propose a methodology for automatic count of leukocytes (only based on canine blood samples) based on a filter of emphasis of darked points (usable leukocytes). The proposed filter is used for segmentation improvement and to derive a template for the automatic count. The filter maximizes the difference between a free leukocyte image (background) and the same image with leukocytes (foreground). It reduces over-segmentation issues commonly presented in watershed segmentation and k-means clustering.

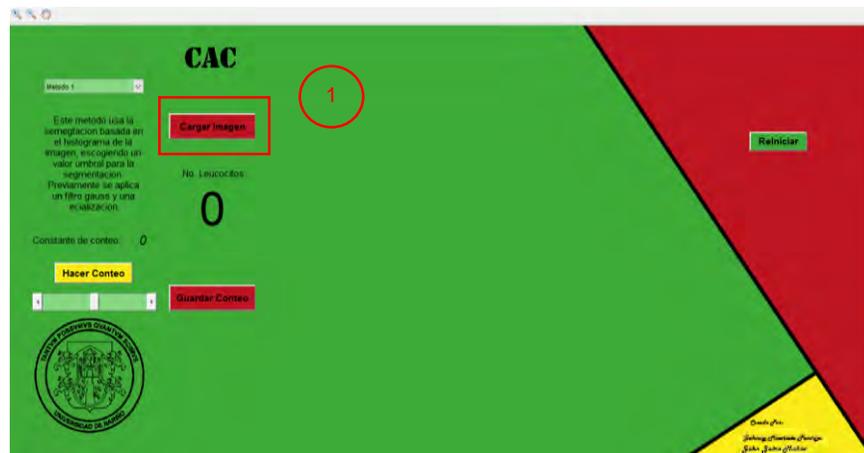
This paper is organized as follows: Section II describes in a detailed level the stages of the methodology: restoration, segmentation, and count. In Section III, discussion of the main outcomes and future work is presented. Finally, concluding remarks are done in Section IV.

II. METHODOLOGY

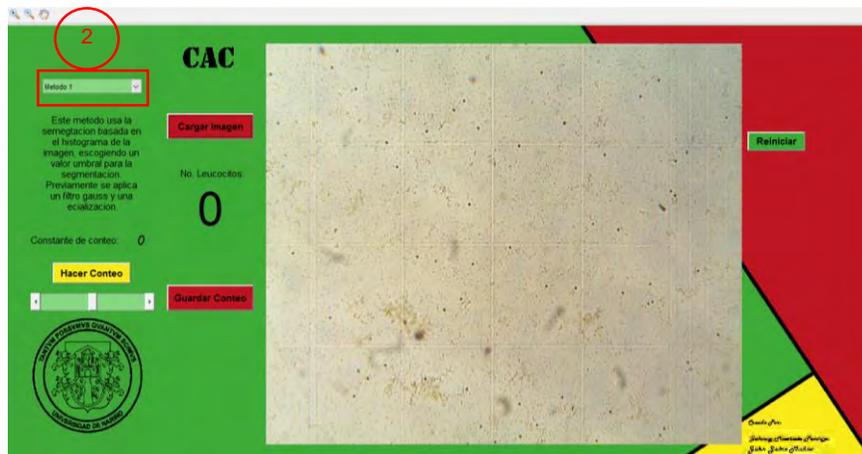
Many laboratories in veterinary clinics still use Neubauer chambers to count blood cells in animal samples. Currently, a diluted sample is put in the chamber and observed with an optical microscope. The leukocyte count, operator focus each corner square (corner square is used for red blood cell count)

ANEXO 3. MANUAL DE USUARIO DE LA APLICACIÓN

1. Al accionar el botón **“cargar Imagen”** se abre una ventana de búsqueda donde se debe buscar la imagen o fotografía donde se muestren los glóbulos blancos (GB). Al cargar la imagen, automáticamente se visualiza en la pantalla.

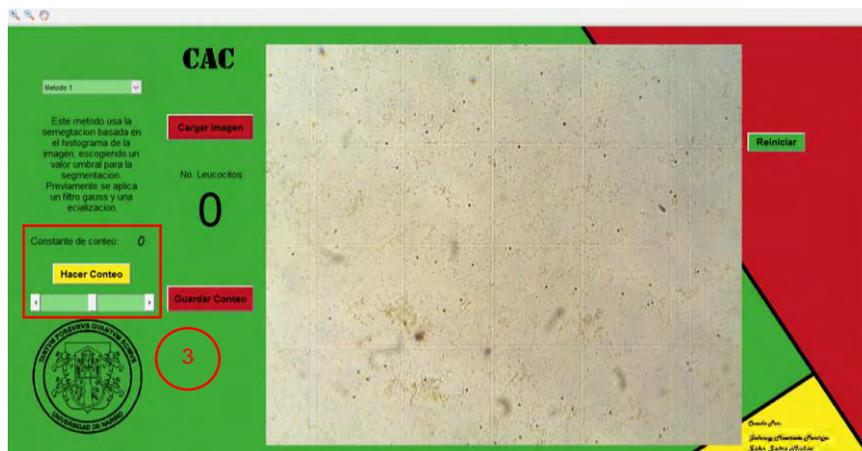


2. Se debe escoger el método de conteo que se prefiera.

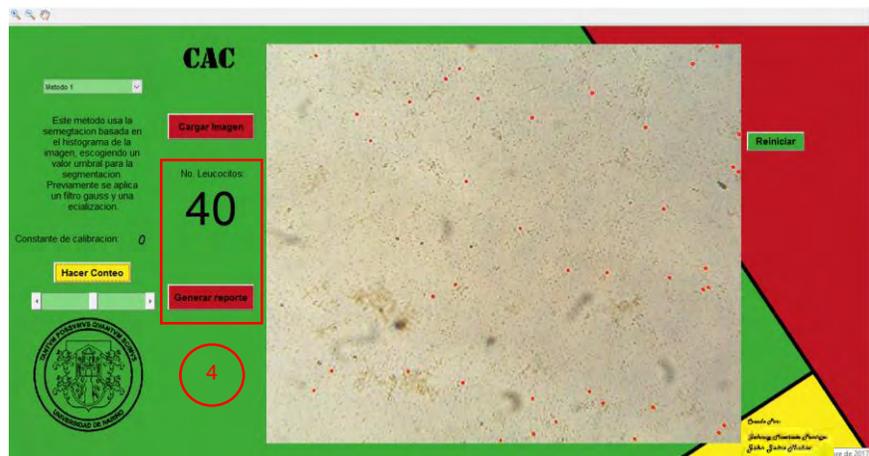


3. Se acciona el botón **“hacer conteo”** para realizar el conteo de las células presentes en la imagen previamente cargada. La barra deslizante que controla la constante de conteo que representa la sensibilidad del método de conteo está configurada para obtener el mejor resultado, pero, existen casos donde las patologías presentes en algunos pacientes animales hacen que las células cambien de tamaño drásticamente, por esta razón a veces es necesario calibrar nuevamente el método para obtener mejores resultados.

Como paso importante del proceso de conteo la interfaz realiza la selección de la región de interés automáticamente ya que no se debe hacer el conteo en toda el área de la imagen si no en una determinada.



4. Cuando el proceso termina y el número de células se muestra en la pantalla se puede accionar el botón **“generar reporte”** el cual guarda en una base de datos.



Se genera un archivo de texto con los siguientes campos: identificador del paciente, fecha, tipo de célula contada, y el número total calculado. Otros campos adicionales son el nombre del encargado de realizar el conteo y el médico que trata al paciente.