

Visión Computacional I

Conteo y Seguimiento de desplazamiento bacteriano.

Edgar Iván Aguilera Hernández
Maestría en Ciencias de la Computación, CIMAT, Gto.

11 de junio de 2025

La caracterización del comportamiento bacteriano, a través de su desplazamiento en el medio, permite entender las interacciones entre estas y otros elementos de su entorno, igualmente para algunas de sus especies, las tendencias en su movimiento logran explicar el origen y evolución de las infecciones. Dentro de la variedad de métodos para su seguimiento, la principal ventaja de la técnica basada en visión computacional es la de evitar cualquier tipo de etiquetado directo, como los métodos de fluorescencia, en las que no siempre es posible aplicar una tinción compatible sobre la membrana de la especie de interés, o en otros casos llegando a modificar el comportamiento de la bacteria.

1. Marco Teórico

1.1. Extracción de fondo

Dado que el objetivo del seguimiento bacteriano se basa en los elementos móviles de la escena registrada por el microscopio, una de las técnicas ajustadas para este tipo de problemas es la segmentación y extracción del fondo en las imágenes obtenidas. Para lograr una velocidad de procesamiento $O(n)$ al mismo tiempo tomar ventaja de las características conocidas en las imágenes adquiridas por el microscopio, es decir, un fondo estable y cambios mínimos o incluso nulos de iluminación, se optó por un método de mezcla de gaussianas (Mixture of Gaussians: MOG) con la que se ha reportado un mejor desempeño bajo este tipo de supuestos [4], específicamente la implementación propuesta por Zivkovic y Heijden [3], como una de las versiones mejoradas de su propia propuesta.

En este tipo de métodos se utiliza la modelación de la intensidad de cada píxel, como una combinación variable (M) de gaussianas, en donde el peso asignado a cada una de estas distribuciones será proporcional al tiempo que dicha intensidad permanece sobre el píxel, de esta forma, cuando el peso de la distribución es bajo, entonces será considerado como parte de los elementos móviles, por la relación bayesiana entre las probabilidades de estas categorías, la clasificación se puede expresar como (1):

$$\frac{p(BG|x^t)}{p(FG|x^t)} = \frac{p(BG|x^t)p(BG)}{p(FG|x^t)p(FG)}$$

donde :

x^t : valor del píxel en el tiempo t

BG : densidad del fondo (background)

FG : densidad del plano móvil (foreground)

(1)

En la determinación de estos pesos, se utiliza como base una muestra de los valores del píxel para un periodo de tiempo T , $X_T = \{x^t, \dots, x^{t-T}\}$. Como no se tiene conocimiento previo de la pertenencia de estos píxeles a una de las 2 categorías, entonces las distribuciones de densidad son re-expresadas como una combinación de BG y FG, caracterizadas por las estimaciones de los parámetros que las componen (2).

$$\hat{p}(x|X_T, BG + FG) = \sum_{m=1}^M \hat{\pi}_m N(x; \hat{\mu}_m, \hat{\sigma}_m I)$$

donde :

$\hat{\mu}_m$: estimación de la media sobre la m-esima gaussiana (2)

$\hat{\sigma}_m I$: estimación de las varianzas (conservando isotropía sobre las covarianzas)

$\hat{\pi}_m$: factor de contribución (no negativo y $\sum_m \hat{\pi}_m = 1$)

Al recibir una nueva muestra x^t , se evalúa la similitud con el resto de componentes gaussianas M , utilizando la distancia cuadrada de Mahalanobis: $D_m^2(x^t) = \frac{\delta_m^T \delta_m}{\hat{\sigma}_m^2}$, donde $\delta_m = x^t - \hat{\mu}_m$, si la similitud es inferior a un umbral definido, entonces el nuevo valor del píxel será representado por las distribuciones construidas hasta ese punto, específicamente estableciendo el componente con la menor distancia, denominado *owner* (o_m^t), en 1, mientras para el resto de las gaussianas se establecen en 0, actualizando en el proceso los estimadores a través de las ecuaciones (3). Cuando la distancia resultante es superior al umbral de tolerancia, la nueva muestra será modelada por un nuevo componente en donde los estimadores son inicializados como $\hat{\mu}_m = x^t, \hat{\sigma}_m = \sigma_0, \hat{\pi}_m = \alpha$, en donde σ_0 será un valor de inicialización elegido previamente, y $\alpha = \frac{1}{T}$ como una re-expresión para el inverso del periodo.

$$\begin{aligned} \hat{\pi}_m b &= \hat{\pi}_m + \alpha(o_m^t - \hat{\pi}_m) \\ \hat{\mu}_m &= \hat{\mu}_m + o_m^t \left(\frac{\alpha}{\hat{\pi}_m} \right) \delta_m \\ \hat{\sigma}_m^2 &= \hat{\sigma}_m^2 + o_m^t \left(\frac{\alpha}{\hat{\pi}_m} \right) (\delta_m^T \delta_m - \hat{\sigma}_m^2) \end{aligned} \quad (3)$$

1.2. Operaciones morfológicas

Con la segmentación de los píxeles sobre cada frame, es posible la presencia de ruido sobre la imagen binarizada, por resultado de falsos positivos o falsos negativos, estas inconsistencia pueden aparecer en forma de píxeles marcados como parte del *foreground* y aislados de los verdaderos elementos móviles de la escena, o en otro caso como píxeles marcados como *background* dentro de los elementos móviles, debido al efecto de este último, es común observar sobre un mismo objeto lo que aparentan ser múltiples elementos independientes (inconexos), distando de la verdadera cantidad de objetos móviles en la escena. Para mitigar este tipo de fenómenos, se recurre a una fase de filtrado especializada en escenas binarizadas, denominadas operaciones morfológicas.

Este tipo de operaciones realiza comparaciones de un píxel, con los de una vecindad delimitada por una matriz (kernel), denominada en esta método como elemento estructurante, esto debido a la variedad de tamaño y formas que pueden ser definidos de acuerdo al tipo de respuesta esperado (fig. 1).

Al utilizar estas estructuras bajo dimensiones que generalmente no exceden las decenas, en combinación con una de las operaciones comparativas, dilatación y erosión, el resultado será una modificación de las fronteras entre los valores 0 y 1 [2]:

- Erosión: de forma generalizada, esta operación asigna el mínimo valor al píxel sobre el cual se encuentra operando (centrado) el elemento estructurante, como en este caso es una imagen binaria, se asigna 0 si alguno de los píxeles vecinos es 0, resultando en la reducción o eliminación, según su dimensión, de los falsos positivos y .“delgazando” las figuras de interés.
- Dilatación: con el efecto opuesto, el píxel bajo operación será marcado como 1 si al menos uno de los vecinos tiene el valor 1. Lo que visualmente resulta en la .“expansión” de las elementos en escena.

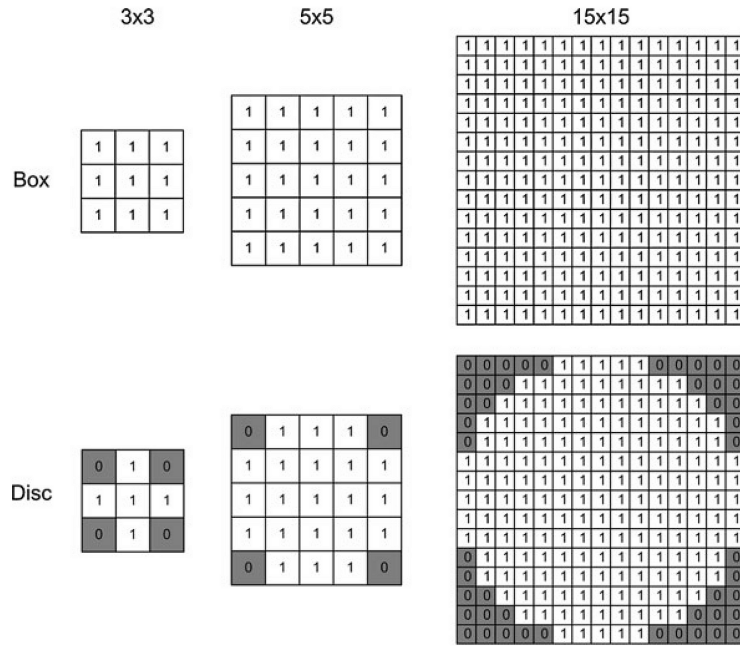


Figura 1: Ejemplo de elementos estructurantes.

Con una secuencia de estos operandos, específicamente con el objetivo de rellenar”los huecos resultantes de los falsos negativos en la extracción del fondo, se consigue la operación de cerradura”, en la que al combinar la dilatación para rellenar los huecos y posteriormente la erosión para conservar el contorno original, es posible reducir a cantidad de elementos inconexos y obtener una representación más precisas del área en el objeto de interés (fig 2).

Binary Morphology

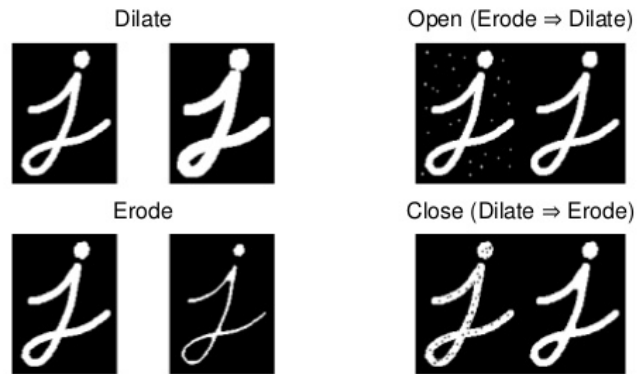


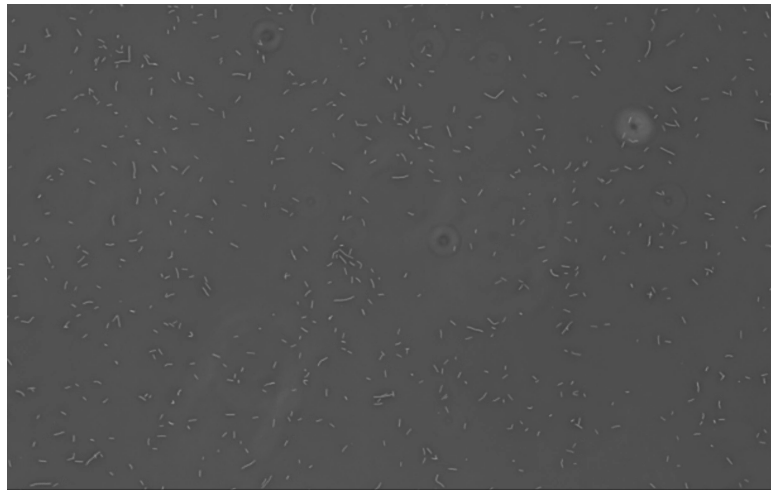
Figura 2: Operaciones morfológicas básicas y su combinación para obtener una imagen filtrada.

2. Metodología

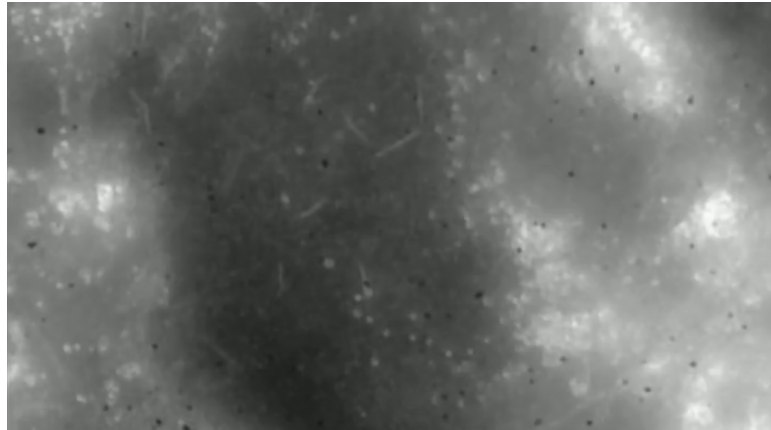
El par de videos utilizados en el presente corresponden a grabaciones presentadas en trabajos similares de conteo y seguimiento de bacterias (fig. 3):

- *Bacillus subtilis* [5]: video con 2:46 min. de movilidad bacteriana sobre un medio de cultivo no especificado, observadas desde un microscopio con objetivo 10x, se utilizó una cámara Aptina CMOS Sensor 18MP 1/2.3 Color, a frecuencia de 30 fps.
- *Leptospira interrogans* [6]: video de 23 seg de movilidad bacteriana sobre un medio de células renales cultivadas para el experimento, se observaron a través de un microscopio de campo oscuro , con una cámara WAT- 910HX CCD 1/2, a una frecuencia de 30 fps.

Todas las operaciones y visualizaciones sobre la secuencia de imágenes fueron aplicadas utilizando la librería de OpenCV [1], mientras el algoritmo de ejecución fue realizado en C++ sobre un sub-sistema operativo WSL Ubuntu ver. 22.04.3., con un procesador de 13va generación Intel Core i7-13620H a 2.40 GHz. El código completo puede ser consultados en el repositorio: <https://github.com/EdgarBanz0/BacteriaTracking>



(a) *Bacillus subtilis*.



(b) *Leptospira interrogans*.

Figura 3: Capturas de video para cada ambiente bacteriano.

La adquisición y procesamiento de cada frame puede ser dividida en 5 fases principales,

1. Preprocesamiento: en la que se transforma la imagen a color a escala de grises y se aplica un suavizado gaussiano con el objetivo de reducir el ruido que resulta tras la aplicación de los métodos de segmentación y operaciones morfológicas.
2. Extracción de fondo: Utilizando el modelo MOG descrito previamente, se logra obtener una imagen binarizada en el que los elementos de interés serán marcados en la totalidad de su área, considerando que se mantendrán algunas imperfecciones de menor escala dentro y fuera de estos elementos.

3. Caracterización de agrupaciones: Con la idea de usar el centro de masa de cada una de las bacterias en movimiento, se aplica inicialmente una serie de filtrados para eliminar los píxeles erróneamente marcados como bacterias, utilizando la operación de apertura, en seguida se rellenan los huecos dentro de los elementos que verdaderamente pertenecen a los cuerpos bacterianos a través de la operación de cerradura, y tras haber obtenido un área completa que pueda ser asociada al elemento móvil, entonces se calcula el área asociada así como el centro de masa con el que se marcará cada cuerpo bacteriano para ser el píxel representativo de la posición en el frame actual.
4. Seguimiento bacteriano entre frames: utilizando una estructura que almacena las posiciones anteriores para cada centro de masa registrado previamente, una bacteria se considerará la misma entre frames consecutivos, si la distancia euclidiana entre el centro de masa para el frame t se encuentra a un máximo de 10 píxeles con el frame $t+1$, en otro caso, se registrará como una nueva bacteria a seguir en las próximas iteraciones. Además, si un registro no puede ser asociado con una de las bacterias en el nuevo frame, se conservará su posición final hasta por 10 frames, a modo de "tiempo de vida", tras exceder esta cifra, se eliminará su historial de desplazamiento permanentemente.
5. Actualización de estadísticas: Conservando el número de bacterias detectadas en el frame actual, así como el máximo observado hasta el frame t , se actualizará el área promedio de las bacterias identificadas para que sirvan de estimación en los próximos frames si se detectan agrupaciones de bacterias bajo un mismo centro de masa, de esta forma el conteo e historial de desplazamiento reflejarán de forma más precisa el posible solapamiento de bacterias.

Algorithm 1: Pseudocódigo de seguimiento bacteriano

```

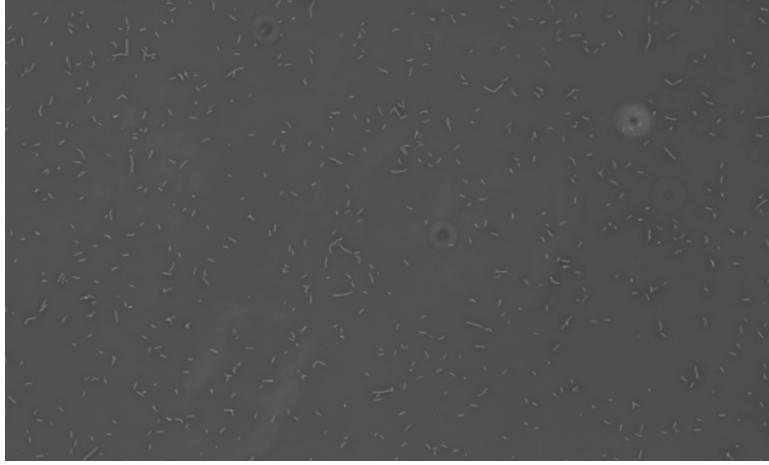
1 Function BacteriaTracking(ImSeq):
   Data: image sequence(video) : ImSeq[width,height,t]
   Result: displacement history : move_mc[bacteria,trajectories]
2   kernel_small[x,y]                ▷ Structuring element (disc) for opening;
3   kernel_large[x,y]                ▷ Structuring element (disc) for closing;
4   new_mc[ ]                          ▷ Detected mass centers on each frame;
5   move_mc[ ]                        ▷ history of positions of mass centers;
6   for frame in ImSeq do
7     clean_frame ← color2gray(frame);
8     clean_frame ← GaussianBlur(clean_frame);
9     bkg_frame ← MOG(clean_frame);
10    clean_frame ← morphologyOpen(bkg_frame, kernel_small);
11    clean_frame ← morphologyClose(clean_frame, kernel_large);
12    new_mc ← massCenter(clean_frame);
13    temp_mc ← move_mc;
14    for i in new_mc do
15      for j in move_mc[:,0] do
16        d ← EuclideanDistance(move_mc[j], new_mc[i]) if distance < 10 then
17          temp_mc[j] ← move_mc[j];
18          ▷ 5. Update avg. mass and element count;
19    statistics(move_mc);
20    res_img ← drawElements(move_mc);
21    show(res_img);

```

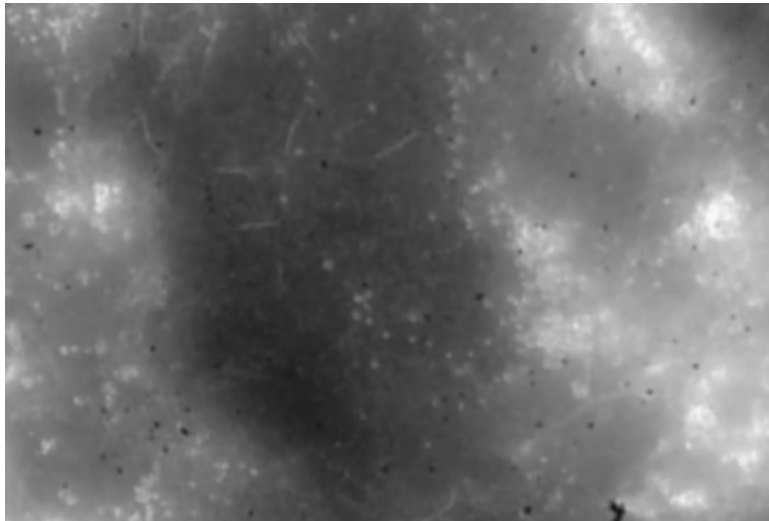
3. Resultados

3.1. Preprocesamiento

Utilizando un kernel de suavizado Gaussiano de tamaño 3×3 sobre cada frame en escala de grises, se consigue atenuar algunos de los elementos del fondo con lo que se podrá efectuar una mejor segmentación del fondo (fig 4).



(a) *Bacillus subtilis*.



(b) *Leptospira interrogans*.

Figura 4: Suavizado Gaussiano sobre cada video.

3.2. Extracción de fondo

Al construir el modelo de segmentación del fondo a partir de la función *createBackgroundSubtractorMOG2* en openCV, esta implementa la versión mejorada del modelo MOG al determinar de forma automática el número adecuado de distribuciones Gaussianas M con las que se representa la intensidad de los píxeles. En la configuración de este método se definió el historial de muestras (X_T) con una longitud de 200 frames, mientras la distancia umbral entre los valores del nuevo píxel y las distribuciones previas fue de 12.0. Con estos parámetros se logró el equilibrio entre la adaptación del modelo a los cambios en el fondo y el ruido restante, especialmente sobre el video *Leptospira interrogans*, en el que las bacterias presentan un desplazamiento no solo sobre el espacio visible, sino también en la profundidad del medio, lo que para un valor de X_T y distancia umbral más altos significarían cambios menos perceptibles por el modelo (fig 5).



(a) Sub-región de *Bacillus subtilis*.

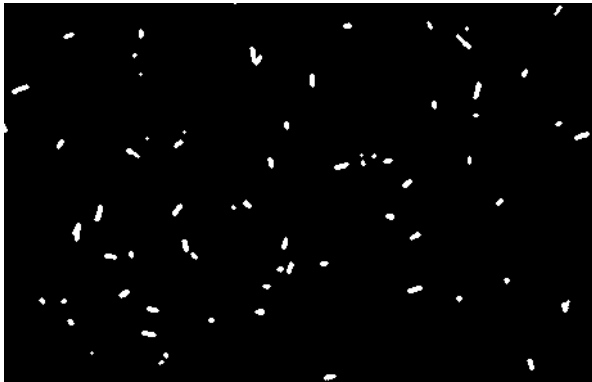


(b) *Leptospira interrogans*.

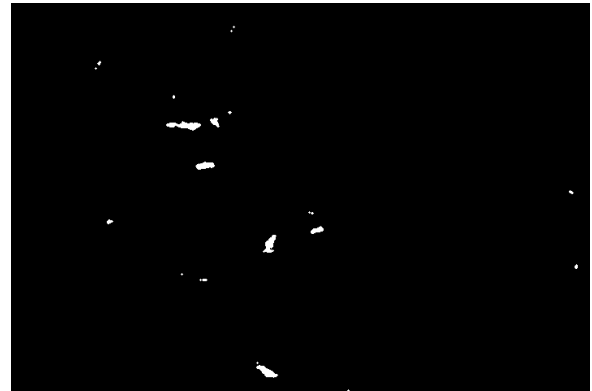
Figura 5: Escenas binarizadas tras la segmentación de fondo.

3.3. Caracterización de agrupaciones

En la eliminación del ruido de la fase anterior, se definieron 2 elementos estructurantes en forma de disco: durante la operación de apertura se utilizó un kernel de 3×3 , y para la apertura se expandió a 11×11 , la razón de esta diferencia radica en la necesidad de eliminar solo puntos pequeños fuera de lo que verdaderamente son masas de bacterias (fig 6), mientras que al rellenar huecos dentro de estas masas, también se intenta unir o en el peor de los casos agrupar elementos de mayor tamaño (fig 7), este último efecto puede ser contrarrestado con la división del área de las masas resultantes entre el área promedio determinada hasta ese momento.



(a) Sub-región de *Bacillus subtilis*.

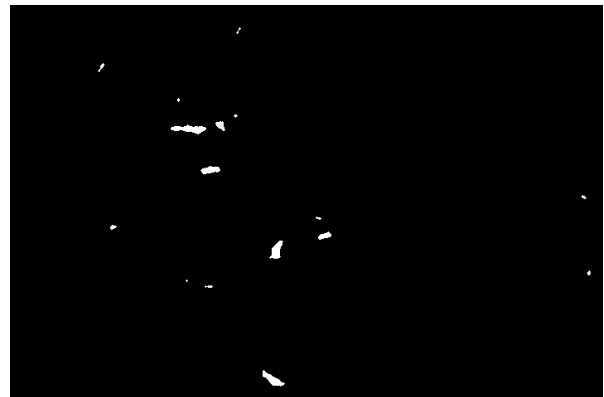


(b) *Leptospira interrogans*.

Figura 6: Operación de apertura para la eliminación de ruido fuera de los elementos principales.



(a) Sub-región de *Bacillus subtilis*.



(b) *Leptospira interrogans*.

Figura 7: Operación de cerradura para la eliminación de ruido dentro de los elementos principales.

En específico sobre la estimación del centro de masa (punto de equilibrio del área en un elemento de la imagen) (fig 8), openCV realiza una implementación de los momentos de una imagen (función *moments*), en donde al igual que en física, un elemento en la imagen delimitado por su contorno, se considera que tiene un peso dado por la intensidad de los píxeles que contiene, así los momentos espaciales para los grados i, j se definen como:

$$m_{i,j} = \sum_{x,y} I(x,y) x^j y^i \quad (4)$$

En la literatura estos momentos y sus distintos grados son utilizados en la caracterización de imágenes, sin embargo para nuestra implementación solo es necesario hasta el primer orden, pues el centro de masa se obtiene:

$$\hat{x} = \frac{m_{1,0}}{m_{0,0}}, \hat{y} = \frac{m_{0,1}}{m_{0,0}} \quad (5)$$



(a) Sub-región de *Bacillus subtilis*.

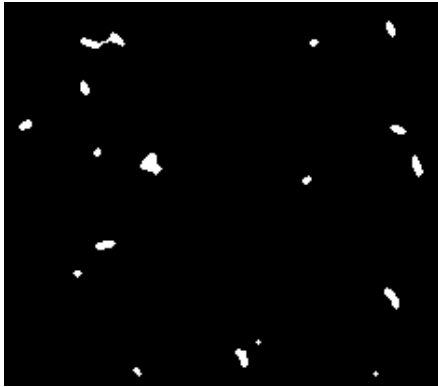


(b) *Leptospira interrogans*.

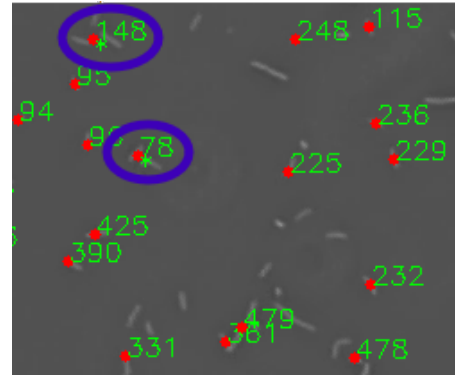
Figura 8: Centro de masa(punto oscuro) obtenido sobre elementos con área definida.

3.4. Seguimiento bacteriano entre frames

Con el posible solapamiento de las bacterias por efecto de su propio movimiento, o por la agrupación realizada durante la operación de cerradura, una forma de lidiar con este problema durante el registro de las trayectorias, es utilizando el área promedio de las bacterias adquirida de los frames anteriores, de esta forma se logra asociar un elemento n_i con tantas bacterias como el área promedio divida las áreas del frame actual $n_i = area_i / (1.5 * area_{avg})$, dado que por la propia morfología y punto de vista en el microscopio estas dimensiones puedes variar fácilmente de un momento a otro, el número n se restringe a un mínimo de 1 y el área promediada es multiplicada 1.5 veces para subestimar las bacterias dentro de una agrupación (fig 9 y 10).



(a) áreas segmentadas tras cerradura

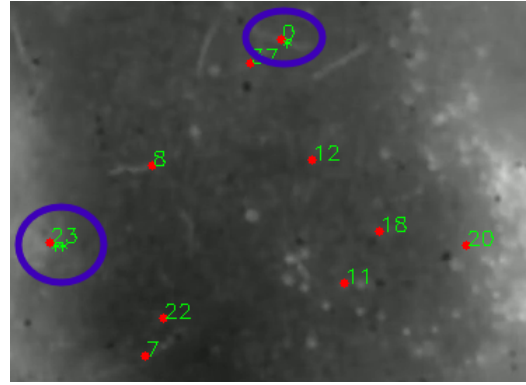


(b) contabilización múltiple (número de
"*.adicionales)

Figura 9: Agrupaciones bacterianas y su contabilización múltiple cuando las áreas son notablemente mayores al promedio (subregión de video *Bacillus subtilis*).



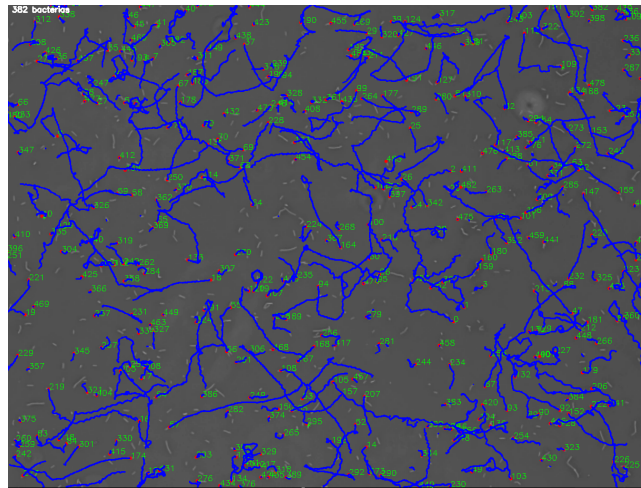
(a) áreas segmentadas tras cerradura



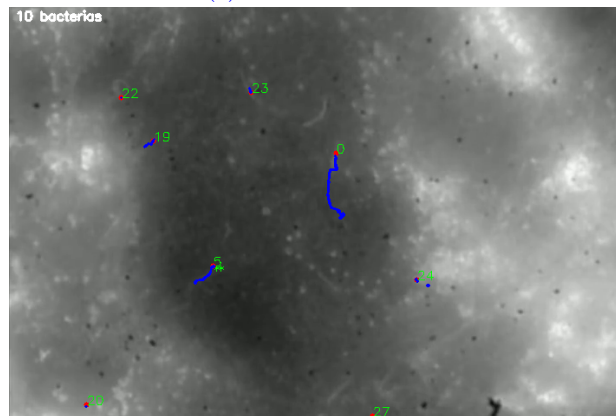
(b) contabilización múltiple (número de ".*.adicionales)

Figura 10: Agrupaciones bacterianas y su contabilización múltiple cuando las áreas son notablemente mayores al promedio (video *Leptospira interrogans*).

Durante la ejecución, el historial de desplazamientos será reiniciado tras haber una colisión, calculando las velocidades a partir de los puntos a descartar y añadiendo los nuevos cálculos al promedio de velocidad global. En la siguiente figura (fig 10) se muestran las trayectorias obtenidas para cada bacteria detectada, por el movimiento en la profundidad del medio que se observa en el video de *Leptospira interrogans*, las trayectorias son menos prolongadas.



(a) *Bacillus subtilis*.



(b) *Leptospira interrogans*.

Figura 11: Agrupaciones bacterianas y su contabilización múltiple (número de ".*.adicionales).

Métrica	Bacillus subtilis.	Leptospira interrogans
Frames	4985	1337
Detecciones Max.	418	24
Detecciones Promedio	348	12
Velocidad Promedio	3.02	5.23

Tabla 1: Resultados de detección y seguimiento bacteriano por video.

4. Discusión

De los 2 videos recopilados para las demostraciones realizadas previamente, unicamente se tiene conocimiento de los resultados cuantitativos reportados por la literatura encargada de efectuar el seguimiento y estimación de las velocidades bacterianas en el video *Leptospira interrogans* [6], sin embargo, dado que este forma parte de una serie de grabaciones efectuadas con diferentes grados de infección en las que se observa ciertas diferencias en las velocidades promedio alcanzadas por las bacterias, el video aquí presentado no pudo ser explícitamente asociado con una de estas pruebas.

Para obtener una comparación de los resultados bajo condiciones similares, se tomó las dimensiones del píxel reportadas por [6] para obtener una velocidad en $\mu m/s$, es este caso: $x = 0.170\mu m/pixel$, $y = 0.187\mu m/pixel$, como en la presente implementación no se diferencia el desplazamiento vertical u horizontal, las dimensiones fueron ajustadas bajo la suposición de un píxel cuadrado $x = y = 0.170\mu m/pixel$. Es así que al tomar el rango de los experimentos en [6], se reportó que las velocidades promedio se encuentran entre $5 \mu m/s - 7 \mu m/s$, mientras que en el presente trabajo se obtuvo de $5.23 \mu m/s$.

5. Conclusiones

El seguimiento de motilidad bacteriana a través de secuencias de video permite caracterizar el comportamiento de las bacterias sin la necesidad de recurrir a métodos de rastreo directos como lo es la tinción de sus membranas, aunque esta técnica de procesamiento de imágenes presenta sus desventajas principalmente en la precisión del algoritmo para registrar la totalidad del desplazamiento individual sobre cada bacteria, en realidad su capacidad de conteo y seguimiento mejora especialmente sobre escenas donde el desplazamiento bacteriana se realiza principalmente sobre el plano bidimensional, como se observó en el video *Bacillus subtilis*. Como punto a mejorar en futuras implementaciones, se podría obtener mayor ventaja del área promedio determinada sobre cada iteración, para ajustar de forma dinámica las dimensiones de los elementos estructurantes, y realizar durante la misma ejecución una separación más precisa de los cuerpos bacterianos.

Referencias

- [1] G. Bradski, "The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, 2000.
- [2] R. Fisher, S. Perkins, A. Walker y E. Wolfart, *HYPERMEDIA IMAGE PROCESSING REFERENCE*, 2000. dirección: <https://homepages.inf.ed.ac.uk/rbf/HIPR2/morops.htm>.
- [3] Z. Zivkovic y F. van der Heijden, "Efficient adaptive density estimation per image pixel for the task of background subtraction," *Pattern Recognition Letters*, vol. 27, n.º 7, págs. 773-780, 2006, ISSN: 0167-8655. DOI: 10.1016/j.patrec.2005.11.005.
- [4] L. A. Marcomini y A. L. Cunha, *A Comparison between Background Modelling Methods for Vehicle Segmentation in Highway Traffic Videos*, 2018. arXiv: 1810.02835 [cs.CV]. dirección: <https://arxiv.org/abs/1810.02835>.
- [5] S. J., O. I., D. J y et al., "YSMR: a video tracking and analysis program for bacterial motility.," *BMC Bioinformatics* 21, 166 (2020), 2020. DOI: 10.1186/s12859-020-3495-9.
- [6] A. K., K. N. y N. Sl, "Machine learning-based motion tracking reveals an inverse correlation between adhesivity and surface motility of the leptospirosis spirochete.," *Nat Commun* 14, 7703 (2023), 2023. DOI: 10.1038/s41467-023-43366-0.