



Institución Universitaria

# **Metodología para la caracterización de la apitoxina desde imágenes de electroforesis bidimensional en gel usando descriptores espaciales**

**Gerardo Antonio Villegas Rivera**

Instituto Tecnológico Metropolitano

Facultad de Ingenierías

Medellín, Colombia

2018

# **Metodología para la caracterización de la apitoxina desde imágenes de electroforesis bidimensional en gel usando descriptores espaciales**

**Gerardo Antonio Villegas Rivera**

Tesis o trabajo de investigación presentada(o) como requisito parcial para optar al título de:

**Magister en Automatización y Control Industrial**

Director (a):

María Constanza Torres Madroño, MSc, PhD<sup>1</sup>

Co-directores:

Edilson Delgado Trejos, MSc, PhD<sup>2</sup>

Sarah Röthlisberger Booth, MSc, PhD<sup>3</sup>

Línea de Investigación:

Máquinas Inteligentes y Reconocimiento de Patrones

Grupos de Investigación:

<sup>1</sup>Grupo de Investigación Automática, Electrónica y Ciencias Computacionales (AEyCC).

<sup>2</sup>Grupo de Investigación en Calidad, Metrología y Producción (CMyP).

<sup>3</sup>Grupo de Investigación e Innovación Biomédica (GI2B).

Instituto Tecnológico Metropolitano

Facultad de Ingenierías

Medellín, Colombia

2018

*A mi esposa Irene,  
mis hijos Daniel y Katherine,  
por su amor, fe, confianza,  
paciencia y comprensión.*

## **Agradecimientos.**

Expreso mis más profundos agradecimientos:

A María Constanza Torres Madroñero, instructora y asesora principal del proyecto, por su paciencia y orientación desinteresada.

A los Co-investigadores y asesores Edilson Delgado Trejos y Sarah Röthlisberger Booth, por haber tenido fe en el proyecto y por haber permitido la consolidación del mismo.

Al personal directivo y de comunicaciones del Instituto Tecnológico Metropolitano, por su acompañamiento,

A todos los profesores que se relacionaron conmigo en el proceso de aprendizaje durante la realización de los estudios de la maestría y que supieron darme las bases necesarias para la realización de este proyecto.

A todas las personas que en una u otra forma colaboraron en la realización de este trabajo.

## Resumen

La electroforesis bidimensional, una de las técnicas más empleadas para el análisis proteómico, permite separar cientos o miles de proteínas en un único gel, mostrando un patrón característico. En el análisis de estas imágenes es muy importante una correcta detección de las proteínas, ya que cualquier error en esta etapa puede llevar a la detección de falsas proteínas, o a obviar proteínas importantes, pero de baja abundancia, lo cual afectaría los resultados del análisis. Técnicas de segmentación son empleadas para separar las proteínas del fondo y encontrar anomalías. Los métodos empleados para la segmentación de imágenes de electroforesis bidimensional en gel (2DGE) se pueden clasificar como: métodos basados en detección de bordes, métodos morfológicos, umbralización, multiumbralización y métodos basados en regiones. Adicional a la detección de proteínas en imágenes 2DGE, en muchos estudios proteómicos se hace necesario la fusión o registro de imágenes para la identificación y comparación de patrones de varias muestras diferentes. Para este proceso de fusión se pueden usar las imágenes originales o los resultados de la segmentación. A pesar de los avances significativos en el campo de procesamiento de imágenes de 2DGE, no se encuentran en la literatura métodos completamente automatizados. Las herramientas comerciales disponibles para el análisis y procesamiento de imágenes 2DGE requieren que el usuario seleccione adecuadamente ciertos parámetros, de los cuales dependen los resultados arrojados por el software. Este proyecto propone una metodología de procesamiento de imágenes 2DGE que incluye la fase de segmentación y fusión. Se realiza una comparación de técnicas de segmentación usando 24 imágenes 2DGE obtenidas del veneno de apitoxina. A partir de esta comparación, los mejores resultados fueron obtenidos con la técnica de multiumbralización automática en 16 y 8 ventanas. Por su parte, la fusión de imágenes se obtiene con base en el promedio de valores de píxeles relacionados en cada par de imágenes comparadas. A partir de la metodología propuesta se logró caracterizar la apitoxina para abejas de interior y exterior, con una identificación automática de 79 de las 115 proteínas conocidas en el patrón, equivalente al 68.7%.

**Palabras Clave:** *Análisis proteómico, Electroforesis bidimensional en gel, Procesamiento de imágenes, Segmentación.*

## Abstract

Two-dimensional electrophoresis, one of the most used techniques for proteomic analysis, allows to separate hundreds or thousands of proteins in a single gel, showing a characteristic pattern. In the analysis of these images, a correct detection of proteins is very important, since any error in this stage can lead to the detection of false proteins, or to obviate important but low-abundances proteins, which would affect the results of the analysis. Segmentation techniques are used to separate proteins from the background and find anomalies. The methods used for segmentation of two-dimensional gel electrophoresis images (2DGE) can be classified as: methods based on edge detection, morphological methods, thresholding, multi-thresholding and methods based on regions. In addition to the detection of proteins in 2DGE images, in many proteomic studies it is necessary to merge or fuse images for the identification and comparison of patterns of several samples. For the fusion process, you can use the original images or segmentation results. Despite significant advances in the field of 2DGE image processing, fully automated methods are not found in the literature. The available commercial tools for the analysis and processing of 2DGE images require that the user select adequately certain parameters, on which the results of the software depend. This project proposes a 2DGE image processing methodology that includes the segmentation and fusion phase. A comparison of segmentation techniques is performed using 24 2DGE images obtained from apitoxina. From this comparison, the best results were obtained with the automatic multi-thresholding technique with 16 and 8 windows. On the other hand, image fusion is based on the average of values of related pixels of each pair of images compared. Using the proposed methodology, the apitoxin from indoor and outdoor bees was characterized, with an automatic identification of 79 from 115 known proteins, equivalent to 68.7%.

**Keywords:** *Proteomic analysis, Two-dimensional gel electrophoresis, Image processing, Segmentation.*

**Contenido**

**Resumen .....v**

**Contenido ..... vii**

**Lista de figuras.....ix**

**Lista de tablas.....xiv**

**1. Introducción .....16**

1.1 Justificación ..... 16

1.2 Planteamiento del problema..... 18

1.3 Objetivos..... 22

1.3.1 Objetivo General ..... 22

1.3.2 Objetivos Específicos ..... 22

1.4 Organización del documento ..... 22

**2. Estado del Arte .....23**

2.1 Análisis Proteómico ..... 23

2.2 Imágenes de electroforesis bidimensional 2DGE ..... 25

2.3 Procesamiento de imágenes 2DGE ..... 27

2.3.1 Preprocesamiento de imágenes 2DGE ..... 28

2.3.2 Detección de proteínas y segmentación..... 30

2.3.3 Detección de bordes ..... 32

2.3.4 Métodos Morfológicos ..... 33

2.3.5 Umbralización ..... 35

2.3.6 Métodos basados en regiones..... 36

2.3.7 Registro y fusión de imágenes ..... 37

2.4 Análisis proteómico de apitoxina ..... 38

2.5 Discusión. .... 39

**3. Marco experimental .....41**

3.1 Base de Datos..... 42

3.1.1 Adquisición de imágenes apitoxina..... 42

3.1.2 Imagen Dr. Dirk de Graaf ..... 44

3.2 Métodos de segmentación..... 45

3.2.1 Filtrado de imágenes..... 46

3.2.2 Detector de bordes Sobel ..... 46

3.2.3 Umbralización automática ..... 47

3.2.4 Multiumbralización automática para 4, 8 y 16 ventanas filtradas..... 48

3.2.5 Transformada Watershed. .... 49

3.2.6 Fronteras y etiquetas boundary y bwlabel ..... 50

3.2.7 Evaluación de los métodos de segmentación. .... 50

3.3 Fusión de imágenes ..... 51

3.4 Contribución. .... 52

**4. Resultados.....53**

---

4.1 Resultados comparación imágenes Dr. Dirk de Graaf .....	53
4.1.1 Discusión .....	58
4.2 Resultados comparación imágenes apitoxina. ....	59
4.2.1 Discusión .....	81
4.3 Resultados comparación imágenes apitoxina fusionadas. ....	82
4.3.1 Otra métrica de medición – Coeficiente Sorensen-Dice .....	102
4.3.2 Discusión .....	111
4.4 Comparación de resultados de imágenes de apitoxina con datos de informe FGN008. ....	112
4.4.1 Discusión .....	123
<b>5. Conclusiones. ....</b>	<b>125</b>
<b>6. Recomendaciones. ....</b>	<b>128</b>
<b>7. Bibliografía .....</b>	<b>130</b>



## Lista de figuras

<b>Figura 1-1:</b> Imágenes de geles de proteínas del veneno de abejas (apitoxina).....	19
<b>Figura 2-1:</b> Procesamiento y análisis de imágenes de electroforesis bidimensional (2DGE)....	24
<b>Figura 2-2:</b> Pasos que seguir en el análisis proteómico del veneno de las abejas .....	26
<b>Figura 2-3:</b> Diagrama de bloques del procesamiento de imágenes 2DGE estándar usando software especializado. (Fuente propia).....	29
<b>Figura 3-1:</b> Metodología para la identificación de objetos en imágenes 2DGE.....	41
<b>Figura 3-2:</b> Metodología para la fusión e identificación de objetos en imágenes 2DGE fusionadas .....	41
<b>Figura 3-3:</b> Imágenes de muestras de veneno de apitoxina: (a) abejas de exterior, (b) abejas de interior y (c) facilitada por Dr. Dirk de Graaf .....	43
<b>Figura 3-4:</b> Imagen 2-DGE de Peiren et al. (2008).....	45
<b>Figura 4-1:</b> Relaciones de objetos segmentados (NS), proteínas de patrón detectadas (VP), objetos segmentados no reconocidos (FP), proteínas de patrón no detectadas en técnicas de segmentación (FN) imagen Dr. Dirk de Graaf. ....	55
<b>Figura 4-2:</b> Relación de precisión (P) y sensibilidad (S) en la comparación de técnicas de segmentación imagen Dr. Dirk Dgraaf. ....	56
<b>Figura 4-3:</b> Marcado de objetos conocidos en las imágenes del Dr Dirk de Graaf empleando las técnicas descritas en la Tabla 4-2. (a) Umbralización automática; (b) Detector de bordes Sobel; (c) Transformada Watershed; (d) Multiumbral 4 ventanas; (e) Multiumbral 8 ventanas; (f) Multiumbral 16 ventanas.....	57
<b>Figura 4-4:</b> Muestras originales de imágenes del veneno de abejas de Interior según numeración de la Tabla 4-5 .....	61
<b>Figura 4-5:</b> Muestras originales de imágenes del veneno de abejas de Exterior según numeración de la Tabla 4-5 .....	62
<b>Figura 4-6:</b> (a) Imagen máster sintética.Tomado de la Figura 13 Informe FGN008. (b) Complemento de la Figura (a) estandarizado .....	63

**Figura 4-7:** Imagen máster de apitoxina donde se resaltan las proteínas con ausencia/presencia. En rojo se observan las proteínas solo presentes en apitoxina de interior, y en verde se observan las proteínas solo presentes en apitoxina de exterior. (a) Apitoxina obtenida de la finca La Isla, (b) apitoxina obtenida de la finca La Loma. Tomado de la Figura 14 del Informe FGN008.....64

**Figura 4-8:** Imagen máster de apitoxina donde se resaltan las proteínas con expresión diferencial (en rojo). (a) Apitoxina obtenida de la finca La Isla, (b) apitoxina obtenida de la finca La Loma. Tomado de la Figura 15 Informe FGN008.....64

**Figura 4-9:** Análisis global de apitoxina de exterior vs. apitoxina de interior. (a) Imagen máster de apitoxina donde se resaltan las proteínas con ausencia/presencia. En rojo se observan las proteínas solo presentes en apitoxina de interior, y en verde se observan las proteínas solo presentes en apitoxina de exterior. (b) Imagen master de apitoxina donde se resaltan las proteínas con expresión diferencial (en rojo). Tomado de la Figura 16 Informe FGN008. ....65

**Figura 4-10:** Imagen muestra Isla interior -2 bastante oscura (MII2BO) y su complemento divididas en 4 ventanas : (a) MII2BO; (b) Ventana 1 de MII2BO; (c) Ventana 2 de MII2BO; (d) Ventana 3 de MII2BO; (e) Ventana 4 de MII2BO; (f) MII2BO complemento; (g) Ventana 1 de MII2BO complemento; (h) Ventana 2 de MII2BO complemento; (i) Ventana 3 de MII2BO complemento; (j) Ventana 4 de MII2BO complemento.....67

**Figura 4-11:** Relaciones de proteínas conocidas, precisión y sensibilidad en técnica umbralización automática imágenes de interior Tabla 4-7 .....70

**Figura 4-12:** Relaciones de proteínas conocidas, precisión y sensibilidad en técnica multiumbralización automática 4 ventanas imágenes de interior Tabla 4-7 .....71

**Figura 4-13:** Relaciones de objetos segmentados (NS), proteínas de patrón detectadas (VP), objetos segmentados no reconocidos (FP), proteínas de patrón no detectadas (FN) en técnica multiumbralización automática 8 ventanas imágenes de interior Tabla 4-7 .....71

**Figura 4-14:** Relaciones de proteínas conocidas, precisión y sensibilidad en técnica multiumbralización automática 8 ventanas imágenes de interior Tabla 4-7 .....72

**Figura 4-15:** Relaciones de proteínas conocidas, precisión y sensibilidad en técnica multiumbralización automática 16 ventanas imágenes de interior Tabla 4-7 .....72

**Figura 4-16:** Imágenes resultantes de los procesos de umbralización automática y multiumbralización para 4 – 8 y 16 ventanas para las muestras Isla interior-2 bastante oscuro (MII2BO) y Loma interior1 bastante oscuro (MLI1BO): (a) MII2BO; (b) Umbral automático MII2BO; (c) Multiumbral 4 Ventanas MII2BO; (d) Multiumbral 8 Ventanas MII2BO; (e) Multiumbral 16 Ventanas MII2BO; (f) MLI1BO; (g) Umbral automático MLI1BO; (h)

Multiumbral 4 Ventanas MLI1BO; (i) Multiumbral 8 Ventanas LI1BO; (j) Multiumbral 16 Ventanas MLI1BO.....74

**Figura 4- 17:** Imágenes resultantes de los procesos de umbralización automática y multiumbralización para 4 – 8 y 16 ventanas para las muestras Isla interior-2 bastante oscuro y Loma interior1 bastante oscuro, habiendo eliminado los objetos con áreas mayores a 1000 pixeles. (a) MII2BO; (b) Umbral automático MII2BO; (c) Multiumbral 4 Ventanas MII2BO; (d) Multiumbral 8 Ventanas MII2BO; (e) Multiumbral 16 Ventanas MII2BO; (f) MLI1BO; (g) Umbral automático MLI1BO; (h) Multiumbral 4 Ventanas MLI1BO; (i) Multiumbral 8 Ventanas MLI1BO; (j) Multiumbral 16 Ventanas MLI1BO.....75

**Figura 4-18:** Relaciones de proteínas conocidas, precisión y sensibilidad en técnica umbralización automática imágenes de exterior Tabla 4-8.....78

**Figura 4- 19:** Relaciones de proteínas conocidas, precisión y sensibilidad en técnica multiumbralización automática 4 ventanas imágenes de exterior Tabla 4-8 .....79

**Figura 4-20:** Relaciones de objetos segmentados (NS), proteínas de patrón detectadas (VP), objetos segmentados no reconocidos (FP), proteínas de patrón no detectada (FN) en técnica multiumbralización automática 8 ventanas imágenes de exterior Tabla 4-8 .....79

**Figura 4-21:** Relaciones de proteínas conocidas, precisión y sensibilidad en técnica multiumbralización automática 8 ventanas imágenes de exterior Tabla 4-8 .....80

**Figura 4-22:** Relaciones de proteínas conocidas, precisión y sensibilidad en técnica multiumbralización automática 16 ventanas imágenes de exterior Tabla 4-8 .....80

**Figura 4-23:** Imágenes resultantes del proceso de fusión en dos muestras seleccionadas.(a) MII2BOX; (b) MII2BOX multiumbral 4 ventanas; (c) MII2BOX ventana 1; (d) ) MII2BOX ventana 2; (e) MII2BOX ventana 3; (f) MII2BOX ventana 4; (g) MLI1BOX; (h) MLI1BOX multiumbral 4 ventanas; (i) MLI1BOX ventana 1; (j) MLI1BOX ventana 2; (k) MLI1BOX ventana 3; (l) MLI1BOX ventana 4; (m) Fusión MII2BOX-MLI1BOX ventanas 1; (n) Fusión MII2BOX-MLI1BOX ventanas 2; (o) Fusión MII2BOX-MLI1BOX ventanas 3;(p) Fusión MII2BOX-MLI1BOX ventanas 4; (q) Fusión MII2BOX-MLI1BOX; (r) Fusión MII2BOX-MLI1BOX sin áreas mayores a 100 pixeles; (s) Detección de objetos en imágenes fusionadas ..85

**Figura 4-24:** Relaciones del consolidado total de fusión de objetos segmentados (NS), proteínas de patrón detectadas (VP), objetos segmentados no reconocidos (FP), proteínas de patrón no detectadas (FN) en técnica umbralización automática Tabla 4-11 .....92

**Figura 4- 25:** Relaciones del consolidado total de fusión de proteínas de patrón detectadas (VP), precisión (P) y sensibilidad (S) en técnica umbralización automática Tabla 4-11 .....93

**Figura 4-26:** Relaciones del consolidado total de fusión de objetos segmentados (NS), proteínas de patrón detectadas (VP), objetos segmentados no reconocidos (FP), proteínas de patrón no detectadas (FN) en técnica multiumbralización automática 4 ventanas Tabla 4-11.....93

**Figura 4-27:** Relaciones del consolidado total de fusión de proteínas de patrón detectadas (VP), precisión (P) y sensibilidad (S) en técnica multiumbralización automática 4 ventanas Tabla 4-11 .....94

**Figura 4-28:** Relaciones del consolidado total de fusión de objetos segmentados (NS), proteínas de patrón detectadas (VP), objetos segmentados no reconocidos (FP), proteínas de patrón no detectadas (FN) en técnica multiumbralización automática 8 ventanas Tabla 4-11.....94

**Figura 4-29:** Relaciones del consolidado total de fusión de proteínas de patrón detectadas (VP), precisión (P) y sensibilidad (S) en técnica multiumbralización automática 8 ventanas Tabla 4-11 .....95

**Figura 4-30:** Relaciones del consolidado total de fusión de objetos segmentados (NS), proteínas de patrón detectadas (VP), objetos segmentados no reconocidos (FP), proteínas de patrón no detectadas (FN) en técnica multiumbralización automática 16 ventanas Tabla 4-11.....95

**Figura 4-31:** Relaciones del consolidado total de fusión de proteínas de patrón detectadas (VP), precisión (P) y sensibilidad (S) en técnica multiumbralización automática 16 ventanas .....96

**Figura 4-32:** Consolidados de la fusión de todas las imágenes de interior. (a) Fusión Isla Interior Umbral automático; (b) Fusión Isla Interior Umbral automático; (c) Fusión Isla- Loma Interior Umbral automático; (d) Detección de objetos fusión Umbral automático; (e) Fusión Isla Interior Multiumbral 4 ventanas; (f) Fusión Loma Interior Multiumbral 4 ventanas; (g) Fusión Isla- Loma Interior Multiumbral 4 ventanas; (h) Detección de objetos fusión multiumbral 4 ventanas; (i) Fusión Isla Interior Multiumbral 8 ventanas; (j) Fusión Loma Interior Multiumbral 8 ventanas; (k) ) Fusión Isla- Loma Interior Multiumbral 8; (l) Detección de objetos fusión multiumbral 8 ventanas; (m) Fusión Isla Interior Multiumbral 16 ventanas; (n) Fusión Loma Interior Multiumbral 16 ventanas; (o) ) Fusión Isla- Loma Interior Multiumbral 16; (p) Detección de objetos fusión multiumbral 16 ventanas .....98

**Figura 4-33:** Consolidados de la fusión de todas las imágenes de exterior. (a) Fusión Esmeralda Exterior umbral automático; (b) ) Fusión Isla Exterior umbral automático; (c) Fusión Loma Exterior umbral automático; (d) Fusión Total Exterior umbral automático; (e) Detección objetos fusión umbral automático; (f) Fusión Esmeralda Exterior multiumbral 4 ventanas; (g) Fusión Isla Exterior multiumbral 4 ventanas; (h) Fusión Loma Exterior multiumbral 4 ventanas; (i) Fusión Total Exterior multiumbral 4 ventanas; (j) ) Detección objetos fusión multiumbral 4 ventanas; (k) Fusión Esmeralda Exterior multiumbral 8 ventanas; (l) Fusión Isla Exterior multiumbral 8 ventanas; (m) Fusión Loma Exterior multiumbral 8 ventanas; (n) Fusión Total Exterior

multiumbral 8 ventanas; (o) ) Detección objetos fusión multiumbral 8 ventanas; (p) Fusión Esmeralda Exterior multiumbral 16 ventanas; (q) Fusión Isla Exterior multiumbral 16 ventanas; (r) Fusión Loma Exterior multiumbral 16 ventanas; (s) Fusión Total Exterior multiumbral 16 ventanas; (t) ) Detección objetos fusión multiumbral 16 ventanas ..... 100

**Figura 4-34:** Consolidados de la fusión de todas las imágenes de interior y exterior. (a) Fusión Int\_Ext umbral automático todas las áreas; (b) Fusión Int\_Ext umbral automático áreas  $\leq 1000$  pixeles; (c) Detección objetos umbral automático áreas  $\leq 1000$  pixeles; (d) Fusión Int\_Ext multiumbral 4 ventanas todas las áreas; (e) Fusión Int\_Ext multiumbral 4 ventanas áreas  $\leq 1000$  pixeles; (f) Detección objetos multiumbral 4 ventanas áreas  $\leq 1000$  pixeles; (g) ) Fusión Int\_Ext multiumbral 8 ventanas todas las áreas; (h) Fusión Int\_Ext multiumbral 8 ventanas áreas  $\leq 1000$  pixeles; (i) Detección objetos multiumbral 8 ventanas áreas  $\leq 1000$  pixeles; (j) ) Fusión Int\_Ext multiumbral 16 ventanas todas las áreas; (k) Fusión Int\_Ext multiumbral 16 ventanas áreas  $\leq 1000$  pixeles; (l) Detección objetos multiumbral 16 ventanas áreas  $\leq 1000$  pixeles. .... 101

**Figura 4-35:** Coeficiente Sorensen-Dice para el consolidado de las fusiones de imágenes de los grupos de interior y de exterior, en las técnicas de umbralización y multiumbralización para 4-8 y 16 ventanas, según datos de la Tabla 4-14 ..... 111

**Figura 4-36:** Marcación y visualización de algunas proteínas con posible identidad según información dada en la Tabla 4-19 ..... 122

**Figura 4-37:** Visualización más detallada de características de 3 proteínas según información dada en la Tabla 4-19 ..... 122

## Lista de tablas

<b>Tabla 4-1:</b> Coordenadas (x, y) para las 62 proteínas en el patrón de la imagen 2DGE de apitoxina Dr. Dirk de Graaf .....	54
<b>Tabla 4-2:</b> Resumen de comparación de técnicas de segmentación con respecto al número de objetos conocidos en la imagen Dr. Dirk Dgraaf.....	55
<b>Tabla 4-3:</b> Coordenadas (x, y) de los 67 objetos con propiedades (NS) encontradas en el proceso para la técnica multiumbral 16 ventanas. ....	56
<b>Tabla 4-4:</b> Características espaciales de los 44 objetos conocidos detectados (VP) en el procesamiento de la técnica multiumbral 16 ventanas en comparación con los 62 spots del patrón de apitoxina Dr. Dirk de Graaf. ....	58
<b>Tabla 4-5:</b> Descripción y dimensiones de las imágenes originales del veneno de las abejas. ....	60
<b>Tabla 4-6:</b> Coordenadas (x,y) para las 115 proteínas del patrón de apitoxina de imágenes de abejas de interior y de exterior. ....	66
<b>Tabla 4-7:</b> Resumen general de comparación técnicas Umbralización automática, Multiumbralización en 4 - 8 y 16 ventanas para las imágenes de interior muestras La Isla y La loma, con respecto al número de objetos conocidos definidos en el patrón de apitoxina mostrados en la Tabla 4-6. ....	68
<b>Tabla 4-8:</b> Resumen general de comparación técnicas Umbralización automática, Multiumbralización en 4 - 8 y 16 ventanas para las imágenes de exterior muestras La Esmeralda, La Isla y La loma, con respecto al número de objetos conocidos definidos en el patrón de apitoxina mostrados en la Tabla 4-6.....	76
<b>Tabla 4-9:</b> Resumen general de comparación técnicas Umbralización automática, Multiumbralización en 4 - 8 y 16 ventanas para las imágenes fusionadas de interior muestras La Isla y La loma, con respecto al número de objetos conocidos definidos en el patrón de apitoxina mostrados en la Tabla 4-6.....	86
<b>Tabla 4-10:</b> Resumen general de comparación técnicas Umbralización automática, Multiumbralización en 4 - 8 y 16 ventanas para las imágenes fusionadas de Exterior muestras La Esmeralda, La Isla y La loma, con respecto al número de objetos conocidos definidos en el patrón de apitoxina mostrados en la Tabla 4-6.....	89
<b>Tabla 4-11:</b> Consolidado general de comparación de técnicas de segmentación.....	91

---

<b>Tabla 4-12:</b> Coeficiente Dice para imágenes de interior para técnicas Umbralización y multiumbralización automáticas para 4, 8 y 16 ventanas.....	103
<b>Tabla 4-13:</b> Coeficiente Dice para imágenes de Exterior para técnicas Umbralización y multiumbralización automáticas para 4, 8 y 16 ventanas.....	105
<b>Tabla 4-14:</b> Coeficiente Dice consolidado para imágenes de Interior y Exterior para técnicas Umbralización y multiumbralización automáticas para 4, 8 y 16 ventanas .....	107
<b>Tabla 4-15:</b> Tabla comparativa de spots detectados en las imágenes de apitoxina de interior y de exterior relacionadas con las proteínas registradas en el informe FGN008 .....	113
<b>Tabla 4-16:</b> Tabla relacional de la cantidad de proteínas totales detectadas en cada muestra y en la fusión de imágenes y la cantidad de proteínas iguales contenidas en las otras muestras relacionales .....	116
<b>Tabla 4-17:</b> Proteínas con expresión diferencial entre la apitoxina de obreras de exterior e interior FGN008 .....	117
<b>Tabla 4-18:</b> Proteínas presentes sólo en exterior e interior con técnicas de segmentación.....	119
<b>Tabla 4-19:</b> Posible identidad de algunas proteínas y sus características procesadas con técnica de segmentación multiumbralización 8 ventanas. ....	121

# 1. Introducción

## 1.1 Justificación

El veneno de las abejas, también conocido con el nombre de apitoxina, está compuesto de proteínas tales como melitina, apamina, adolapina, fosfolipasa, histamina y dopamina, entre otras. Estas proteínas tienen acciones neurotóxicas, analgésicas, vasoactivas, hemolíticas y citotóxicas (Peiren et al., 2005). A través del estudio de la apitoxina se busca determinar toxinas que permitan el desarrollo de fármacos, así como, comprender el proceso de envenenamiento para el desarrollo de tratamiento de picaduras masivas de abejas africanizadas.

Su acción analgésica y anti-inflamatoria ha impulsado el uso de este veneno como terapia alternativa en casos de artritis (J.-D. Lee, Park, Chae, & Lim, 2005), y por sus propiedades inmunoactivantes, se han realizados estudios para usarlo como coadyuvante en el tratamiento de la esclerosis múltiple (Mirshafiey, 2007). Recientemente, un grupo de científicos evidenciaron que la apitoxina puede inhibir la infección por VIH. Usando nano partículas cargadas con melitina se observó alteración en la membrana viral y por consiguiente una reducción en la tasa de infección, sin observar ninguna toxicidad en los tejidos normales (Hood, Jallouk, Campbell, Ratner, & Wickline, 2012). También se han divulgado estudios sobre los potenciales beneficios de la apitoxina en el tratamiento de Parkinson; en particular, se demostró que la apamina tiene efectos protectores en las neuronas dopaminérgicas in vitro (Alvarez-Fischer et al., 2013). De acuerdo con lo anterior, la apitoxina tiene una tendencia creciente a ser empleada en tratamientos medicinales complementarios o alternativos, en la llamada apiterapia o apitoxoterapia. Los efectos de la apitoxina sobre la salud aún es objeto de estudio, por lo que todavía no se ha autorizado formalmente como medicamento. Por lo tanto, no deben escatimarse recursos para el fomento de investigaciones científicas formales sobre la caracterización de la apitoxina orientado al análisis exploratorio de las propiedades medicinales o terapéuticas, a fin de aportar al desarrollo tecnológico en el área de la salud y de otras áreas afines.



Las abejas productoras de miel viven en sociedades comunales altamente complejas que incluyen labores específicas para cada uno de sus miembros: la reina, las abejas obreras y los zánganos. En la división social, las obreras de interior son abejas cuyos trabajos incluyen la limpieza, cuidado de los jóvenes y mantenimiento de la colmena, mientras que las abejas obreras de exterior cumplen roles en la defensa de la colmena y en la recolección de néctar y polen (Amdam, Csondes, Fondrk, & Page, 2006). Muchos apicultores coinciden en que la picadura de una abeja africanizada de exterior es más dolorosa que la picadura de una obrera de interior, lo cual es de esperarse ya que una de las funciones de las abejas de exterior es la defensa de la colmena. Sin embargo, no se ha realizado ninguna investigación de la composición proteica de la apitoxina de las abejas africanizadas de interior y de exterior de una misma colmena, que permita explicar la potencia de la picadura entre estas dos clases de abejas. Estudios realizados por (Schumacher & Egen, 1995) mostraron las diferencias en la letalidad en ratones causada por el veneno de las abejas africanizadas y dos poblaciones de abejas europeas de la especie *Apis mellifera*. Un estudio reciente que empleó espectrometría de masas demostró que la apitoxina de las abejas africanizadas contiene 4 proteínas únicas (Major royal jelly protein 5 precursor (MRJPs), Alpha-glucosidase precursor, Hypothetical protein LOC725114 isoform 1, Major royal jelly protein 1), que no se encuentran en las otras dos poblaciones de abejas europeas comparadas (Resende, Vasilj, Santos, Palma, & Shevchenko, 2013). En la literatura, no se encuentran estudios que permita determinar las diferencias proteómicas de las abejas africanizadas del interior y del exterior. Este proyecto busca establecer una metodología de comparación que permita encontrar marcadores diferenciales de la apitoxina a partir del análisis de imágenes de electroforesis bidimensional en gel (2-DGE, por sus siglas en inglés two-dimensional gel eletrophoresis).

La electroforesis bidimensional en gel 2DGE es una de las técnicas de preferencia en estudios proteómicos a nivel mundial, debido a su fácil manejo, el gran número de proteínas que se pueden analizar, la habilidad de detectar moléculas de alto peso molecular y porque permite diferenciar isoformas de una misma proteína que ha sufrido modificaciones (Görg, Weiss, & Dunn, 2005). El auge que ha tenido las imágenes 2DGE para análisis proteómico se evidencia por los diversos software que asiste el procesamiento de estas imágenes. Varios paquetes de software de análisis de imágenes 2D-PAGE se han desarrollado y son ampliamente utilizados por los biólogos expertos en laboratorios tales como: TotalLab 100 (Lin et al., 2009), PDQuest (BioRad) (Palagi,

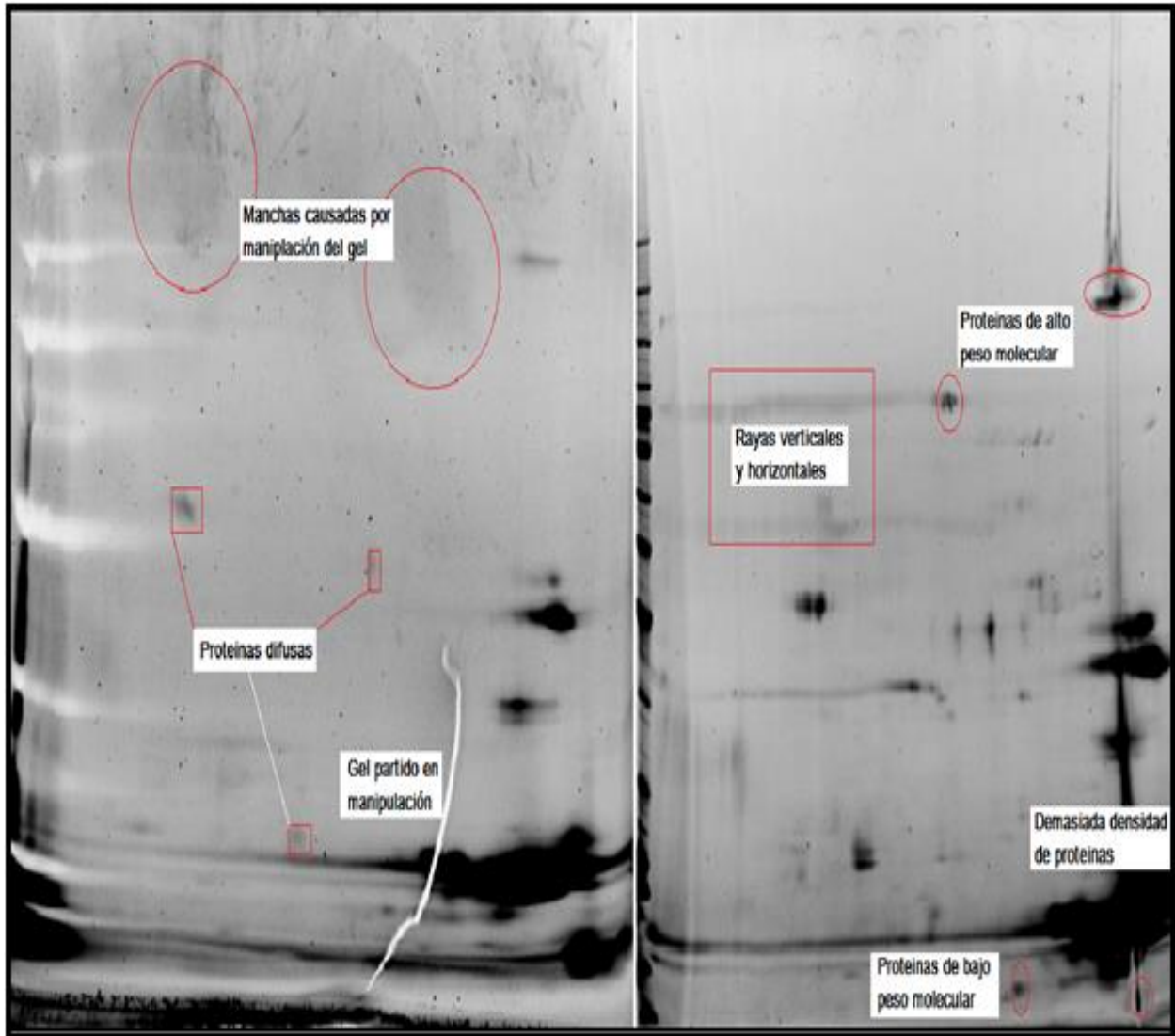
Hernandez, Walther, & Appel, 2006), ImageMaster (GE Healthcare) (Caruso et al., 2008), Melanie II (based on Melanie) (Appel, Vargas, Palagi, Walther, & Hochstrasser, 1997), PDQuest (Berth, Moser, Kolbe, & Bernhardt, 2007), Z3 (Raman, Cheung, & Marte, 2002), ProteomeWeaver (Shen, Fan, Zeng, & Cheng, 2005), Delta2D (Clark & Gutstein, 2008). Estos software usan técnicas básicas de pre-procesamiento para mejorar la apariencia de las imágenes, así como técnicas de segmentación y registro que asisten en la identificación de proteínas (Li, Seillier-Moiseiwitsch, & Korostyshevskiy, 2011). Sin embargo, estos software son en su mayoría costosos y requieren la intervención manual de un experto, por lo cual el procesamiento de imágenes 2DGE está lejos de ser un proceso automático (Brandão, Barbosa, & Arruda, 2010). El avance significativo en las capacidades computacionales para el procesamiento de imágenes y el creciente interés por el análisis proteómico basado en imágenes 2DGE, crean la necesidad de estudiar y desarrollar una metodología que permita procesar la información contenida en múltiples imágenes 2DGE de forma automática y confiable.

Este trabajo de investigación tiene como objetivo principal la elaboración de una metodología basada en procesamiento digital de imágenes para la caracterización de la apitoxina de abejas del interior y del exterior desde imágenes de electroforesis bidimensional en gel (2DGE). El proyecto pretende desarrollar una metodología automática, donde se requiera la mínima intervención humana para el procesamiento digital de las imágenes 2DGE

## 1.2 Planteamiento del problema

En el análisis de imágenes 2-DGE, se presentan muchos problemas para obtener una buena caracterización de sus componentes, razón por la cual se debe utilizar técnicas de procesamiento automático para reducir el ruido, corregir el fondo, disminuir el efecto de “streaking” (rayas verticales, rayas horizontales), y mejorar la detección de proteínas de baja abundancia (puntos difusos o con baja intensidad). En las mezclas de los reactivos se pueden presentar sustancias no deseadas que se pueden observar en la imagen como manchas. Lo mismo sucede con el veneno de las abejas, el cual puede contener componentes no propios en la solución como partículas de polvo, que se deben eliminar. En la manipulación del gel se pueden presentar: contactos con este

que se ven en la imagen como manchas, movimientos bruscos en el gel que causan roturas o particiones del mismo. La Figura 1-1 muestra algunas de las anomalías más comunes en las imágenes 2-DGE.



**Figura 1-1:** Imágenes de gels de proteínas del veneno de abejas (apitoxina)

Otro problema en el análisis de imágenes, con respecto a la detección de los puntos de interés, es la cantidad de parámetros que se deben tener en el método de segmentación, para identificar y cuantificar las manchas individuales y usualmente el usuario no está capacitado para manejar dichos parámetros, obteniéndose como resultado una segmentación deficiente de la imagen. Los

métodos paramétricos se basan en un modelo, haciendo suposiciones fuertes sobre la apariencia de los puntos de interés, pero a menudo son flexibles para representar todos los puntos que pueden estar presentes en una imagen. Los métodos no paramétricos no hacen suposiciones sobre la apariencia del punto, e imponen pocas restricciones en la detección de los puntos de interés, lo que permite más flexibilidad, pero reducen la robustez cuando los datos de la imagen son complejos métodos. El modelo de detección debe permitir generar nuevas formas de puntos, que pertenezcan a un misma distribución estadística, en donde la apariencia y la forma de los puntos de interés sean capaces de ajustarse a los datos de las imágenes, mejor que las parametrizaciones de puntos realizadas con modelos gaussianos y de difusión (Rogers, Graham, & Tonge, 2003).

La detección de puntos y la segmentación son componentes fundamentales de análisis de imágenes 2DGE, pero siguen siendo tareas arduas y difíciles debido a las características de las imágenes. Varios paquetes de software y los enfoques académicos están disponibles para la detección de puntos de imagen 2DGE y la segmentación. Cada uno tiene sus respectivas ventajas y desventajas y alcanza un nivel diferente de éxito en el tratamiento de los retos de análisis de imágenes 2DGE. Una característica común de los métodos disponibles es su dependencia de la intervención del usuario con el fin de lograr resultados óptimos, un proceso que puede conducir a resultados subjetivos y no reproducibles (E. Kostopoulou, Katsigiannis, & Maroulis, 2015). El análisis de los datos proteómicos, presentados en forma de electroferogramas 2DGE (gráfico realizado con los resultados de un análisis por electroforesis), sigue siendo un verdadero desafío y requiere nuevas mejoras. Tres problemas principales en el análisis de datos se asocian con: (1) 'varianza introducida en el software' (debido a los diferentes métodos de pre-procesamiento de datos o de sus diferentes parámetros de entrada), (2) análisis de significancia univariable, y (3) los requisitos de detección de manchas y cuantificación (Daszykowski, Færgestad, Grove, Martens, & Walczak, 2009).

Los resultados que se obtienen con la utilización de diferentes software comerciales, en el análisis de imágenes 2DGE presentan muchas diferencias, que dificultan una adecuada parametrización de la información, por lo cual las comparaciones realizadas de los resultados no tienen una evaluación cualitativa en cuanto a precisión de sus procedimientos y poca capacidad para replicar

resultados obtenidos dificultando la comparación con diferentes software (Wheelock & Buckpitt, 2005).

Actualmente no existe un método de segmentación que alcance resultados aceptables para todo tipo de imagen 2DGE. No existen métodos que sean generales y que puedan ser aplicados a cualquier variedad de datos. De cualquier forma, los métodos que son especializados para aplicaciones particulares pueden obtener mejores resultados tomando en cuenta conocimiento a priori. Por lo tanto, la selección de un método apropiado para un problema de segmentación puede ser muy difícil.

El procesamiento y análisis automático de imágenes de electroforesis bidimensional en gel 2DGE tiene grandes desafíos, debidos a la variabilidad implícita en las muestras biológicas, la variabilidad introducida por la preparación de las muestras, el sistema de adquisición de la imagen y el proceso de identificación de las proteínas. Como se puede apreciar en la literatura, son escasos los estudios formales aplicados al análisis proteómico, en donde se desarrollen aplicaciones efectivas de procesamiento de imágenes 2DGE, que estén enfocados a la identificación de marcadores diferenciales, que permitan la caracterización proteómica de una clase de individuos, lo cual constituye vacíos en el procesamiento de imágenes 2DGE, que exigen el desarrollo de nuevas y efectivas metodologías de análisis, que tomen ventajas de las herramientas computacionales para el diseño de nuevos aplicativos en esta área de conocimiento. Las colmenas de abejas son divididas en dos poblaciones o clases (de interior y de exterior) y para cada una de las clases de individuos se toman muestras de veneno (apitoxina). De acuerdo a la caracterización de las imágenes de geles bidimensionales de apitoxina, se pretende establecer una comparación entre estas dos clases que permitan determinar variaciones de la composición del veneno según la actividad social de las abejas obreras.

De acuerdo con la caracterización de los componentes proteicos del veneno de la abeja africanizada, mediante marcadores diferenciales morfológicos derivados del procesamiento digital de imágenes, se logrará establecer una metodología de comparación que ayude a determinar las variaciones de la composición de la apitoxina según la actividad social de las abejas obreras (de interior o de exterior).

## 1.3 Objetivos

### 1.3.1 Objetivo General

Proponer una metodología para la identificación de marcadores espacio-morfológicos basada en procesamiento digital de imágenes 2DGE aplicado al análisis proteómico para la caracterización del veneno de las abejas africanizadas.

### 1.3.2 Objetivos Específicos

- Identificar descriptores morfológicos y espaciales en las imágenes de electroforesis bidimensional en gel, obtenidos desde el veneno de las abejas llamado apitoxina.
- Analizar el espacio de representación para establecer la relación proteómica del veneno de las abejas obreras: del interior y del exterior, buscando las características comunes y discriminantes desde los descriptores espaciales extraídos de las imágenes de electroforesis bidimensional en gel 2DGE.

## 1.4 Organización del documento

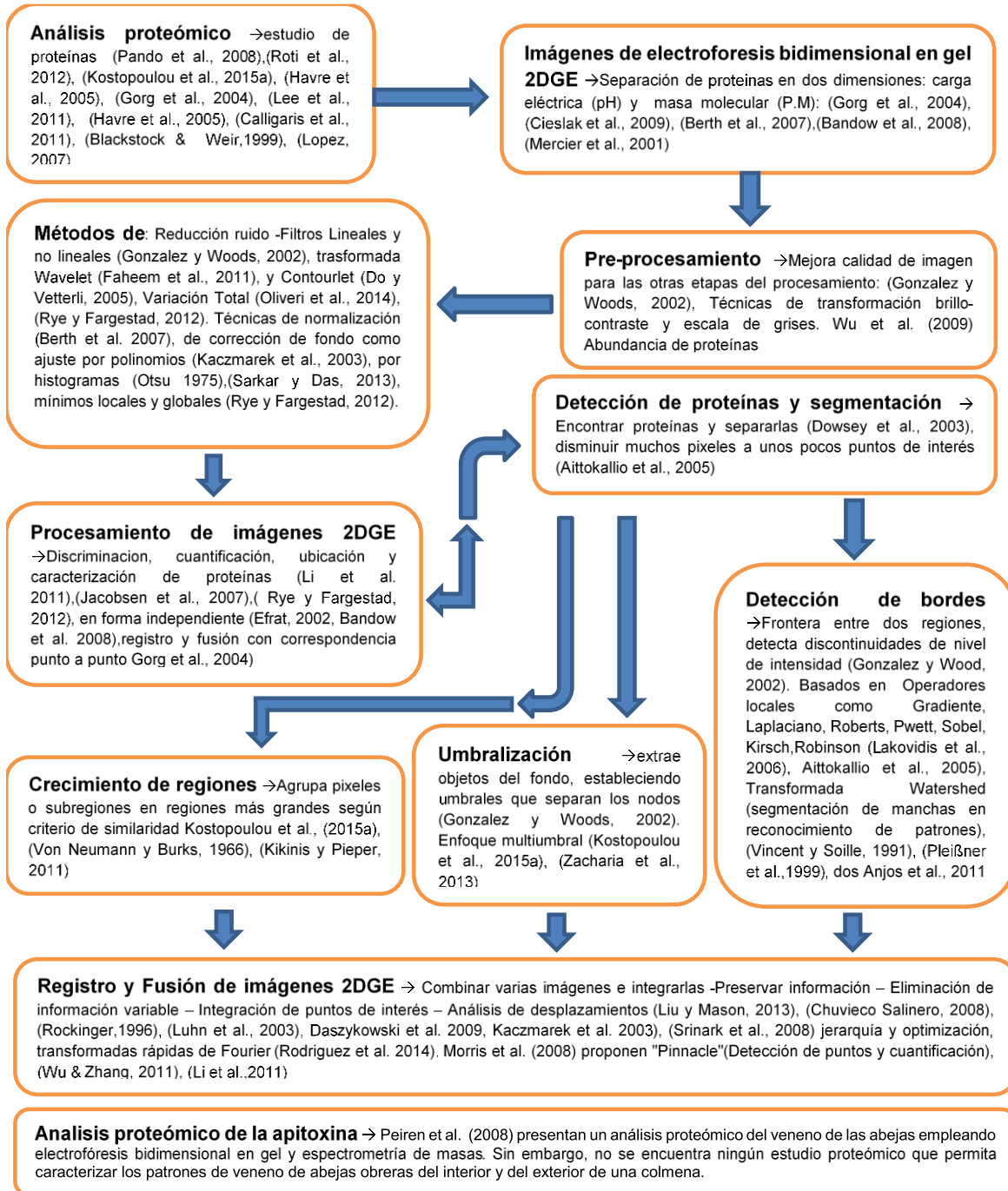
El Capítulo 2 presenta una revisión detallada de literatura sobre los métodos y procedimientos para el procesamiento y análisis de imágenes de electroforesis bidimensional en gel –2DGE. El Capítulo 3 presenta el marco experimental formulado para el desarrollo de este trabajo de investigación; en este capítulo se incluye la descripción de las bases de datos usadas, la metodología de procesamiento de imágenes 2DGE propuesta y la descripción de los experimentos realizados para su validación. El Capítulo 4 incluye los resultados que se obtienen al procesar imágenes 2DGE con la metodología propuesta, así como, el análisis y discusión de los resultados. Por último, el Capítulo 5 presenta las conclusiones de este trabajo y recomendación para trabajos futuros.

## 2. Estado del Arte

Este capítulo presenta una revisión sobre las corrientes, técnicas y algoritmos empleados para el procesamiento y análisis de imágenes de electroforesis bidimensional (2DGE). Inicialmente, se describe el análisis proteómico y el proceso para la adquisición de imágenes 2DGE. El procesamiento de imágenes 2DGE incluye: preprocesamiento, detección de proteínas y fusión de imágenes. Esta revisión presenta las principales técnicas y algoritmos empleados en cada una de estas etapas. Por último, se revisan trabajos previos en el análisis proteómico del veneno de las abejas, lo cual permite establecer un referente para validar la contribución de este proyecto de tesis (Ver Figura 2-1).

### 2.1 Análisis Proteómico

La proteómica es el estudio de las proteínas que se encuentran en células, tejidos y órganos de los organismos biológicos (Pando-Robles & Lanz-Mendoza, 2009; Roti & Stegmaier, 2012). Las proteínas son las encargadas de ejecutar y controlar la gran mayoría de los procesos celulares. Actualmente, existen diversas aplicaciones basadas en el estudio del proteoma (E. Kostopoulou et al., 2015). Por ejemplo, el análisis sistemático del proteoma permite identificar proteínas cuya presencia, ausencia o alteración están correlacionados con determinados estados fisiológicos (Görg et al., 2005). El campo de la proteómica es ampliamente utilizado en el estudio a gran escala de proteínas en las células (Havre, Singhal, Payne, Weir Lipton, & Webb-Robertson, 2005), con el fin de contribuir a la evaluación de nuevos fármacos, la identificación de nuevos biomarcadores y la exploración de eventos biológicos (J. Lee, Han, Altwerger, & Kohn, 2011). El mayor desafío está en comprender la relación entre las alteraciones en los perfiles proteómicos y la expresión de enfermedades, permitiendo la identificación de nuevos biomarcadores y el desarrollo de nuevos medicamentos (Roti & Stegmaier, 2012).



**Figura 2-1:** Procesamiento y análisis de imágenes de electroforesis bidimensional (2DGE).

El análisis proteómico se puede dividir en tres etapas: la separación de proteínas, la identificación de proteínas y el análisis de datos (Blackstock & Weir, 1999) (Ver Figura 2-2). En la separación de proteínas se encuentran métodos basados en electroforesis y cromatografía. Algunos de los métodos de electroforesis más destacados son (Calligaris, Villard, & Lafitte, 2011): electroforesis



bidimensional de gel (2DGE), y electroforesis de elución continua. Por su parte, la cromatografía se basa en el principio de retención selectiva. Algunos de los métodos empleados para la separación de proteínas por cromatografía son (Calligaris et al., 2011): cromatografía líquida de fase inversa (RPLC), cromatografía de interacción hidrofílica (HILIC), cromatografía de intercambio iónico, focalización y cromatografía líquida de ultra-alto rendimiento (UHPLC). Una vez separadas las proteínas, se utilizan técnicas basadas en espectrometría de masa para su identificación. La espectrometría de masas permite la identificación de moléculas y la composición de diferentes elementos químicos que forman un compuesto. Finalmente, el análisis de los datos obtenidos en las etapas de separación e identificación permite: estudiar los cambios globales de las proteínas o determinar la función de una proteína específica (Blackstock & Weir, 1999). El análisis de datos es soportado por diferentes herramientas computacionales tales como ProSight, Dymension, ProteinMine, DeCyder, entre otros.

## **2.2 Imágenes de electroforesis bidimensional 2DGE**

La electroforesis bidimensional es una técnica de alta resolución que permite la separación en dos etapas de mezclas complejas de proteínas (Görg et al., 2005). En primera instancia las proteínas son separadas en función de su carga a lo largo de un gel con gradiente de pH (variación del pH en una determinada dirección) (Cieslak & Ribera, 2009). Cada proteína avanza en el campo eléctrico hasta alcanzar un valor de pH igual a su punto isoeléctrico (cuando la proteína alcanza una carga neta 0). En la segunda dimensión las proteínas son separadas en función de su masa molecular. A partir de esta separación se genera un patrón de proteínas que es digitalizado como una imagen mediante el escaneo del gel.

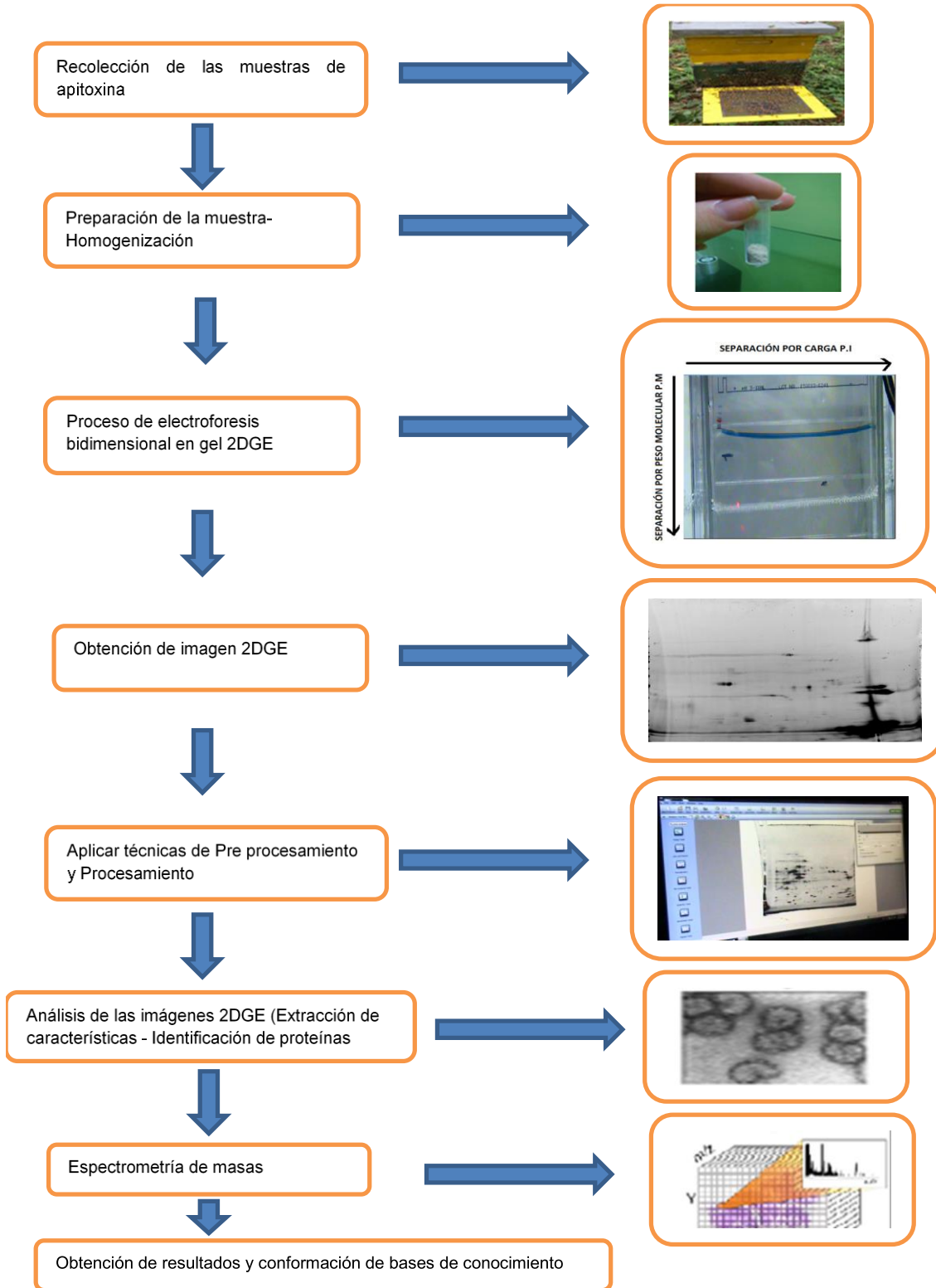


Figura 2-2: Pasos que seguir en el análisis proteómico del veneno de las abejas

Estas imágenes permiten el análisis e identificación de proteínas (Bandow et al., 2008; Berth et al., 2007) , a partir de lo cual se puede determinar cambios en las proteínas y abundancias de las mismas (Calligaris et al., 2011), así como identificar patrones proteicos de diferentes grupos (Görg et al., 2005). El proceso de electroforesis se ha utilizado en diferentes áreas de aplicación, tales como la investigación del cáncer, el diagnóstico de enfermedades infecciosas y sus tratamientos, así como la investigación en farmacología (J. Lee et al., 2011).

Cada proteína se caracteriza por su posición vertical y horizontal en la imagen 2DGE, así como por su tamaño e intensidad. Sin embargo, la posición, y abundancia de una proteína puede verse afectada por diversas causas. La variación en la posición de una sola especie de proteína a través de un conjunto de geles se debe a una combinación de efectos globales (cambio y rotación del gel en el proceso de escaneado, expansión y contracción durante la rehidratación y secado) y los efectos locales (falta de homogeneidad en el gel o variaciones en las condiciones debido a cambios de temperatura o corriente de fuga (Mercier & Slater, 2001)). En general, en el proceso de electroforesis, la preparación y manipulación del gel, así como la digitalización del mismo, genera algunas anomalías en las imágenes 2DGE. Entre las anomalías más comunes se encuentran: ruido generado por polvo o residuos que quedan depositados en el gel y que puede llegar a confundirse con pequeñas proteínas; diferentes niveles de saturación debido a la iluminación del sistema de adquisición que puede resaltar algunas proteínas y atenuar otras; y rayas verticales y horizontales, generadas por la concentración de los reactivos empleados para la separación de proteínas. Estas anomalías hacen necesario el uso de técnicas de procesamiento de imágenes digitales para mejorar la calidad de las imágenes 2DGE y reducir el efecto de las anomalías en la detección e identificación de las proteínas. El procesamiento de imágenes se considera una etapa fundamental para el análisis proteómico a partir de imágenes 2DGE. La siguiente sección presenta las principales corrientes de procesamiento empleadas en imágenes 2DGE.

### **2.3 Procesamiento de imágenes 2DGE**

El procesamiento de imágenes 2DGE es importante y ampliamente usado para el análisis proteómico, debido a la dificultad para discriminar manchas de las proteínas reales, la presencia

de ruido, la cuantificación de los niveles de proteínas y la ubicación de las proteínas (Li et al., 2011). Se pueden encontrar en la literatura dos corrientes de procesamiento de un conjunto de imágenes 2DGE (Bandow et al., 2008; Jacobsen et al., 2007; Rye & Fargestad, 2012). La primera corriente realiza una detección e identificación de las proteínas en cada gel de forma independiente (Bandow et al., 2008; Efrat, Hoffmann, Kriegel, Schultz, & Wenk, 2002). Una vez esta detección es realizada, se identifican proteínas comunes en el conjunto de geles analizados (Figura 2-3(a)). La segunda corriente realiza un registro de las imágenes, que mediante transformaciones geométricas alinean las imágenes de tal manera que se obtiene una correspondencia punto a punto (Bandow et al., 2008; Görg et al., 2005). Una vez se tienen las imágenes registradas, se obtiene una única imagen que contiene toda la información del conjunto de geles analizados (Figura 2-3(b)). Sobre esta única imagen se realiza la detección e identificación de proteínas, y se usa como plantilla para identificar las proteínas en cada gel analizado.

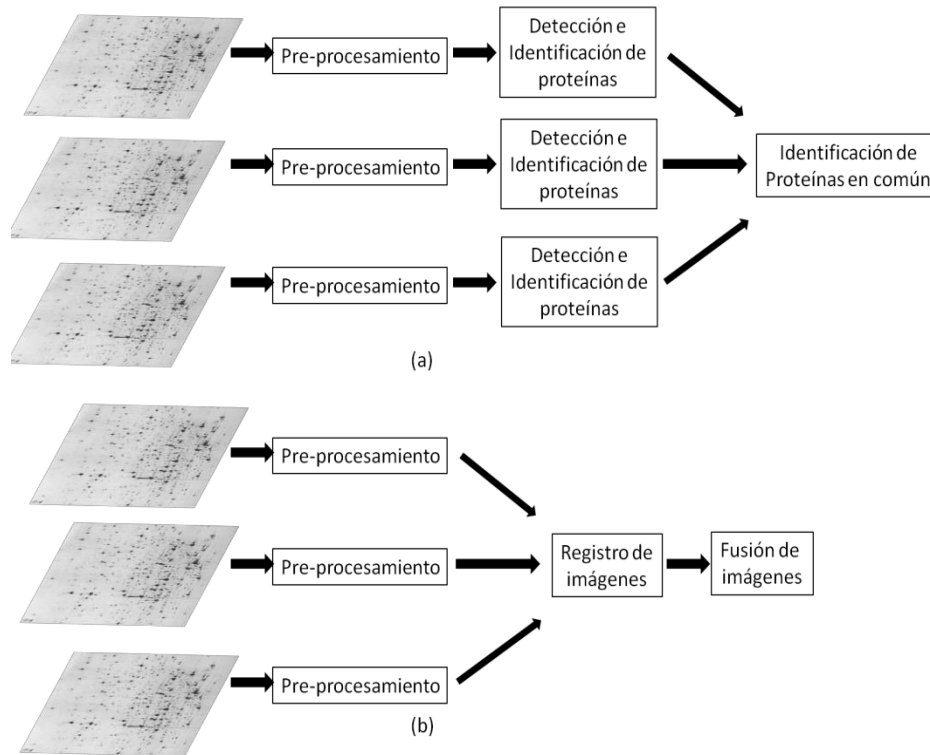
A partir de estas dos corrientes (Figura 2-3), se puede notar que en el procesamiento de imágenes 2DGE se realizan tres etapas: pre-procesamiento, identificación de proteínas y fusión de imágenes. Estas dos últimas se alternan de acuerdo a la corriente de procesamiento que se esté empleando.

### **2.3.1 Preprocesamiento de imágenes 2DGE**

El preprocesamiento de imágenes busca mejorar la calidad de la imagen de forma que aumente la posibilidad de éxito en las etapas posteriores (Gonzalez & Woods, 2002). “Wu et al. (2009) presentan un estudio sobre las técnicas de transformación de ajuste de brillo, contraste y rango de escala de grises”, para identificar su impacto en la información contenida en las imágenes 2DGE empleando TotalLab 100, una herramienta computacional que soporta el análisis de geles. De acuerdo con Wu et al. (2009), este tipo de transformaciones modifican la información proporcionada por las imágenes 2DGE, principalmente la información relacionada con la abundancia de la proteína.

En general, los métodos de pre-procesamiento para imágenes 2DGE se pueden clasificar en: métodos de reducción ruido, normalización de intensidades y corrección de fondo (Rye &

Fargestad, 2012); los cuales buscan mejorar los detalles de los puntos, reducir el ruido y homogeneizar el fondo para facilitar la detección de las proteínas.



**Figura 2-3:** Diagrama de bloques del procesamiento de imágenes 2DGE estándar usando software especializado. (Fuente propia).

En la reducción de ruido se pueden emplear filtros lineales o no lineales. Sin embargo, los filtros lineales reducen la intensidad de las proteínas y hacen que los bordes difuminados de las proteínas tengan un efecto secundario no deseado (Gonzalez & Woods, 2002). Por su parte, los filtros no lineales permiten preservar los bordes, pero al ser técnicas más robustas requieren de más parámetros. Los filtros no lineales más empleados para la reducción de ruido en imágenes 2DGE son las técnicas basadas en la transformada Wavelet (Faheem, Rashwan, Sarhan, & Youssef, 2011), la transformada Contourlet (Do & Vetterli, 2005) y Variación Total (Oliveri, Anselmi, & Massa, 2014).

Por otra parte, la normalización de imágenes 2DGE busca mejorar el contraste de las proteínas. Las técnicas más empleadas de normalización de imagen 2DGE se pueden dividir en dos categorías. La primera usa la información de los píxeles en un único gel (Berth et al., 2007), y la segunda categoría emplea la información desde múltiples geles para normalizar la intensidad (Rye & Fargestad, 2012). Ambas categorías se basan en el histograma para realizar la normalización de la imagen.

Las técnicas de corrección de fondo de imágenes 2DGE permiten obtener un fondo más uniforme que favorece la detección de las proteínas. Idealmente, se espera que los píxeles en el fondo de una imagen 2DGE sean cercanos al valor máximo de intensidad, correspondiente al color blanco. Las técnicas más comunes para la corrección del fondo se basan en ajuste por polinomios (Kaczmarek, Walczak, Jong, & Vandeginste, 2003), métodos de umbrales multinivel basados en ajuste por histogramas, que equivale a segmentar una imagen de nivel de gris en varias regiones, mejorando la separación entre los objetos y logrando mayor eficiencia, comparada con la técnica basada en histograma 1D (Sarkar & Das, 2013); método de Otsu mejorado, basado en un histograma bidimensional que mejora la división de las regiones, presenta más resistencia a los ruidos, mejor consistencia de las regiones segmentadas y fronteras más precisas (Otsu, 1979) y métodos de mínimos locales y globales (Rye & Fargestad, 2012).

### **2.3.2 Detección de proteínas y segmentación**

Los resultados del análisis proteómico dependen en gran medida de la etapa de detección de proteínas (Dowsey, Dunn, & Yang, 2003). Detectar falsas proteínas o no reconocer otras, conduce a conclusiones erróneas sobre las características proteómicas de una célula o tejido. Adicionalmente, dado que las imágenes 2DGE tienen cientos de proteínas, la detección permite reducir la cantidad de datos desde una gran cantidad de píxeles a algunos puntos característicos (Aittokallio, Salmi, Nyman, & Nevalainen, 2005). Sin embargo, la detección de proteínas debe enfrentar varios desafíos. En muchos casos, los puntos tienen un bajo contraste que dificulta la visualización de las proteínas. Además, los puntos adyacentes se superponen y por lo tanto son

difíciles de separar. Por último, la calidad de estas imágenes se degrada debido a la existencia de ruido, manchas y fondo no homogéneo.

Para la detección de estos puntos característicos se pueden identificar dos metodologías en la literatura: la detección de proteínas independientemente en una imagen 2DGE o identificación de características comunes desde una imagen fusionada desde varias muestras de geles (ver Figura 2-3). En ambos casos se emplean técnicas de segmentación para separar las proteínas del fondo y las anomalías en la imagen 2DGE.

La segmentación divide una imagen en sus partes constituyentes u objetos. Típicamente estos algoritmos se basan en determinar discontinuidades o en algún criterio de similitud (Gonzalez & Woods, 2002). En la primera categoría, la imagen se divide teniendo en cuenta los cambios bruscos de intensidades. Ejemplos de estas técnicas son la detección de puntos aislados, líneas y bordes de una imagen. Por otra parte, los métodos basados en similitud dividen la imagen de acuerdo a la homogeneidad de la intensidad. Los métodos de similitud más comunes son la umbralización, crecimiento de regiones, unión y división de regiones y métodos morfológicos (Gonzalez & Woods, 2002).

Particularmente para imágenes 2DGE, los métodos de segmentación que se encuentran con más frecuencia en la literatura son: la detección de bordes (Appel et al., 1997; P.F. Lemkin & Lipkin, 1981; Peter F. Lemkin & Lipkin, 1983; Raman et al., 2002; Solomon & Harrington, 1993), métodos morfológicos (dos Anjos, Moller, Ersboll, Finnie, & Shahbazkia, 2011; Kim, Kim, Won, & In, 2003; Eirini Kostopoulou, Zacharia, & Maroulis, 2014; Savelonas, Mylona, & Maroulis, 2012; Srinark & Kambhamettu, 2008; Sun & Wang, 2010), umbralización (Cutler, Heald, White, & Ruan, 2003; Eirini Kostopoulou, Katsigiannis, & Maroulis, 2015; Sengar, Upadhyay, Singh, & Gadre, 2016; Zacharia, Kostopoulou, Maroulis, Anagnou, & Pappa, 2013) y métodos basados en regiones (Iakovidis, Maroulis, Zacharia, & Kossida, 2006; Eirini Kostopoulou, Zacharia, & Maroulis, 2012; Sengar, Upadhyay, Singh, & Gadre, 2012). A continuación, se presenta una breve revisión de cada una de estas corrientes en la segmentación de imágenes 2DGE.

### 2.3.3 Detección de bordes

La detección de bordes es el método más común para detectar discontinuidades de intensidad (Gonzalez y Wood, 2002). Un borde se define como la frontera entre dos regiones con intensidades de gris relativamente distintas. Los cambios abruptos en los valores de la función intensidad de una imagen están asociados a valores elevados de la primera derivada de dicha función. Para la detección de discontinuidades se emplea una máscara que se desplaza a través de una imagen calculando un operador local de derivación, tales como: el gradiente, Laplaciano, Roberts, Prewitt, Sobel, Kirsch, Robinson, entre otros (Gonzalez y Wood, 2002).

Matemáticamente, el gradiente de una función de dos variables (en este caso, la función de intensidad de la imagen) para cada punto es un vector bidimensional cuyos componentes están dados por las primeras derivadas de las direcciones verticales y horizontales. Para cada punto de la imagen, el vector gradiente apunta en dirección del incremento máximo posible de la intensidad, y la magnitud del vector gradiente corresponde a la cantidad de cambio de la intensidad en esa dirección (Patnaik, S. and Yang, Y.M. (2012).

El detector de borde Prewitt es una forma apropiada de estimar la magnitud y la orientación de un borde y se involucran a los vecinos de filas / columnas adyacentes para proporcionar mayor inmunidad al ruido. El operador Prewitt está limitado a ocho posibles orientaciones; sin embargo los experimentos muestran que las estimaciones de la orientación directa no es tan precisa (Prewitt, 1970).

El operador Sobel es un operador diferencial discreto que calcula una aproximación al gradiente de la función de intensidad de una imagen. Para cada punto de la imagen a procesar, el resultado del operador Sobel es tanto el vector gradiente correspondiente como la norma de este vector. Así, para cada punto, este operador da la magnitud del mayor cambio posible, la dirección de este y el sentido desde oscuro a claro. (Patnaik, S. and Yang, Y.M. (2012).



El operador de bordes Roberts obtiene buena respuesta ante bordes diagonales. Ofrece buenas prestaciones en cuanto a localización. El gran inconveniente de este operador es su extremada sensibilidad al ruido y por tanto tiene pobres cualidades de detección (Patnaik, S. and Yang, Y.M. (2012).

Varios software especializados en el procesamiento de imágenes 2DGE emplean detectores de bordes para la segmentación de las proteínas (Appel et al., 1997; Iakovidis et al., 2006; P.F. Lemkin & Lipkin, 1981; Peter F. Lemkin & Lipkin, 1983; Raman et al., 2002; Solomon & Harrington, 1993). Sin embargo, uno de los principales problemas que presenta la segmentación de proteínas usando detectores de borde es que estos no diferencian entre proteínas y anomalías, tales como manchas verticales y horizontales, ruido o roturas del gel (Aittokallio et al., 2005; Iakovidis et al., 2006). Adicionalmente, los métodos basados en detección de borde no son capaces de diferenciar proteínas superpuestas.

### 2.3.4 Métodos Morfológicos

La transformada Watershed se ha empleado ampliamente para mejorar los resultados de segmentación de imágenes 2DGE. La transformada Watershed, introducida en 1977 en morfología matemática, se emplea principalmente en problemas de segmentación de imágenes en escala de grises. Una watershed se puede definir como una línea divisoria que separa diferentes cuencas (*cathment basin*) en una imagen en escala de grises. Una cuenca es un conjunto de píxeles asociado a un punto mínimo tal que desde cada píxel se puede alcanzar el mínimo siguiendo un camino descendente (Vincent & Soille, 1991).

La principal desventaja de la transformada Watershed para la segmentación de imágenes 2DGE es su alta sensibilidad al ruido que resulta en una sobre-segmentación. Establecer los mínimos locales se dificulta por la presencia del ruido en las imágenes 2DGE (Sun & Wang, 2010). Para solucionar este problema existen diferentes alternativas que combinan la transformada Watershed con otras técnicas como la umbralización (dos Anjos et al., 2011; Kim et al., 2003), la unión y separación de regiones (Srinark & Kambhamettu, 2008), y el uso de marcadores (Sun & Wang, 2010).

“Kim et al., (2003) usan la transformada Watershed en combinación con umbralización para la segmentación de imágenes 2DGE”. Uno de los problemas de la segmentación basada en Watershed es establecer los puntos mínimos en la imagen, dado que las regiones usualmente se forman por nivel de gris continuos; lo que conlleva a una sobre-segmentación de la imagen. Para solucionar este problema, en este mismo trabajo realizan primero la umbralización de la imagen con un valor alto de nivel de gris produciendo una región muy grande a la cual se aplica la transformada Watershed. Si la transformada obtiene dos regiones, el proceso de umbralización-Watershed se vuelve aplicar en la región más grande. Con este proceso se busca eliminar la necesidad de establecer mínimos y realizar un post-procesamiento para unir o separar regiones.

(Srinark & Kambhamettu, 2008) parten de la transformada Watershed para dividir la imagen 2DGE en múltiples regiones. Cada región es segmentada usando K-means identificando el fondo y pixeles que pertenecen a proteínas. Posteriormente, se identifica el centroide de cada proteína, a partir de los cuales se realiza un proceso de división o unión de regiones para evitar problemas de sobre-segmentación.

El uso de marcadores que restringen la transformada Watershed ayuda a mejorar los problemas de sobre-segmentación que se obtienen al aplicar esta técnica en imágenes 2DGE. (Sun & Wang, 2010) usan una técnica de filtrado basado en la distribución de intensidad de los pixeles para definir los marcadores, los cuales se emplean como los mínimos locales para calcular las Watersheds.

Adicional a la transformada Watershed, se han introducido otros métodos morfológicos para la segmentación automática de proteínas en imágenes 2DGE. Por ejemplo, (Savelonas et al., 2012) propone un plan basado en la mejora de los contornos activos y su explotación para la segmentación de proteínas, cuyos objetivos son la detección de regiones formadas por multicapas, la detección de proteínas sobrepuesta y su segmentación, adaptación de histogramas y reconstrucciones morfológicas para eliminar ruidos no deseados, rayas y control de las manchas asociadas con la intensidad de los objetos y de la región. La evaluación del esquema es direccionada con la información de bases de datos reales e imágenes sintéticas 2DGE, como

también con las comparaciones cuantitativas presentes en el estado del arte y en algunos paquetes de software empleados en el análisis de imágenes 2DGE.

### 2.3.5 Umbralización

La umbralización (thresholding) es un método que busca segmentar imágenes escalares creando una partición binaria de las intensidades de las imágenes. Una umbralización trata de determinar un valor de intensidad, llamado umbral (threshold), que separa las clases deseadas. La segmentación se logra agrupando todos los píxeles con mayor intensidad al umbral en una clase, y todos los otros píxeles en otra clase. La determinación de más de un valor umbral es un proceso llamado multiumbralización (multithresholding). La umbralización es una técnica efectiva para obtener la segmentación de imágenes donde estructuras diferentes tienen intensidades contrastantes u otras características diferenciables. La partición usualmente es generada interactivamente, pero también existen métodos automáticos. Para imágenes escalares, los métodos interactivos pueden estar basados en la apreciación visual del usuario ya que la operación de umbralización es implementable en tiempo real. Generalmente, la umbralización es el paso inicial de una secuencia de operaciones de procesamiento de imágenes. Su principal limitación es que en su forma más simple solo se generan dos clases y por lo tanto no se puede aplicar a imágenes multicanal. En adición, la umbralización usualmente no toma en cuenta las características espaciales de la imagen. Esto causa que sea sensible al ruido o a intensidades no homogéneas (Sahoo et al., 1988).

La umbralización es uno de los métodos más importantes y sencillos para la segmentación de imágenes. Esta técnica permite separar objetos del fondo a partir de un umbral sobre la intensidad de los píxeles (Gonzalez y Woods, 2002). En procesamiento de imágenes 2DGE es común encontrar enfoques multiumbral para la detección de las proteínas (Eirini Kostopoulou, Katsigiannis, & Maroulis, 2015; Sengar et al., 2016; Zacharia et al., 2013). Por ejemplo, (Eirini Kostopoulou, Katsigiannis, Maroulis, Pappa, & Anagnou, 2015) dividen la imagen en ventanas de tamaño fijo, en cada una de las cuales se aplica el proceso de umbralización.

El principal problema que presentan los algoritmos de umbralización es su alta sensibilidad al ruido, anomalías y fondo no homogéneo de las imágenes 2DGE, lo cual resulta en la detección de falsas proteínas (Eirini Kostopoulou, Katsigiannis, & Maroulis, 2015). Por esta razón, se encuentra en la literatura la combinación de las técnicas de umbralización con otros métodos de procesamiento de imágenes. Por ejemplo, (Zacharia et al., 2013) combina un filtro basado en la información de textura multidireccional con un detector de la máxima intensidad de regiones (umbralización local). Por su parte, (Sengar et al., 2016) emplean un método de umbralización que no usa el valor de intensidades para separar las proteínas del fondo, sino emplea los coeficientes wavelets para separar los puntos, su borde y el fondo de las imágenes 2DGE.

### **2.3.6 Métodos basados en regiones**

El crecimiento de regiones busca agrupar píxeles o subregiones dentro de regiones más grandes de acuerdo algún criterio de similitud. El método se basa en la agregación de píxeles similares y espacialmente próximos a partir de un conjunto de semillas aleatorias. La similitud entre píxeles se determina a partir del nivel de intensidad, color o textura (Gonzalez y Woods, 2002). En segmentación de proteínas se usa el crecimiento de regiones combinado con otras técnicas. (E. Kostopoulou et al., 2015) utilizan un algoritmo de corte-crecimiento combinada con crecimiento de regiones para segmentar proteínas en imágenes 2DGE.

Para la segmentación de proteínas en imágenes 2DGE también encontramos métodos paramétricos, los cuales modelan las proteínas a través de una función, usualmente una distribución Gaussiana bidimensional (Iakovidis et al., 2006). Sin embargo, este tipo de modelos paramétricos tiene problemas para detectar proteínas saturadas, así como proteínas superpuestas. (Iakovidis et al., 2006) proponen el uso de múltiples funciones para modelar las proteínas en las imágenes 2DGE, empleando un algoritmo genético para la estimación de los parámetros que mejor se adecuen a la imagen. A través del multi-modelos (Iakovidis et al., 2006) logran detectar proteínas superpuestas, sin embargo, el costo computacional del proceso de optimización es bastante alto.

### 2.3.7 Registro y fusión de imágenes

La fusión de imágenes consiste en la combinación de datos desde varias imágenes. Los objetivos de la fusión de imágenes son: realzar elementos que no son visibles, complementar datos para mejorar una clasificación, detectar cambios en una zona usando datos multitemporales, reemplazar datos anómalos, adicionar datos faltantes. El proceso de fusión debe satisfacer tres condiciones (Rockinger, 1996): preservación de toda la información pertinente, eliminación de información irrelevante y ruido, minimización de los artefactos y las inconsistencias en la imagen fusionada.

En el análisis proteómico de imágenes 2DGE, es altamente deseable conocer un conjunto significativo de posiciones estándar que permitan identificar proteínas en una colección de geles (Luhn, Berth, Hecker, & Bernhardt, 2003). Una posición estándar debe cumplir los siguientes requisitos: (1) la posición estándar de una especie de proteína debe estar razonablemente cerca de su posición sobre cualquier imagen de gel; (2) las relaciones espaciales entre los puntos se deben preservar lo mejor posible; (3) debería ser posible encontrar puntos correspondientes en un conjunto de geles basados en sus posiciones normales; y (4) la posición estándar debe permitir la predicción de la especie de proteína de un lugar no identificado, en función de su posición. Sin embargo, la posición de una proteína en la imagen 2DGE se puede ver afectada por anomalías presentes en la imagen debido al proceso de separación de las proteínas y adquisición de la imagen. Por esta razón se hace necesario el uso de técnicas de procesamiento digital de imágenes que permitan detectar de una manera precisa y confiable las proteínas dentro de un gel (Luhn et al., 2003).

Existe una amplia variedad de investigaciones realizadas sobre el registro de imágenes 2DGE, principalmente enfocadas al problema de la selección automática de puntos de referencia para la posterior transformación de la imagen (algunos ejemplos son (Daszykowski, Bierzynska-Krzysik, Silberring, & Walczak, 2010; Kaczmarek, Walczak, Jong, & Vandeginste, 2002; Kaczmarek et al., 2003; Srinark & Kambhamettu, 2008)) proponen un algoritmo de identificación de puntos que es una integración de un método jerárquico que encuentra correspondencia entre pares de puntos.

(Morris, Clark, & Gutstein, 2008) proponen "Pinnacle", un método de detección de puntos y cuantificación. Elabora una imagen promedio sin ruido de un conjunto de imágenes 2DGE correctamente alineadas y esto se logra mediante la detección de los mínimos locales en la imagen y la combinación de la media sin ruido dentro de una proximidad definida. La principal ventaja de Pinnacle se encuentra sobre la detección de focos que se superponen. Sin embargo, ocasionalmente resultan en la detección puntos no esenciales, es decir, puntos de falsos positivos (Y. Wu & Zhang, 2011).

A pesar de los avances en el campo de registro y fusión de imágenes, a la fecha aún muchos investigadores seleccionan manualmente puntos de referencia para realizar el registro y la fusión de imágenes 2DGE (Li et al., 2011). Por lo tanto, surge la necesidad, desde hace mucho tiempo de desarrollar un procedimiento, junto con el software, que sea simple y robusto que facilite el proceso de fusión de imágenes 2DGE.

## **2.4 Análisis proteómico de apitoxina**

Existen pocos estudios proteómicos previos sobre la apitoxina. "Peiren et al.(2008) presentan un análisis proteómico del veneno de las abejas empleando electroforesis bidimensional en gel y espectrometría de masas". Este estudio reveló en los componentes de la apitoxina cuatro enzimas antioxidantes. Un total de 49 puntos extirpados fueron analizados por espectrometría de la masa; 39 de ellos dio lugar a la identificación de 6 proteínas ya conocidas del veneno de las abejas y 3 proteínas que no han sido descritos en las muestras que se habían tomado antes. La primera nueva proteína del veneno tiene un derivado de plaquetas y de crecimiento endotelial, la segunda proteína no muestra homología con ninguna proteína conocida y la tercera coincide con una hipotética proteína similar a la principal proteína de la jalea real. Sin embargo, no se encontró ningún estudio proteómico que permita caracterizar los patrones de veneno de abejas obreras del interior y del exterior de una colmena.

## 2.5 Discusión.

El procesamiento de imágenes se aplica para discriminar las manchas o ruidos de las proteínas reales, cuantificar las proteínas y sus niveles, poder ubicarlas en un espacio determinado. Se hace notorio la presencia de dos corrientes en el procesamiento de un conjunto de imágenes 2DGE: detección e identificación de proteínas de forma independiente y la realización de un registro de las imágenes alineadas, de tal manera que se obtenga una correspondencia punto a punto. El procesamiento normalmente se realiza en tres etapas: preprocesamiento, identificación de proteínas y fusión de imágenes. El primero busca mejorar la calidad de la imagen, aplicando técnicas de transformación y ajuste, de disminución de ruido, normalización de intensidades y corrección de fondo.

Al aplicar métodos de detección y segmentación, se busca identificar las proteínas, reduciendo los datos a pocos puntos característicos. La segmentación divide una imagen en sus objetos y partes constituyentes. Los métodos más comunes son: umbralización, crecimiento de regiones, detección de bordes y métodos morfológicos. El detector de bordes Prewitt es el que ofrece mayor inmunidad al ruido. El detector de bordes Sobel presenta mayor sensibilidad a los bordes diagonales que el de Prewitt, presentándose muy poca diferencia entre ellos. El detector de bordes Roberts tiene buena respuesta ante bordes diagonales, presenta mucha sensibilidad al ruido y es pobre en detección. En general los mayores problemas que enfrentan los detectores de bordes usados en segmentación, son la poca capacidad que tienen para diferenciar proteínas de anomalías y no diferenciar proteínas superpuestas. Para los procesos definidos en el proyecto, se utiliza el detector de bordes Sobel, que proporciona buenas cualidades de detección de puntos de interés en las imágenes 2DGE procesadas. La transformada watershed mejora los resultados de la segmentación de imágenes, pero presenta alta sensibilidad al ruido que da como resultado una sobre-segmentación. La umbralización no toma en cuenta las características espaciales de la imagen, causando sensibilidad al ruido o a intensidades no homogéneas. A veces, una deficiente o desigual iluminación en la imagen, hace que no sea conveniente el uso de un umbral global para toda la imagen. Una solución puede ser subdividir la imagen y encontrar un umbral apropiado para cada sub-imagen. Como el umbral usado en cada píxel depende de la localización del mismo,

esta umbralización se llama adaptativa o multiumbralización, la cual fue adaptada en el proyecto e implementada también en los procesos de multiumbralización para 4, 8 y 16 ventanas.

Por último, la etapa de fusión de imágenes permite realzar los elementos que no son visibles, complementar datos para una mejor clasificación, reemplazar datos anómalos y adicionar datos faltantes. En la revisión de literatura presentada se evidencia la necesidad de métodos automáticos para el procesamiento de imágenes 2-DGE. Este trabajo de tesis se centra en proponer una metodología para la identificación de patrones característicos desde múltiples geles usando técnicas automáticas de visión artificial.



### 3. Marco experimental

Las Figuras 3-1 y 3-2 muestran cada uno de los pasos seguidos en la metodología implementada para la identificación de objetos en las imágenes 2DGE y en la fusión de las mismas.

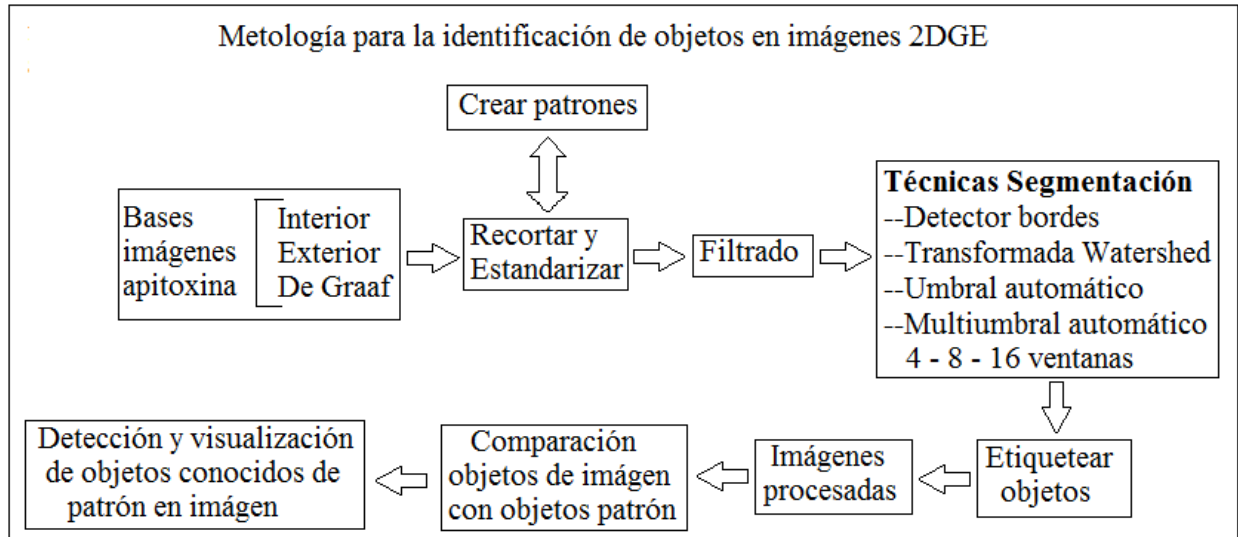


Figura 3-1: Metodología para la identificación de objetos en imágenes 2DGE

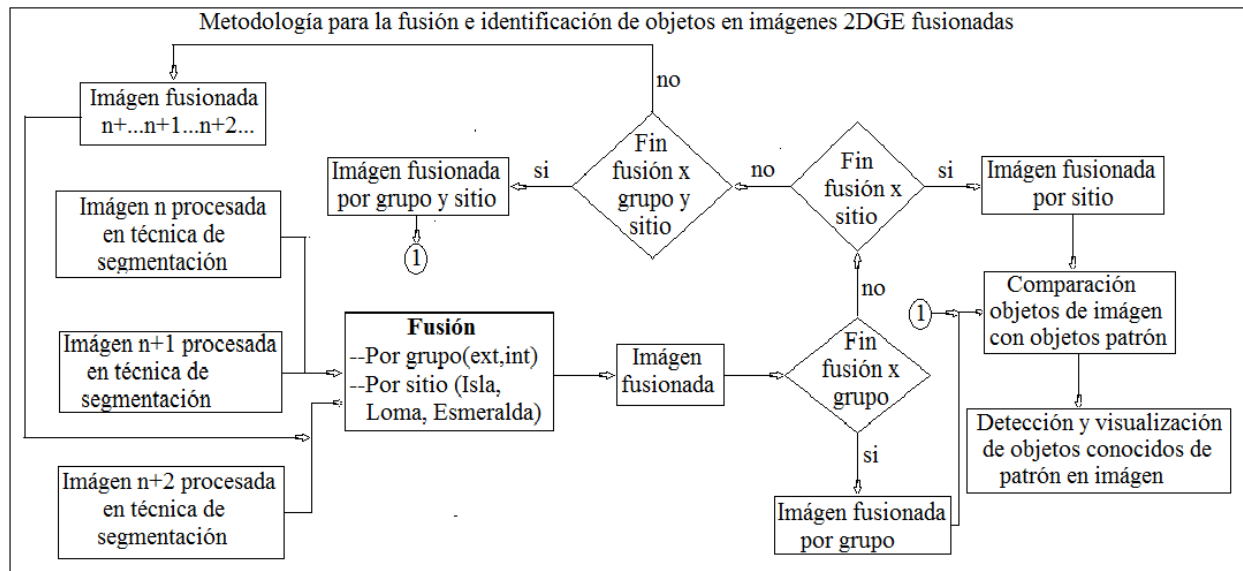


Figura 3-2: Metodología para la fusión e identificación de objetos en imágenes 2DGE fusionadas

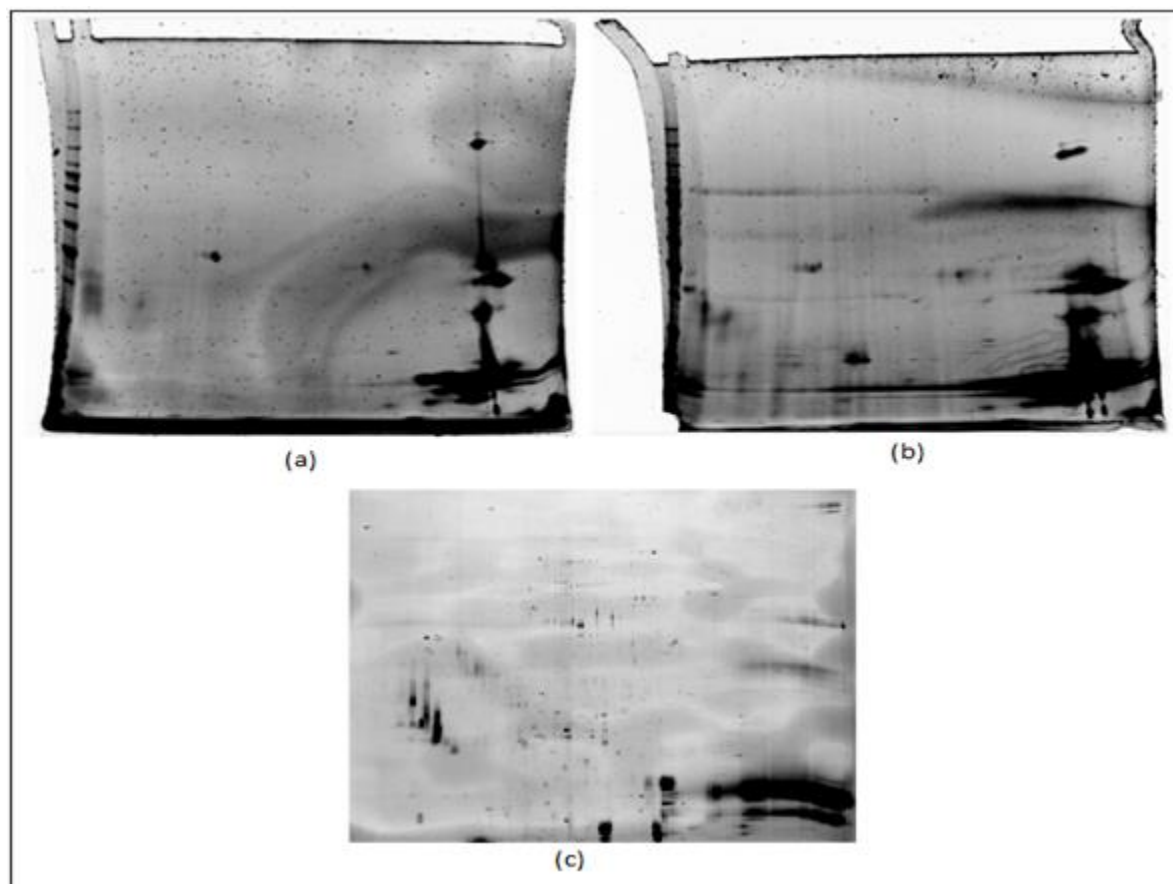
### 3.1 Base de Datos

Una de las bases de datos empleadas para el desarrollo de este proyecto incluye imágenes del veneno de las abejas (apitoxina) generadas en el laboratorio de Biología Celular y Molecular del ITM. La sección 3.1.1 describe en detalle la adquisición de estas imágenes. Adicionalmente, se usó una imagen de apitoxina facilitada por el Dr. Dirk de Graaf de la Universidad de Gante (Bélgica). De las dos bases de datos se conformaron dos patrones básicos en los procesos de comparación de técnicas de segmentación empleadas en este proyecto. El primer patrón se extrae de la imagen del veneno de abejas facilitada por el Dr. Dirk de Graaf, en donde se identifican 62 proteínas conocidas en toda la región de la imagen (ver Figura 3-4) y el segundo patrón se crea con la información contenida en la imagen master sintética compuesta de la sumatoria de todas las imágenes 2DGE obtenidas de las muestras de veneno de interior y de exterior extraídas de las fincas de las veredas La Esmeralda, La Isla y La Loma del municipio de Anserma Caldas, procesadas en el laboratorio de Ciencias Biológicas del Instituto Tecnológico Metropolitano y descritas en el trabajo de investigación: “Informe FGN008 Formato de Informe Final P14227”, páginas 27-28-29. (Röthlisberger, Torres y Delgado (2017) (ver Figuras 4 – 4; 4 – 5; 4 – 6; 4 – 7)

#### 3.1.1 Adquisición de imágenes apitoxina

En las fincas Medellincito vereda La Loma, Guadalcanal vereda La Esmeralda y vereda La Isla del municipio de Anserma Caldas, se tienen 78 colmenas, Se identificó con el apoyo de un apicultor experto las colonias pobladas con abejas de interior (obreras hasta los 21 días de edad) y abejas de exterior (obreras de los 22 días de edad en adelante) (De Felipe & Vandame 1999), (Mendizabal, 2005). Se empleó un extractor de apitoxina marca Ketitlen, el cual extrae el veneno de la abeja melífera (*Apis mellifera*), sin dañarlas o matarlas.

Se realizó una búsqueda en la literatura en cuanto a la extracción de proteínas de la apitoxina y se encontraron cuatro métodos distintos para este fin (Fujita et al., 2010; Li et al., 2011; Matysiak,



**Figura 3-3:** Imágenes de muestras de veneno de apitoxina: (a) abejas de exterior, (b) abejas de interior y (c) facilitada por Dr. Dirk de Graaf

Schmelzer, Neubert, & Kokot, 2011; Yiou et al., 2013; Zhang et al., 2012). El **Método 1** consiste en resuspender la apitoxina en Urea 7 M. Para esto, se pesaron 20 mg de apitoxina los cuales fueron resuspendidos en 200  $\mu$ l de Urea 7 M con agitación con vortex (1 min), sonicación en frío (10 min) y nuevamente agitación con vortex (1 min). El **Método 2** adaptado de (Fujita et al., 2010; Yiou et al., 2013) consiste en resuspender la apitoxina en buffer de lisis, precipitar con TCA y resuspender en Urea 7 M y CHAPS 4 %. Para esto se pesaron 20 mg de apitoxina los cuales fueron resuspendidos en 200  $\mu$ l de buffer de lisis (Urea 8 M, Tiourea 2 M, CHAPS (3-[(3-Cholamidopropyl) dimethylammonio]-1-propanesulfonate) al 4 %, Tris-base 20 mM, DTT (Dithiothreitol) 30 mM, coctel inhibidor de proteasa), seguido por agitación con vortex (1 min), sonicación en frío (10 min) y nuevamente agitación con vortex (1 min). El **Método 3**, adaptado de (Zhang et al., 2012; Li et al., 2011), consiste en resuspender la apitoxina en buffer de lisis, precipitar con acetona y resuspender en Urea 7 M y CHAPS 4 %. Para esto, se pesaron 20 mg de

apitoxina los cuales fueron resuspendidos en 200  $\mu$ l de buffer de lisis (Urea 8 M, Tiourea 2 M, CHAPS (3-[(3-Cholamidopropyl) dimethylammonio]-1-propanesulfonate) al 4 %, Tris-base 20 mM, DTT (Dithiothreitol) 30 mM, coctel inhibidor de proteasa), seguido por agitación con vortex (1 min), sonicación en frío (10 min) y nuevamente agitación con vortex (1 min). En el **Método 4**, descrito en (Matysiak et al., 2011), se resuspendieron 2 mg de apitoxina en 200  $\mu$ l de agua destilada con agitación con vortex (1 min), sonicación en frío (10 min) y nuevamente agitación con vortex (1 min). (extraído del trabajo principal: “*Caracterización del veneno de la abeja africanizada mediante análisis proteómico y procesamiento digital de imágenes orientado a la exploración de sus propiedades* Pag13). (Röthlisberger et al., (2017).

La tinción de proteínas se hace usando SYPRO® Ruby Protein Gel Stain, Las proteínas en los geles se visualizan usando luz UV y el gel que contiene las proteínas separadas espacialmente en sus dos dimensiones (nivel isoelectrico (ph) y peso molecular) mediante el proceso de electroforesis bidimensional, es sometido al proceso de escaneo y procesamiento digital, para la obtención final de las imágenes 2DGE. Las Figuras 4-4 y 4-5 presentan las 23 imágenes 2-DGE obtenidas desde la apitoxina. En total, se cuenta con 12 imágenes para abejas del interior: 5 imágenes de la vereda La Isla y 7 imágenes de la vereda La Loma. Se analizan 11 imágenes para abejas del exterior: 3 imágenes de la vereda La Esmeralda, 2 imágenes de la vereda La Isla y 6 imágenes de la vereda La Loma.

### 3.1.2 Imagen Dr. Dirk de Graaf

Para este trabajo también se usó una imagen 2-DGE de apitoxina facilitada por el Dr. Dirk de Graaf, uno de los autores del artículo “*Proteomic analysis of the honey bee worker venom gland focusing on the mechanisms of protection against tissue damage*” (Peiren et al., 2008), la cual es mostrada en la Figura 3-3(c). Para esta imagen adicionalmente se cuenta con proteínas identificadas, las cuales se muestran en la Figura 3-4. La imagen tiene una dimensión de 550 x 509 pixeles y en ella se detectaron aproximadamente 350 puntos de proteína, de los cuales fueron identificados 62 puntos dominantes mediante análisis de espectrometría de masas. Más de un tercio (23 de 62) corresponden a proteínas conocidas de veneno de abeja. Los restantes 39 puntos

representan 27 proteínas de la glándula de veneno y del tejido de depósito del mismo (Peiren et al., 2008).

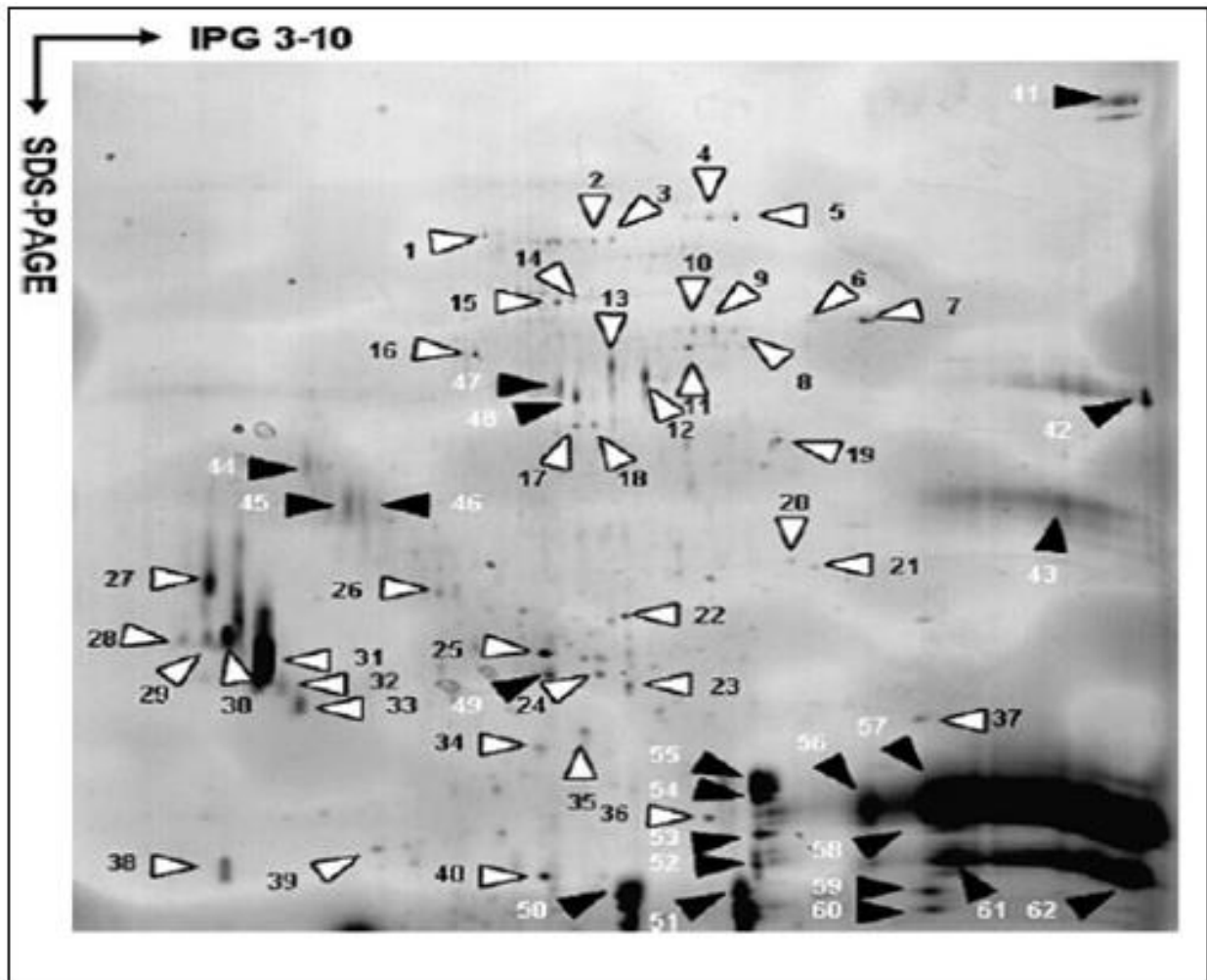


Figura 3-4: Imagen 2-DGE de Peiren et al. (2008).

### 3.2 Métodos de segmentación

Las imágenes originales: Muestras de apitoxina de Interior, de Exterior y Dr. Dirk de Graaf, fueron sometidas a los procesos de:

- Filtrado de imágenes
- Detector de bordes Sobel

- Umbralización automática
- Multiumbralización para 4, 8 y 16 ventanas filtradas.
- Transformada watershed
- Fronteras y etiquetas Boundary\_bwlabel

### 3.2.1 Filtrado de imágenes

En esta fase de experimentación, las imágenes 2DGE originales son sometidas a un proceso de filtrado, que se lleva a cabo utilizando una técnica de filtrado digital (median filter), filtro de difusión no lineal basado en operaciones no lineales sobre un cuadrante de píxeles determinado (3x3), que definen una respuesta para el punto central del cuadrante. Cada píxel de salida contiene el valor mediano en una vecindad 3 por 3 alrededor del píxel correspondiente en la imagen de entrada.

### 3.2.2 Detector de bordes Sobel

Para el procesamiento de imágenes, existen muchas clases de detectores, que son máscaras que se desplazan a través de una imagen calculando un operador local de derivación, tales como: el gradiente, Laplaciano, Roberts, Prewitt, Sobel, Kirsch, Robinson, entre otros (Gonzalez y Wood, 2002). El detector de bordes que más se ajustó a las imágenes 2DGE estudiadas fue el detector de bordes Sobel, que emplea la siguiente máscara:

$$G_{mx} = \begin{vmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{vmatrix} \text{ y } G_{my} = \begin{vmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{vmatrix}$$

Donde  $G_{mx}$  y  $G_{my}$ , corresponden al proceso de convolución realizado en la detección de bordes Sobel en el procesamiento de imágenes. La imagen resultante del proceso Detector de bordes Sobel es sometida al proceso de comparación de los objetos totales con los objetos conocidos del

patrón, obteniéndose una última imagen resultante con los objetos conocidos del patrón detectados y marcados en la imagen segmentada (ver Figura 4-3(b)).

### 3.2.3 Umbralización automática

La umbralización es uno de los métodos de similitud más comunes, importante y sencillo para la segmentación de imágenes. Permite la separación de objetos del fondo a partir de un umbral en la intensidad de los píxeles. En este método se empleó el umbral de imagen global utilizando el método Otsu, que calcula un umbral general (nivel) que puede ser usado para convertir la intensidad de imagen en imagen binaria. El nivel es un valor de intensidad normalizado que se encuentra en el intervalo  $[0,1]$ . Con el método de Otsu, se elige el umbral para minimizar la varianza entre clases de los píxeles en blanco y negro (Otsu, 1979).

Inicialmente se realiza en el procesamiento de las imágenes originales Dr. Dirk de Graaf y de la base de apitoxina, un filtrado en ventanas de  $3 \times 3$  y luego son sometidas al proceso umbralización automática método Otsu, el cual elige automáticamente el umbral minimizando la varianza entre clases de los píxeles en blanco y negro. Se usó la función `graythresh` de Matlab, que devuelve la métrica global de umbralización `Level` y de efectividad `EM`, como el segundo argumento de salida. La métrica de efectividad es un valor en el rango  $[0, 1]$  que indica la efectividad del umbral de la imagen de entrada. Un valor de nivel 0.5 es intermedio entre el blanco y el negro. El límite inferior es alcanzable sólo por imágenes que tienen un solo nivel de gris, y el límite superior es alcanzable sólo por imágenes de dos valores (Otsu, 1979). Realizando pruebas con los umbrales automáticos de `Level` y `EM` y con el complemento de cada uno de ellos ( $1-\text{Level}$  y  $1-\text{EM}$ ), calculados para las imágenes de entrada, se eligió realizar los procesos de umbralización y multiumbralización con el umbral  $1-\text{EM}$ , que fue el que mejores resultados mostró en la visualización de la segmentación de objetos, que para el caso convierte los píxeles mayores o iguales que el umbral complemento en 1 (blanco) o en caso contrario lo convierte en 0 (negro).

La imagen resultante del proceso umbralización automática es sometida al proceso de comparación de los objetos totales con los objetos conocidos del patrón, obteniéndose una última

imagen resultante con los objetos conocidos detectados del patrón y marcados en la imagen segmentada (ver Figura 4-3(a)).

### 3.2.4 Multiumbralización automática para 4, 8 y 16 ventanas filtradas.

En procesamiento de imágenes 2DGE se encuentran enfoques multiumbral para la detección de las proteínas (Sengar et al., 2016; Kostopoulou et al., 2015b; Zacharia et al., 2013). Por ejemplo, Kostopoulou et al. (2015c) divide la imagen en ventanas de tamaño fijo, en cada una de las cuales se aplica el proceso de umbralización. Para este proyecto, las imágenes originales fueron sometidas a un proceso de multiumbralización automática, consistente en la división de las imágenes en forma aleatoria en 4, 8 y 16 ventanas de tamaño fijo, con el fin de tener comparaciones por ventanas de cada par de imágenes, fusionar estas ventanas y obtener una nueva imagen fusionada con iguales o mejores características que con la fusión de pares de imágenes completas. Tomando el tamaño estándar de las imágenes de apitoxina de  $1020 \times 1364$  *pixeles*, la división en 4 ventanas tiene tamaños por ventana de  $510 \times 682$  *pixeles*, para 8 ventanas los tamaños son de  $510 \times 341$  *pixeles* por ventana y de  $255 \times 341$  *pixeles* por ventana para una división en 16 ventanas. Cada ventana es sometida al proceso de umbralización automática con los diferentes controles (level, EM) y sus complementos, obteniéndose 4 – 8 y 16 ventanas multiumbralizadas, que son integradas en una imagen resultante multiumbralizada. Como ejemplo, los valores Level y EM calculados automáticamente para la imagen de la muestra interior Isla 2 bastante oscura (MII2BOX), dividida en 4 ventanas son: Ventana 1. Level: 0.4431 equivalente a valor en *pixeles* en escala 255 de 113, Level complemento: 0.5569; EM: 0.6903; EM complemento: 0.3097. Ventana 2. Level: 0.6706; Level complemento: 0.3294; EM: 0.4745; EM complemento: 0.5255. Ventana 3. Level: 0.3216; Level complemento: 0.6784; EM: 0.5921; EM complemento: 0.4079. Ventana 4. Level: 0.2667; Level complemento: 0.7333; EM: 0.8301; EM complemento: 0.1699. Ver división en 4 ventanas de la imagen MII2BOX en la Figura 4-10.

Adicionalmente a este proceso se le agrega el filtrado y conteo de objetos de cada una de las ventanas originales y de las ventanas resultantes del proceso de multiumbralización, obteniendo otra imagen resultante filtrada. Las imágenes resultantes del proceso multiumbralización



automática para 4 – 8 y 16 ventanas son sometidas al proceso de comparación de los objetos totales con los objetos conocidos de cada uno de los dos patrones (ver Tablas 4-1 y 4-6), obteniéndose una última imagen resultante con los objetos conocidos del patrón detectados y marcados en la imagen segmentada (ver Figura 4-3(d-e-f)).

### 3.2.5 Transformada Watershed.

Fue introducida inicialmente en morfología matemática y es empleada en la solución de problemas de segmentación en imágenes en escala de grises. Presenta una alta sensibilidad al ruido que resulta en una sobre segmentación. Para solucionar este problema se combina con otras técnicas como la umbralización, la unión y separación de regiones.

La transformada watershed encuentra "cuencas de captación" o "líneas de crestas de cuenca" en una imagen tratándola como una superficie donde los píxeles de luz representan alturas elevadas y los píxeles oscuros representan bajas elevaciones. Los elementos de la matriz resultante de la imagen procesada con esta técnica, tienen valores enteros mayores o iguales a 0. Los elementos etiquetados con 0 no pertenecen a una región de cuenca única.

La segmentación utilizando la transformada watershed funciona mejor si puede identificar, o "marcar", los objetos de primer plano y las ubicaciones de fondo. Los pasos que se siguieron para este proceso fueron: Leer la imagen en color y convertirla a escala de grises; utilizar la magnitud de gradiente como la función de segmentación (se utiliza la máscara de borde Sobel), se filtra el arreglo multidimensional y se calcula la magnitud del gradiente mediante la raíz cuadrada de la suma de los cuadrados de las coordenadas  $(x, y)$  y se procede a segmentar la imagen aplicando seguidamente a esta magnitud gradiente la transformada watershed. El gradiente es alto en los bordes de los objetos y bajo en su mayoría dentro de los objetos. Para evitar sobre-segmentación se procede a marcar los objetos de primer plano (que deben ser conjuntos de píxeles conectados dentro de cada uno de los objetos de primer plano), utilizando técnicas morfológicas llamadas "apertura por reconstrucción" y "cierre por reconstrucción". La apertura es una erosión seguida de una dilatación, mientras que el cierre es una erosión seguida de una reconstrucción morfológica. Los anteriores procesos de apertura y cierre ayudan a quitar algunas manchas. Se calculan

marcadores de fondo, en donde los píxeles oscuros pertenecen al fondo. Por último, se procede a marcar los píxeles oscuros que pertenecen al fondo utilizando una función de umbral. La imagen resultante del proceso de watershed es sometida al proceso de comparación de los objetos totales con los objetos conocidos del patrón, obteniéndose una última imagen resultante con los objetos conocidos del patrón detectados y marcados en la imagen segmentada (ver Figura 4-3(c)).

### 3.2.6 Fronteras y etiquetas `boundary` y `bwlabel`

En los procesos de experimentación se utiliza la función de matlab `bwboundaries` para mostrar las fronteras de los objetos en la imagen y trazar los límites exteriores de los objetos, así como los límites de los agujeros dentro de estos objetos, en una imagen binaria. `Bwboundaries` también descende al interior de los objetos más externos (padres) y traza a sus hijos (objetos completamente encerrados por los padres) y devuelve una matriz de celdas de ubicaciones de píxeles límites. Si se utilizan parámetros adicionales, se puede especificar la conectividad a utilizar al trazar los límites del padre y del hijo, o incluir los límites de los agujeros dentro de otros objetos. Para etiquetar los componentes conectados en una imagen binaria bidimensional se utiliza `bwlabel` para una conectividad de 8 y 4 componentes. En todos los procedimientos de comparación de las técnicas de segmentación del proyecto se empleó la conectividad de 4 componentes, que es la que mejor define la forma ovoide de las proteínas.

### 3.2.7 Evaluación de los métodos de segmentación.

Para la evaluación de los métodos de segmentación, se recopiló la información básica en dos patrones de referencia, en donde se identifica la ubicación de las proteínas contenidas en la imagen del Dr. Dirk De Graaf (62 proteínas conocidas) y de las imágenes proporcionadas en el trabajo “Informe FGN008 Formato de Informe Final P14227” (115 proteínas conocidas). Los dos patrones básicos se utilizan para realizar la comparación de las proteínas contenidas en las imágenes procesadas en las diferentes técnicas de segmentación empleadas: umbralización automática, multiumbralización automática, detección de bordes y transformada Watershed. Todas las 24

imágenes (1 del patrón del Dr- Dirk de Graaf y 23 del patrón de apitoxina) son estandarizadas en un tamaño igual al tamaño de las imágenes de donde se extrajeron las proteínas del patrón, habiéndoles realizado un proceso de recorte y de coincidencia de al menos un punto de referencia con las coordenadas del patrón. Se procede a realizar el filtrado, se somete cada imagen a cada una de las técnicas de segmentación y se contabilizan (etiquetan) los objetos para cada una de ellas. El resultado de las imágenes procesadas es sometido a la comparación con los objetos conocidos del patrón, obteniéndose valores calculados para: los objetos totales y los objetos conocidos detectados con propiedades (área, centroide y `boundingbox`), la relación porcentual de objetos conocidos del patrón relacionados con los objetos totales con propiedades, la relación porcentual de objetos conocidos detectados relacionados con los objetos totales con propiedades (precisión) y la relación porcentual de los objetos conocidos detectados relacionados con los objetos del patrón (sensibilidad). En el proceso final se realiza la detección y visualización para cada imagen o su fusión de los objetos conocidos detectados del patrón de referencia (ver metodología Figuras 3-1 y 3-2). Para los anteriores indicadores, se toma como valor de referencia de análisis más significativo el indicador porcentual de la relación de objetos detectados con respecto a los objetos conocidos del patrón (sensibilidad (S%)) (ver Tablas 4-2 y 4-11).

### 3.3 Fusión de imágenes

Las imágenes resultantes de la aplicación de cada una de las técnicas de segmentación descritas en el proyecto son sometidas al proceso de fusión que se lleva a cabo mediante la unión de los objetos de cada par de imágenes. En el proceso de fusión se extraen para cada una de las imágenes los gradientes geométricos por: dilatación, erosión y morfológico. Se realiza el barrido de píxeles en cada una de las imágenes con gradientes y sin cálculo de gradientes y para cada par de píxeles se calcula el promedio de la sumatoria de los valores que están ubicados espacialmente en las mismas coordenadas de ambas imágenes comparadas y cuyos valores resultante de la unión se lleva a los píxeles que conforman la nueva imagen fusionada. La imagen resultante fusionada es sometida nuevamente a procesos de fusión con el resto de las imágenes de la base, hasta realizar todas las fusiones deseadas por grupo, por sitio y sus combinaciones. Las imágenes finales fusionadas son sometidas al proceso de comparación de objetos de la misma imagen con los objetos conocidos

del patrón de referencia, realizándose finalmente la detección y visualización de los objetos conocidos detectados del patrón dentro de la imagen procesada (ver Figuras 3-2 (metodología para la fusión) y 4-34 (Consolidado de la fusión de todas las imágenes de interior y exterior)).

### 3.4 Contribución.

En los procesos de multiumbralización automática, en los cuales se divide cada imagen en 4, 8 y 16 ventanas, se ejecuta la fusión de cada par de imágenes, tanto en forma completa (fusión imagen 1 con imagen 2), como también la fusión de cada par de ventanas (fusión de ventana 1 de la imagen 1 con la ventana 2 de la imagen 2) realizándose al final la unión de cada ventana fusionada conformándose la imagen completa fusionada. Un ejemplo de este proceso es mostrado en la Figura 4-23, en donde se fusionan dos muestras de imágenes: muestra Isla Interior 2 bastante oscura estandarizada ((a) MII2BOX) y muestra Loma Interior 1 bastante oscura estandarizada ((g) MLI1BOX). Las imágenes b y h corresponden a MII2BOX y MLI1BOX después de haber sido sometidas al proceso de multiumbralización 4 ventanas y en cada una de las muestras anteriores se ejecuta la división en 4 ventanas (imágenes c, d, e, f para ventanas 1, 2, 3, 4 de MII2BOX e imágenes i, j, k, l para ventanas 1, 2, 3, 4 de MLI1BOX). Las imágenes m, n, o, p corresponden a la fusión llevada a cabo con cada par de ventanas multiumbralizadas correspondientes de las muestras MII2BOX y MLI1BOX.

## 4. Resultados

Para la comparación de las técnicas de segmentación se utilizó las dos bases de datos descritas en la sección 3.1. En las proteínas detectadas desde los métodos de segmentación y en los patrones de las bases de datos, se miden tres características espaciales por proteína: área, centro de masa y el `boundingbox` (caja contenedora de menor área que encierra cada una de las proteínas). Posteriormente, se realiza el análisis de las imágenes 2DGE de abejas de interior y exterior, a partir del estudio de la fusión de muestras para cada grupo de abejas; las imágenes fusionadas se comparan con el patrón descrito en la sección 3.1.1.

### 4.1 Resultados comparación imágenes Dr. Dirk de Graaf

Todas las imágenes de resultados del Dr. Dirk de Graaf y procesadas en las diferentes técnicas son redimensionadas a un tamaño de  $656 \times 790$  *pixeles*, esto con el fin de poder comparar la imagen con el patrón. En el proceso de comparación para hallar la ubicación de las proteínas, se define una aproximación espacial de  $\pm 10$  *pixeles* a partir de las coordenadas  $(x, y)$  referenciados sobre el centro de masa de las proteínas detectadas desde cada técnica de segmentación aplicada. Las proteínas detectadas son mostradas con la marca (\*) sobre la imagen resultante.

La Tabla 4-1 muestra las coordenadas de las 62 proteínas identificadas en la imagen del patrón Dr. Dirk de Graaf (Figura 3-4) con las cuales se realizan los procesos de comparación de las diferentes técnicas de segmentación, que incluyen: umbralización, multiumbralización automática, detección de bordes y transformada Watershed.

**Tabla 4-1:** Coordenadas (x, y) para las 62 proteínas en el patrón de la imagen 2DGE de apitoxina Dr. Dirk de Graaf

Spot	(x, y)	Spot	(x, y)	Spot	(x, y)	Spot	(x, y)
1	(297,133)	17	(362,275)	33	(164,485)	49	(343,461)
2	(375,136)	18	(375,274)	34	(336,517)	50	(401,635)
3	(388,133)	19	(506,285)	35	(368,504)	51	(486,636)
4	(458,120)	20	(515,376)	36	(458,570)	52	(495,602)
5	(476,120)	21	(532,380)	37	(608,495)	53	(493,582)
6	(528,194)	22	(397,418)	38	(111,608)	54	(498,558)
7	(570,195)	23	(400,471)	39	(220,593)	55	(491,539)
8	(474,201)	24	(379,462)	40	(340,613)	56	(570,561)
9	(460,203)	25	(340,446)	41	(751,31)	57	(689,579)
10	(446,204)	26	(265,400)	42	(766,253)	58	(615,567)
11	(443,216)	27	(100,389)	43	(711,335)	59	(613,624)
12	(411,237)	28	(81,436)	44	(170,305)	60	(618,646)
13	(387,229)	29	(99,435)	45	(199,334)	61	(631,600)
14	(360,180)	30	(116,424)	46	(209,333)	62	(758,612)
15	(349,182)	31	(139,445)	47	(351,245)		
16	(290,221)	32	(151,470)	48	(362,252)		

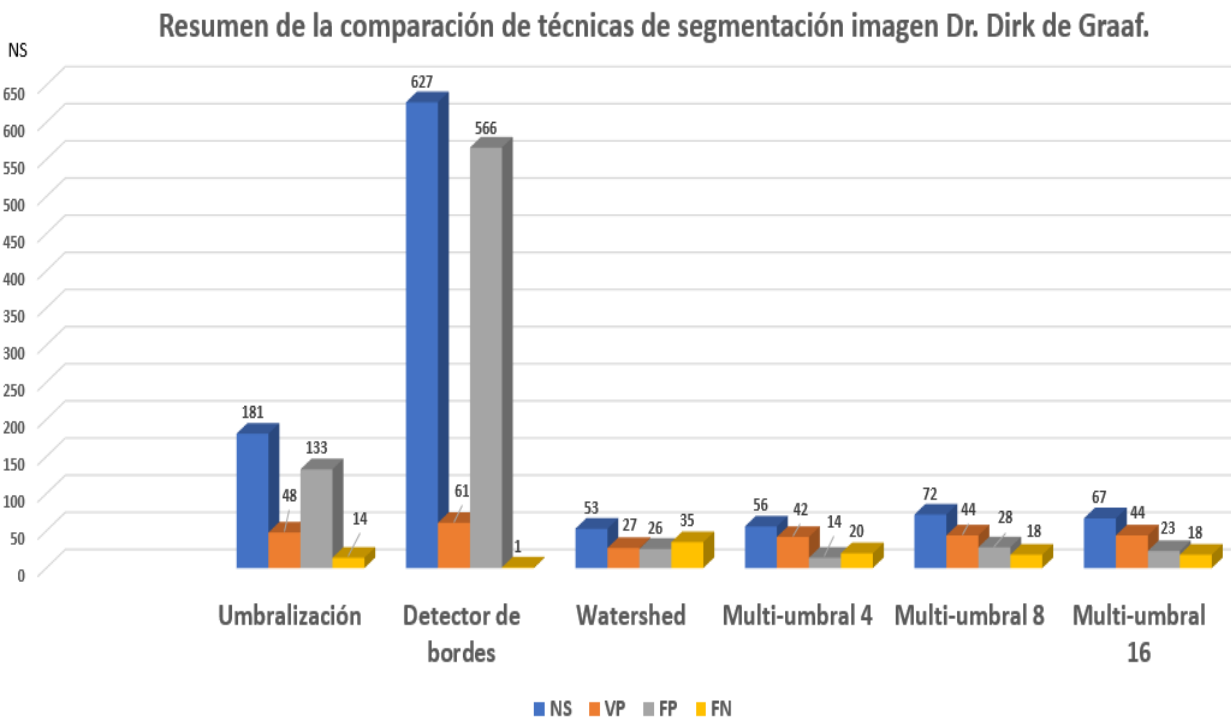
La Tabla 4-2 presenta una comparación de las proteínas conocidas y los objetos segmentados empleando técnicas de umbralización automática, detector de bordes, transformada Watershed y usando multiumbralización con ventanas de 4, 8 y 16. En esta tabla se presenta el número total de objetos segmentados (NS), el número de objetos segmentados que coinciden con proteínas conocidas (verdaderos positivos - VP), el número de objetos segmentados que no corresponden a proteínas conocidas (falsos positivos – FP), y el número de proteínas conocidas que no fueron segmentadas (falsos negativos – FN). Para evaluar la capacidad de detección se presenta la precisión ( $P = VP/NS$ ), la relación de falsos positivos ( $RF = FP/NS$ ), la relación de falsos negativos ( $RFN = FN/62$ ) y la sensibilidad del detector o relación de verdaderos positivos ( $S = VP/62$ ).

De acuerdo con la Tabla 4-2 y la Figura 4-2, los métodos de multiumbralización automática con 4, 8 y 16 ventanas logran una mayor precisión que los otros métodos comparados (ver Figura 4-2). Para el caso de 4 ventanas se obtiene el 75% de precisión. Por su parte la menor precisión se obtiene con el método de segmentación basado en detección de bordes, el cual solo logra un 9.7% de precisión dada la alta sobre-segmentación de la imagen que contiene 627 objetos (ver Figura 4-1). El método de Watershed presenta la más alta relación de omisión de proteínas con un porcentaje de relación de falsos negativos del 56.5% y es el que presenta menor sensibilidad y menor cantidad

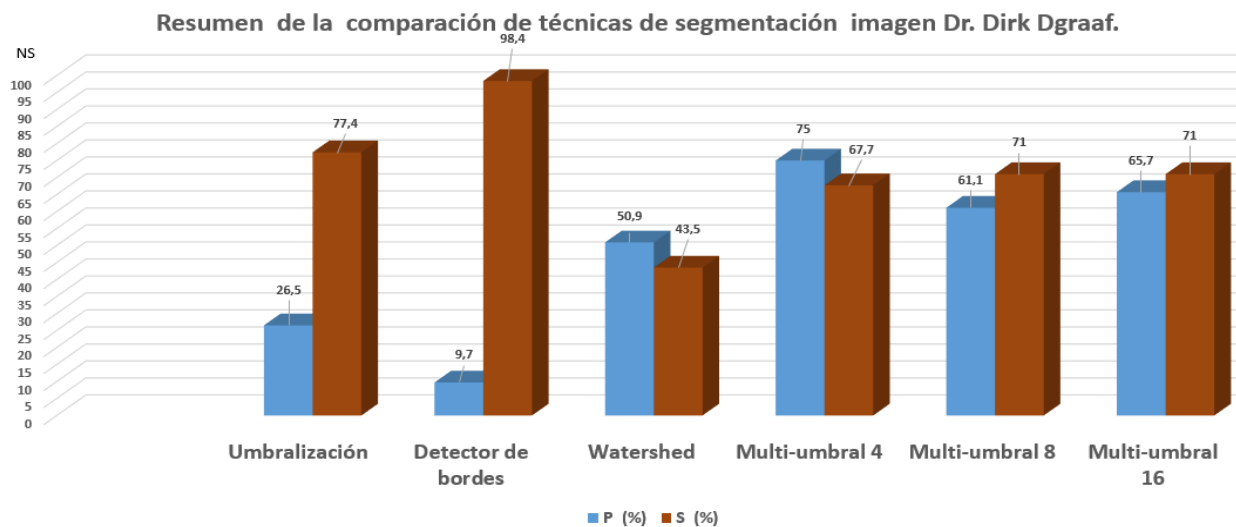
de objetos segmentados. Para los métodos de multiumbralización el porcentaje de la relación de falsos negativos es cercano al 30% para las tres configuraciones de ventanas empleadas. La Figura 4-3 presenta las imágenes segmentadas obtenidas desde cada técnica de segmentación; las marcas indican los puntos segmentados que corresponden a proteínas conocidas.

**Tabla 4-2:** Resumen de comparación de técnicas de segmentación con respecto al número de objetos conocidos en la imagen Dr. Dirk Dgraaf.

Técnica de Segmentación	NS	VP	FP	FN	P (%)	RF (%)	RFN (%)	S (%)
Umbralización	181	48	133	14	26.5	73.5	22.6	77.4
Detector de bordes	627	61	566	1	9.7	90.3	1.6	<b>98.4</b>
Watershed	53	27	26	35	50.9	49.1	56.5	43.5
Multiumbral 4	56	42	14	20	<b>75.0</b>	25.0	32.3	67.7
Multiumbral 8	72	44	28	18	61.1	38.9	29.0	71.0
Multiumbral 16	67	44	23	18	65.7	34.3	29.0	71.0



**Figura 4-1:** Relaciones de objetos segmentados (NS), proteínas de patrón detectadas (VP), objetos segmentados no reconocidos (FP), proteínas de patrón no detectadas en técnicas de segmentación (FN) imagen Dr. Dirk de Graaf.



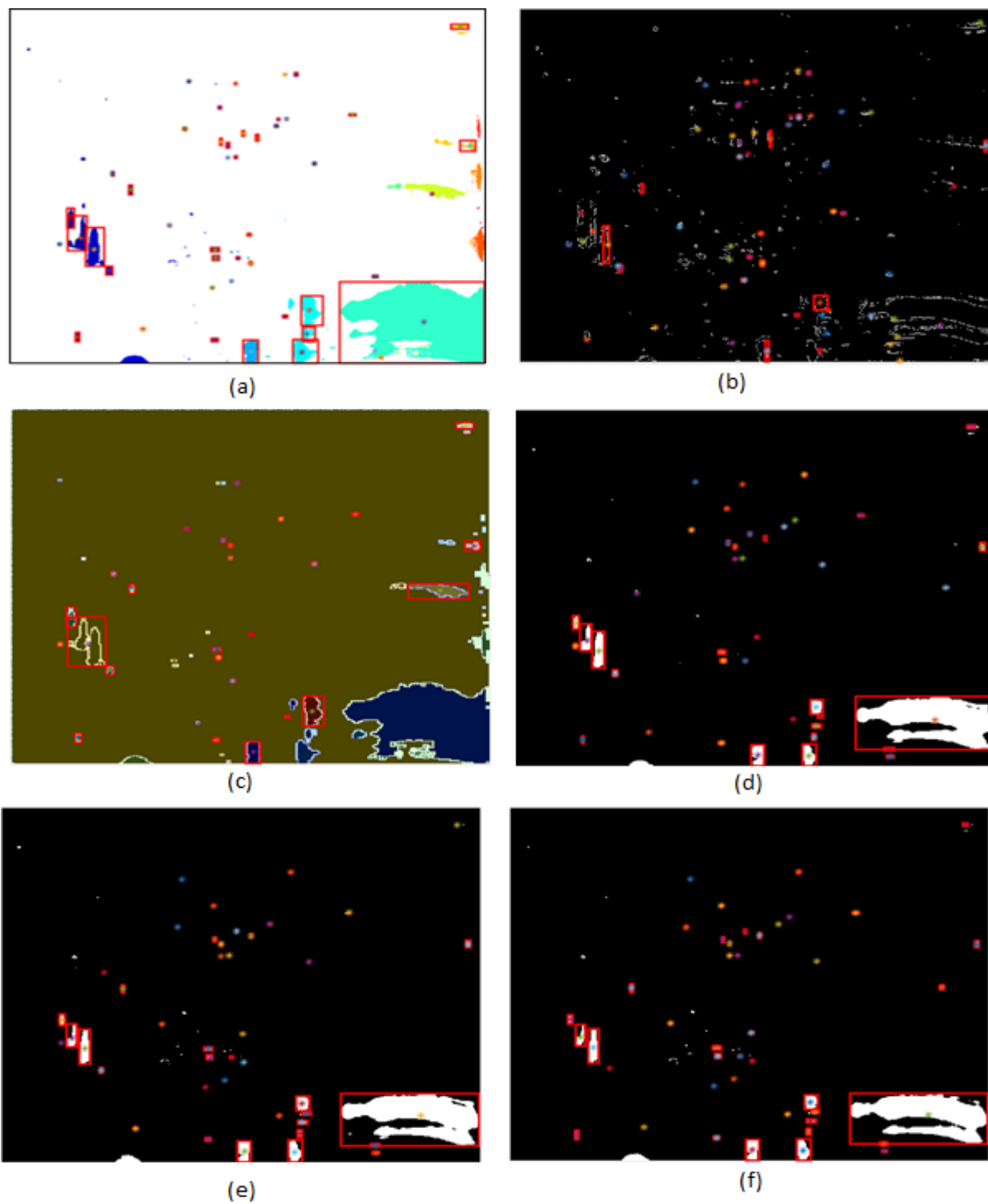
**Figura 4-2:** Relación de precisión (P) y sensibilidad (S) en la comparación de técnicas de segmentación imagen Dr. Dirk Dgraaf.

La Tabla 4-3 presenta las coordenadas del centro de masa para los 67 objetos segmentados (NS) en la imagen Dr. Dirk de Graaf con la técnica multiumbral con 16 ventanas.

**Tabla 4-3:** Coordenadas (x, y) de los 67 objetos con propiedades (NS) encontradas en el proceso para la técnica multiumbral 16 ventanas.

Spot	(x, y)	Spot	(x, y)	Spot	(x, y)	Spot	(x, y)
1	(30,73)	18	(291,222)	35	(376,275)	52	(494,569)
2	(29,75)	19	(291,442)	36	(379, 450)	53	(504,566)
3	(100,394)	20	(296,458)	37	(380,463)	54	(507,286)
4	(111,609)	21	(298,134)	38	(388,230)	55	(691,572)
5	(117,427)	22	(302,380)	39	(401,636)	56	(571,196)
6	(121,278)	23	(301,458)	40	(398,419)	57	(617,626)
7	(138,446)	24	(325,415)	41	(397,462)	58	(619,639)
8	(159,170)	25	(341,447)	42	(401,472)	59	(713,334)
9	(165,485)	26	(337,518)	43	(412,238)	60	(752,32)
10	(207,651)	27	(341,614)	44	(444,217)	61	(753,43)
11	(201,335)	28	(344,462)	45	(458,571)	62	(756,546)
12	(221,594)	29	(350,183)	46	(462,204)	63	(759,545)
13	(266,469)	30	(352,246)	47	(483,638)	64	(762,547)
14	(265,401)	31	(363,253)	48	(477,120)	65	(763,32)
15	(274,468)	32	(363,276)	49	(496,548)	66	(771,548)
16	(276,475)	33	(369,451)	50	(497,583)	67	(771,255)
17	(278,470)	34	(370,505)	51	(492,603)		





**Figura 4-3:** Marcado de objetos conocidos en las imágenes del Dr Dirk de Graaf empleando las técnicas descritas en la Tabla 4-2. (a) Umbralización automática; (b) Detector de bordes Sobel; (c) Transformada Watershed; (d) Multiumbral 4 ventanas; (e) Multiumbral 8 ventanas; (f) Multiumbral 16 ventanas.

**Tabla 4-4:** Características espaciales de los 44 objetos conocidos detectados (VP) en el procesamiento de la técnica multiumbral 16 ventanas en comparación con los 62 spots del patrón de apitoxina Dr. Dirk de Graaf.

Spot	Spot Dr Graaf	(x, y)	Spot	Spot Dr Graaf	(x, y)	Spot	Spot Dr Graaf	(x, y)
1	1	(298,134)	16	25	(341,447)	31	45	(201,335)
2	5	(477,120)	17	26	(265,401)	32	46	(201,335)
3	7	(571,196)	18	27	(100,394)	33	47	(352,246)
4	9	(462,204)	19	30	(117,427)	34	48	(363,253)
5	11	(444,216)	20	31	(138,446)	35	49	(344,462)
6	12	(413,238)	21	33	(164,485)	36	50	(401,636)
7	13	(388,230)	22	34	(337,518)	37	51	(483,638)
8	15	(350,183)	23	35	(370,505)	38	52	(492,603)
9	16	(291,222)	24	35	(458,571)	39	53	(497,583)
10	17	(363,276)	25	38	(111,609)	40	54	(504,566)
11	18	(376,275)	26	39	(221,594)	41	54	(496,548)
12	19	(507,286)	27	40	(341,614)	42	57	(691,572)
13	22	(398,419)	28	41	(752,32)	43	59	(617,626)
14	23	(401,472)	29	42	(771,255)	44	60	(619,639)
15	24	(380,463)	30	43	(713,334)			

### 4.1.1 Discusión

Aunque el mejor resultado de acuerdo con la sensibilidad se obtiene con la técnica detector de bordes (98.4%), se evidencia que no es la mejor técnica para la segmentación y detección de proteínas en imágenes 2DGE. La detección de bordes obtiene una alta sobre-segmentación que conlleva un alto valor de objetos falsos positivos (566) y una baja precisión (9.7%). Por su parte, la técnica watershed obtiene la más baja sensibilidad (43.5%) y porcentajes cercanos al 50% de precisión y relación de falsos positivos.

Se destaca que los resultados dependen del tipo de imagen que se esté analizando. Por ejemplo, para la imagen Dr. Dirk de Graaf, el mejor resultado de acuerdo con la sensibilidad o al número de objetos detectados se obtiene con el detector de Sobel (98.4%). Pero analizando la consistencia en la forma de la proteína, se observa el efecto del ruido en este detector que introduce falsos positivos al proceso (ver Figura 4-3(b)). Lo mismo sucede con la transformada Watershed, que en

el análisis comparativo con la imagen Dr. Dirk de Graaf arrojó los resultados más bajos, debido al menor número de objetos segmentados (NS) y que para otros tipos de imágenes puede presentar un mejor comportamiento. Visualmente se observa que las otras técnicas definen mejor las proteínas, sobre todo el método de umbralización automática, también con un alto porcentaje de sensibilidad (77.4%), pero con el segundo menor valor en precisión (26.5%) debido a la cantidad de objetos segmentados (181). En las técnicas de multiumbralización los resultados muestran que el mejor desempeño no está relacionado con el mayor número de ventanas, sino que las ventanas seleccionadas contengan la mayor cantidad de grupos de proteínas comunes. En el análisis comparativo relacionado con la sensibilidad, el mejor comportamiento se presentó en las técnicas para 8 y 16 ventanas con un 71%. En general con la identificación de propiedades espaciales como el centro de masa, podemos relacionar múltiples imágenes.

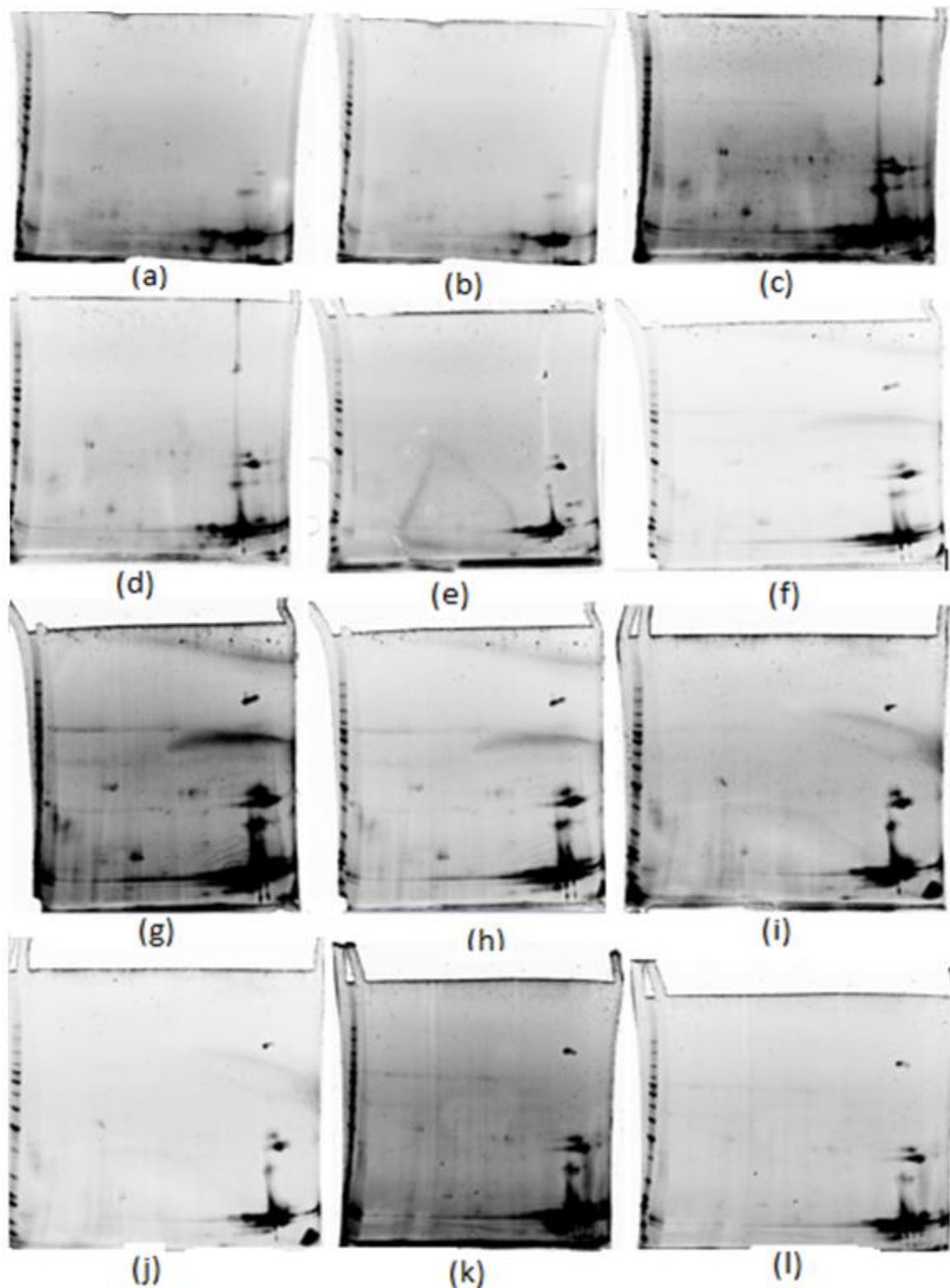
## **4.2 Resultados comparación imágenes apitoxina.**

Debido a la diferencia en las dimensiones de las muestras de imágenes procesadas en el laboratorio de Ciencias Biológicas del Instituto Tecnológico Metropolitano, todas las imágenes de apitoxina de esta base de datos son redimensionadas en un tamaño estándar de  $1020 \times 1364$  *pixeles*, tamaño calculado sobre el promedio dado por la sumatoria de las filas y de las columnas de todas las imágenes estudiadas. El proceso trabaja una aproximación espacial de  $\pm 10$  *pixeles* para los puntos de coordenadas  $(x, y)$  referenciados sobre el centro de masa de las proteínas y realiza las respectivas marcas de proteínas detectadas sobre la imagen resultante según la técnica aplicada. La Tabla 4-5 describe los nombres y tamaños de las 23 imágenes trabajadas en los diferentes procesos de comparación (12 imágenes para el veneno de las abejas de interior y 11 para las de exterior).

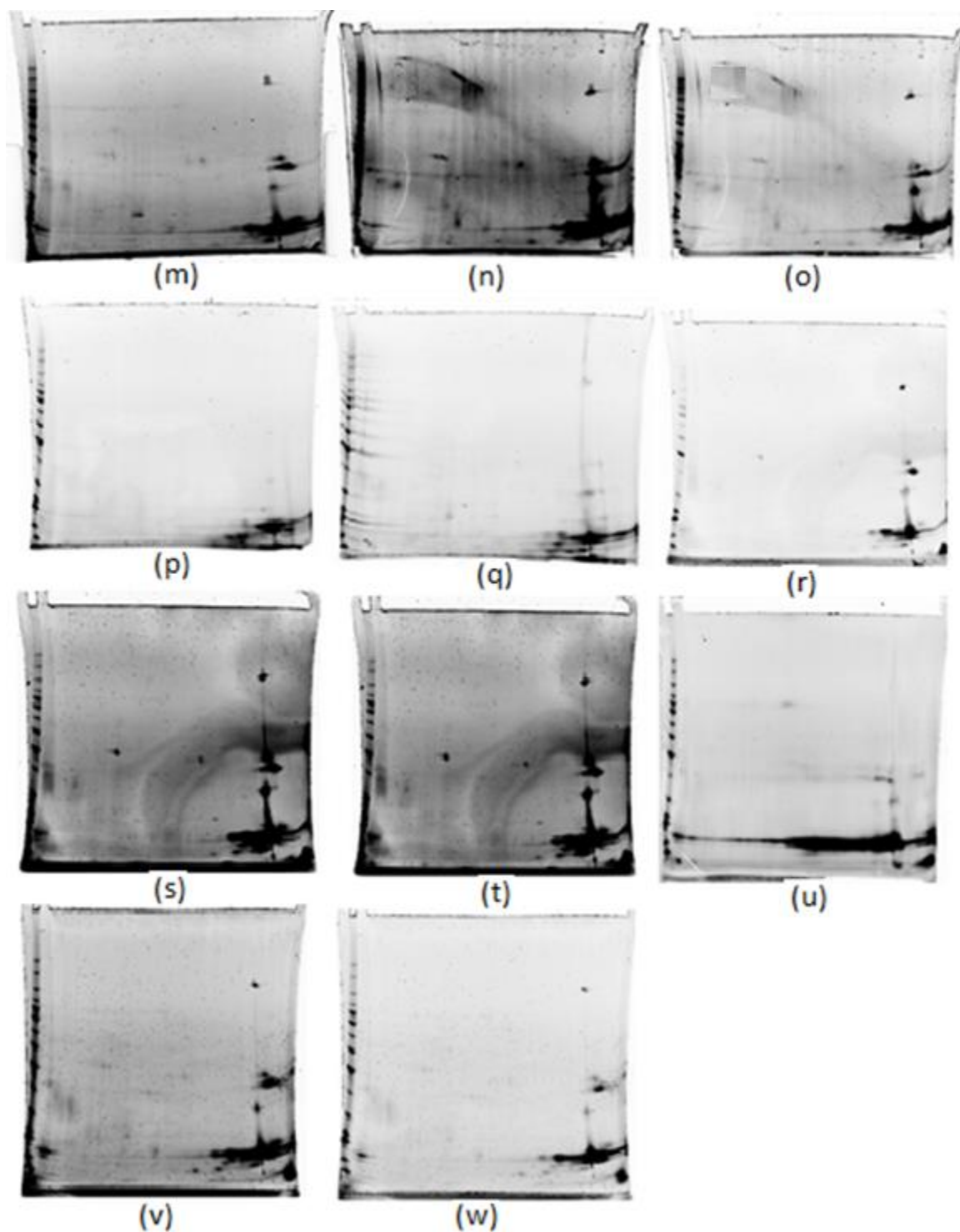
**Tabla 4-5:** Descripción y dimensiones de las imágenes originales del veneno de las abejas.

<b>Imágenes de muestras de veneno de abejas de Interior</b>				
<b>Sec</b>	<b>Nral</b>	<b>Nombre de imagen</b>	<b>Descripción de la imagen</b>	<b>Dimensiones</b>
1	a	MII1DO	Muestra Isla Interior 1 demasiado oscuro	981x1314x4
2	b	MII1SO	Muestra Isla Interior 1 super oscuro	981x1314x4
3	c	MII2BO	Muestra Isla Interior 2 bastante oscuro	981x1314x4
4	d	MII2DO	Muestra Isla Interior 2 demasiado oscuro	981x1314x4
5	e	MII3BO	Muestra Isla Interior 3 bastante oscuro	981x1314x4
6	f	MLI1O	Muestra Loma Interior 1 oscuro	981x1314x4
7	g	MLI1BO	Muestra Loma Interior 1 bastante oscuro	981x1314x4
8	h	MLI1DO	Muestra Loma Interior 1 demasiado oscuro	981x1314x4
9	i	MLI2BO	Muestra Loma Interior 2 bastante oscuro	1023x1370x4
10	j	MLI2DO	Muestra Loma Interior 2 demasiado oscuro	1023x1370x4
11	k	MLI3BO	Muestra Loma Interior 3 bastante oscuro	1067x1429x4
12	l	MLI3DO	Muestra Loma Interior 3 demasiado oscuro	1067x1429x4
<b>Imágenes de muestras de veneno de abejas de Exterior</b>				
<b>Sec</b>	<b>Nral</b>	<b>Nombre de imagen</b>	<b>Descripción de la imagen</b>	<b>Dimensiones</b>
1	m	MEE1DO	Muestra Esmeralda Exterior 1 demasiado oscuro	1153x1544x4
2	n	MEE2BO	Muestra Esmeralda Exterior 2 bastante oscuro	1153x1544x4
3	o	MEE2DO	Muestra Esmeralda Exterior 2 demasiado oscuro	1153x1544x4
4	p	MIE2O	Muestra Isla Exterior 2 oscuro	992x1328x4
5	q	MIE3O	Muestra Isla Exterior 3 oscuro	992x1328x4
6	r	MLE1O	Muestra Loma Exterior 1 oscuro	1067x1429x4
7	s	MLE1BO	Muestra Loma Exterior 1 bastante oscuro	1067x1429x4
8	t	MLE2BO	Muestra Loma Exterior 2 bastante oscuro	981x1314x4
9	u	MLE3DEE	Muestra Loma Exterior 3 DEE	1007x1348x4
10	v	MLE4BO	Muestra Loma Exterior 4 bastante oscuro	926x1240x4
11	w	MLE4DO	Muestra Esmeralda Exterior 4 demasiado oscuro	926x1240x4

Las imágenes originales descritas en la Tabla 4-5 son mostradas en las Figuras 4-4 y 4-5



**Figura 4-4:** Muestras originales de imágenes del veneno de abejas de Interior según numeración de la Tabla 4-5

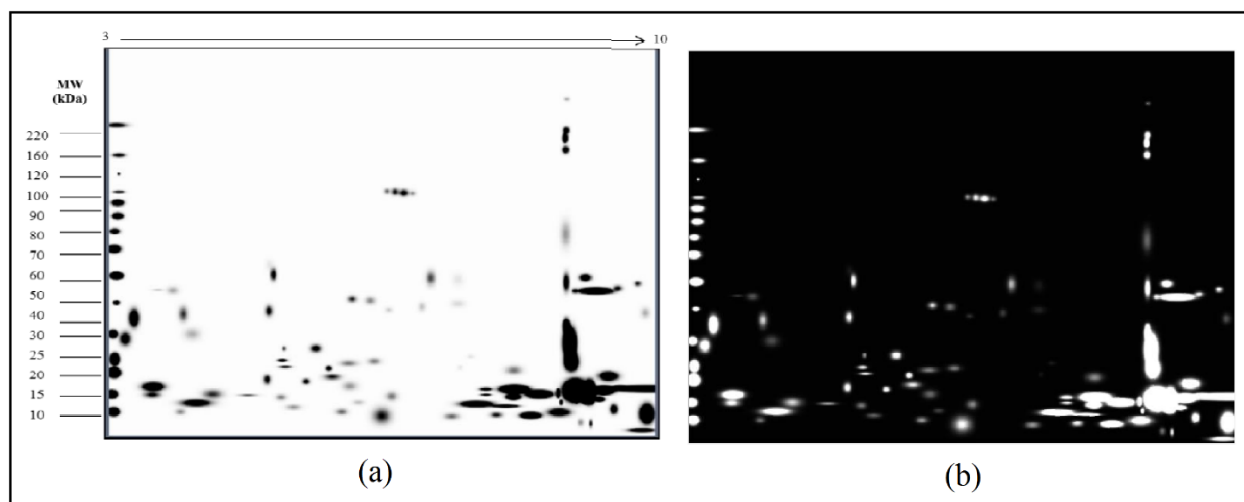


**Figura 4-5:** Muestras originales de imágenes del veneno de abejas de Exterior según numeración de la Tabla 4-5

El patrón general de comparación para las imágenes de apitoxina está compuesto por 115 proteínas, fue generado con la información contenida en el trabajo de investigación: “Informe

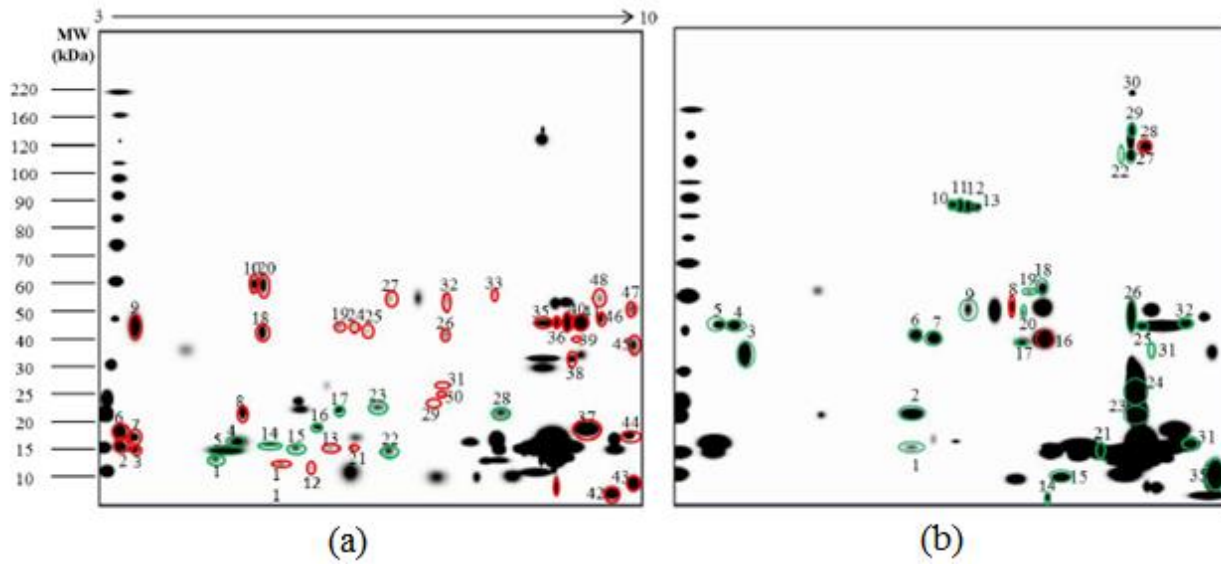
FGN008 Formato de Informe Final P14227”, llevado a cabo por: Grupo de Investigación e Innovación Biomédica - GI2B y Grupo de Investigación en Automática, Electrónica y Ciencias Computacionales – AEyCC, del Instituto Tecnológico Metropolitano. Se toma como referencia el análisis comparativo de las imágenes de los geles realizado con el programa PDQuest™ Advanced 8.0.1. Software (Bio-Rad) tomadas del informe FGN008 páginas 27, 28, 29.

La imagen de la Figura 4-6 “Imagen máster sintética” compuesta de la sumatoria de todas las imágenes 2DGE, se acondiciona a la dimensión estándar de  $1020 \times 1364$  píxeles, se le aplica el proceso para hallar su complemento, tal como lo muestra la Figura 4-6(b), ya que, para los efectos de encontrar objetos, estos se deben mostrar en píxeles de color blanco y fondo negro. Esta imagen se somete al proceso de umbralización automática para encontrar las características de las proteínas, dando como resultado 62 objetos totales, los cuales conforman parte de la base para construir el patrón de comparación de todas las imágenes.

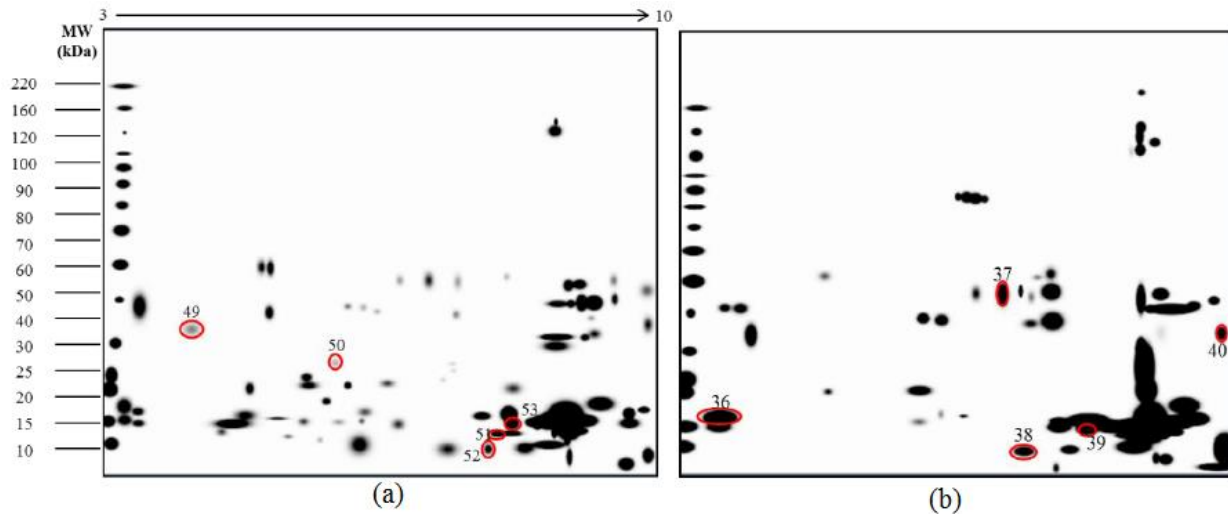


**Figura 4-6:** (a) Imagen máster sintética. Tomado de la Figura 13 Informe FGN008. (b) Complemento de la Figura (a) estandarizado

El resto de las proteínas que conforman el patrón está dado por la extracción de proteínas que no se encuentran en la Figura 4-6 y que se encuentran en las Figuras 4-7, 4-8 y 4-9 tomadas del trabajo Informe FGN008.

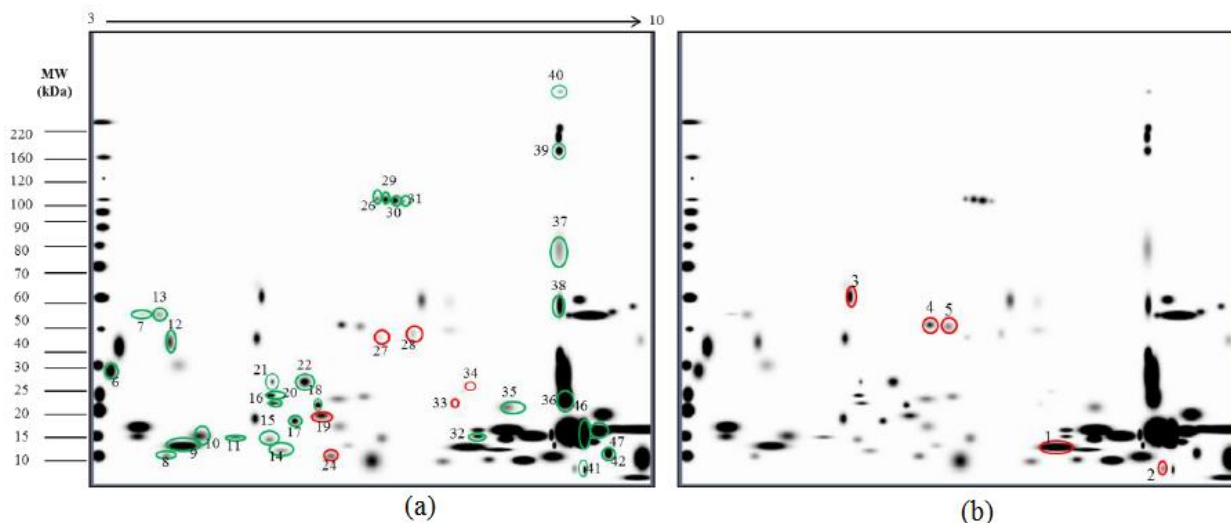


**Figura 4-7:** Imagen máster de apitoxina donde se resaltan las proteínas con ausencia/presencia. En rojo se observan las proteínas solo presentes en apitoxina de interior, y en verde se observan las proteínas solo presentes en apitoxina de exterior. (a) Apitoxina obtenida de la finca La Isla, (b) apitoxina obtenida de la finca La Loma. Tomado de la Figura 14 del Informe FGN008



**Figura 4-8:** Imagen máster de apitoxina donde se resaltan las proteínas con expresión diferencial (en rojo). (a) Apitoxina obtenida de la finca La Isla, (b) apitoxina obtenida de la finca La Loma. Tomado de la Figura 15 Informe FGN008





**Figura 4-9:** Análisis global de apitoxina de exterior vs. apitoxina de interior. (a) Imagen máster de apitoxina donde se resaltan las proteínas con ausencia/presencia. En rojo se observan las proteínas solo presentes en apitoxina de interior, y en verde se observan las proteínas solo presentes en apitoxina de exterior. (b) Imagen master de apitoxina donde se resaltan las proteínas con expresión diferencial (en rojo). Tomado de la Figura 16 Informe FGN008.

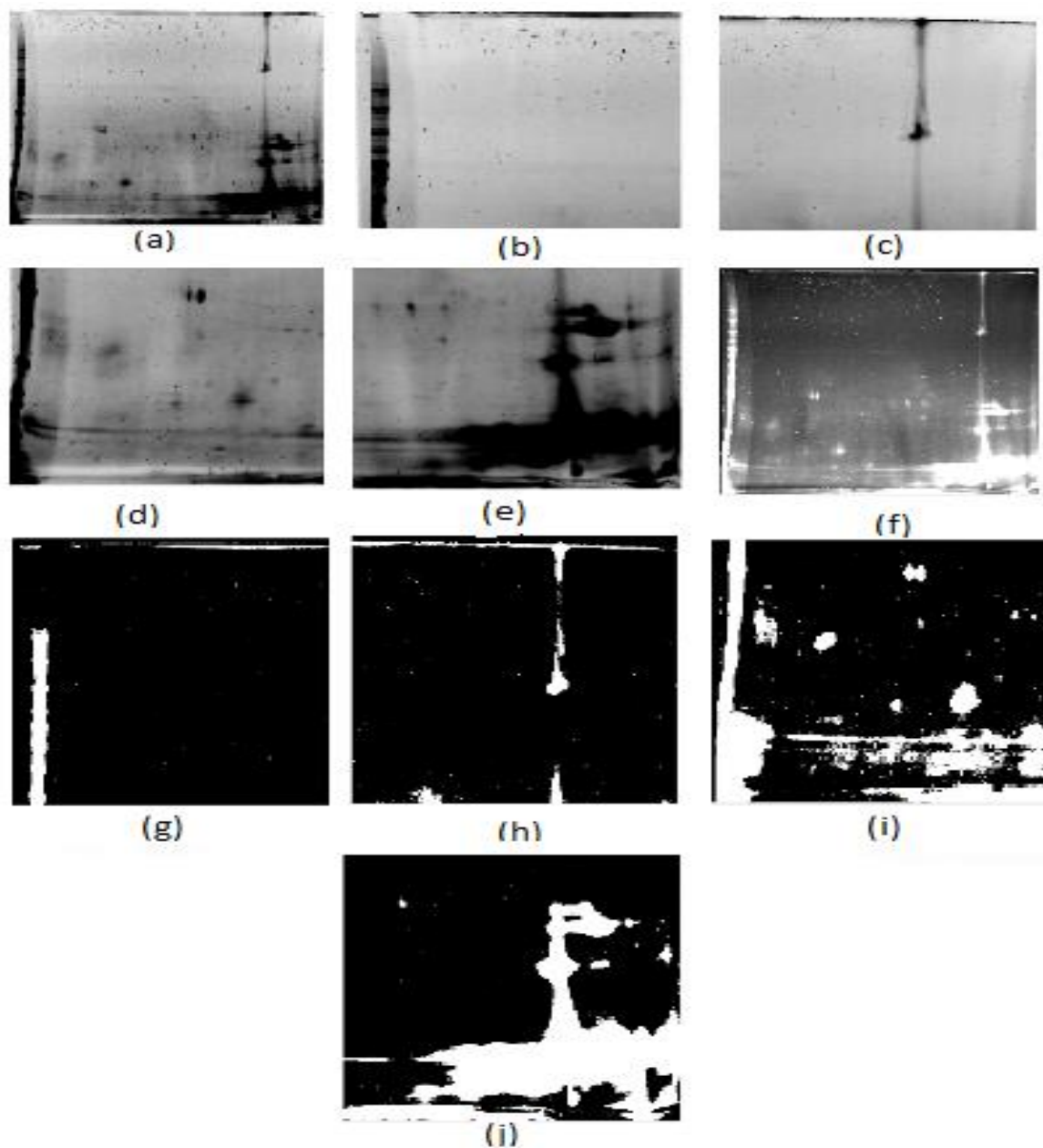
La Tabla 4-6 muestra las 115 coordenadas y características espaciales del patrón de apitoxina conformada con todas las proteínas comunes y no comunes tomadas de las imágenes de las Figuras 4-6, 4-7, 4-8 y 4-9, utilizadas para realizar la comparación de las técnicas de comparación estudiadas en este proyecto.

Como se observa en la Tabla 4-6, el campo Nro. spot corresponde a la numeración que presentan las proteínas identificadas en las imágenes de las Figuras 4-6; 4-7; 4-8; 4-9, en el cuál pueden haber 2 o 3 proteínas comunes y que pertenecen a diferentes grupos poblacionales (de interior y de exterior), el ítem f13 del campo “Grupo” corresponde a proteínas contenidas en la Figura 13 Informe FGN008. Existen algunos spots en el patrón que fueron encontrados en la imagen máster de la Figura 4-6 y que no fueron encontrados en el resto de las imágenes pertenecientes al informe FGN008.

La comparación de las imágenes se realizó únicamente con las técnicas Umbralización automática y Multiumbralización para 4, 8 y 16 ventanas, ya que para el tipo de imágenes estudiadas éstas técnicas eran las que mejor aplicaban en resultados y visualización final de los objetos encontrados. Ver Figura 4-3.

**Tabla 4-6:** Coordenadas (x,y) para las 115 proteínas del patrón de apitoxina de imágenes de abejas de interior y de exterior.

#	Nro Spot	(x, y)	Grupo	#	Nro Spot	(x, y)	Grupo	#	Nro Spot	(x, y)	Grupo	#	Nro Spot	(x, y)	Grupo
1	1-51	(942,943)	Int-Ext	30	32-52	(945,916)	Int-Ext	59	13	(586,929)	Int	88	8	(846,595)	Int
2	2	(1175,985)	Int	31	33	(883,840)	Int	60	14	(432,921)	Ext	89	14	(933,1008)	Ext
3	3-20	(413,599)	Int	32	34	(921,798)	Int	61	31-39	(1190,690)	Int-Ext	90	16	(927,666)	Int
4	4-24	(609,663)	Int-Ext	33	35	(1015,850)	Ext	62	32-40	(1222,641)	Int-Ext	91	18	(921,557)	Ext
5	5	(650,665)	Int	34	22-36	(1155,789)	Ext	63	15	(499,929)	Ext	92	19	(888,565)	Ext
6	6	(41,768)	Ext	35	37	(1139,491)	Ext	64	35-43	(1347,967)	Int-Ext	93	15	(971,969)	Ext
7	5-7	(117,638)	Ext	36	6-35-38	(1146,619)	Ext	65	25-36	(1168,640)	Int-Ext	94	22	(1144,231)	Ext
8	8	(175,954)	Ext	37	27-39	(1145,270)	Ext	66	16	(553,885)	Ext	95	23	(1158,832)	Ext
9	9	(219,939)	Ext	38	6-40	(1148,138)	Ext	67	18	(415,679)	Int	96	28	(1175,252)	Int
10	10	(259,915)	Ext	39	41	(1208,992)	Ext	68	19	(607,667)	Int	97	29	(1146,220)	Ext
11	11	(918,349)	Ext	40	42	(1266,955)	Ext	69	21	(643,928)	Ext	98	52	(857,972)	Ext
12	3-12-49	(186,702)	Ext	41	46	(1162,908)	Ext	70	22	(731,934)	Ext	99	36	(110,900)	Ext
13	4-13	(161,637)	Ext	42	47	(1236,899)	Ext	71	23	(704,841)	Ext	100	37	(806,608)	Ext
14	14	(456,943)	Ext	43	1	(297,952)	Ext	72	28	(1013,853)	Ext	101	38	(859,974)	Ext
15	15	(429,919)	Ext	44	2-17	(601,831)	Ext	73	29	(840,832)	Int	102	39	(1058,971)	Ext
16	16	(442,842)	Ext	45	9-27	(737,603)	Int	74	30	(864,814)	Int	103	40	(1343,697)	Ext
17	17	(494,881)	Ext	46	2	(58,924)	Int	75	31	(866,794)	Int	104	F13	(1194,607)	Sin_def
18	18	(551,845)	Ext	47	3	(97,933)	Int	76	33	(997,596)	Int	105	F13	(62,713)	Sin_def
19	19	(560,869)	Int	48	4	(351,914)	Ext	77	34	(1153,1013)	Int	106	F13	(397,876)	Sin_def
20	20	(435,824)	Ext	49	17-26	(869,674)	Int	78	37	(1227,887)	Int	107	F13	(401,694)	Pat-f13
21	21	(439,792)	Ext	50	5	(321,934)	Ext	79	38	(1190,736)	Int	108	F13	(684,973)	Pat-f13
22	22-50	(519,793)	Ext	51	6	(58,891)	Int	80	41	(1213,658)	Int	109	F13	(710,921)	Pat-f13
23	24	(583,962)	Int	52	7	(93,905)	Int	81	42	(1291,1029)	Int	110	F13	(945,902)	Pat-f13
24	10-26	(697,380)	Ext	53	8	(365,853)	Int	82	44	(1336,902)	Int	111	F13	(1129,963)	Pat-f13
25	27	(700,688)	Int	54	9	(95,665)	Int	83	45	(1347,706)	Int	112	F13	(1133,937)	Pat-f13
26	28	(780,680)	Int	55	32	(873,608)	Int-Ext	84	46	(1265,649)	Int	113	F13	(1253,867)	Pat-f13
27	11-29	(713,378)	Ext	56	10	(394,574)	Int	85	47	(1342,628)	Int	114	F13	(1334,1010)	Pat-f13
28	12-30	(731,382)	Ext	57	11	(460,963)	Int	86	48	(1260,606)	Int	115	F13	(1326,622)	Pat-f13
29	13-31	(758,382)	Ext	58	12	(539,970)	Int	87	1	(602,893)	Ext				



**Figura 4-10:** Imagen muestra Isla interior -2 bastante oscura (MII2BO) y su complemento divididas en 4 ventanas : (a) MII2BO; (b) Ventana 1 de MII2BO; (c) Ventana 2 de MII2BO; (d) Ventana 3 de MII2BO; (e) Ventana 4 de MII2BO; (f) MII2BO complemento; (g) Ventana 1 de MII2BO complemento; (h) Ventana 2 de MII2BO complemento; (i) Ventana 3 de MII2BO complemento; (j) Ventana 4 de MII2BO complemento.

**Tabla 4-7:** Resumen general de comparación técnicas Umbralización automática, Multiumbralización en 4 - 8 y 16 ventanas para las imágenes de interior muestras La Isla y La loma, con respecto al número de objetos conocidos definidos en el patrón de apitoxina mostrados en la Tabla 4-6.

<b>TÉCNICA UMBRALIZACIÓN AUTOMÁTICA IMÁGENES DE INTERIOR</b>									
<b>Imágenes Estandarizadas OC=115</b>		<b>NS</b>	<b>VP</b>	<b>FP</b>	<b>FN</b>	<b>P (%)</b>	<b>RF (%)</b>	<b>RFN (%)</b>	<b>S (%)</b>
La Isla	MII1SOX	90	7	83	108	7,8%	92,2%	93,9%	6,1%
	MII1DOX	188	11	177	104	5,9%	94,1%	90,4%	9,6%
	MII2BOX	1448	22	1426	93	1,5%	98,5%	80,9%	19,1%
	MII2DOX	204	13	191	102	6,4%	93,6%	88,7%	11,3%
	MII3BOX	115	10	105	105	8,7%	91,3%	91,3%	8,7%
La Loma	MLI1BOX	2495	29	2466	86	1,2%	98,8%	74,8%	25,2%
	MLI1DOX	202	6	196	109	3,0%	97,0%	94,8%	5,2%
	MLI1OX	153	4	149	111	2,6%	97,4%	96,5%	3,5%
	MLI2BOX	264	9	255	106	3,4%	96,6%	92,2%	7,8%
	MLI2DOX	115	3	112	112	2,6%	97,4%	97,4%	2,6%
	MLI3BOX	1428	18	1410	97	1,3%	98,7%	84,3%	15,7%
	MLI3DOX	125	11	114	104	8,8%	91,2%	90,4%	9,6%
<b>Promedio Total</b>		569	12	557	103	4,4%	95,6%	89,6%	10,4%
<b>TÉCNICA MULTIUMBRAIZACIÓN AUTOMÁTICA 4 VENTANAS IMÁGENES DE INTERIOR</b>									
<b>Imágenes Estandarizadas OC=115</b>		<b>NS</b>	<b>VP</b>	<b>FP</b>	<b>FN</b>	<b>P (%)</b>	<b>RF (%)</b>	<b>RFN (%)</b>	<b>S (%)</b>
La Isla	MII1SOX	65	3	62	112	4,6%	95,4%	97,4%	2,6%
	MII1DOX	141	20	121	95	14,2%	85,8%	82,6%	17,4%
	MII2BOX	705	37	668	78	5,2%	94,8%	67,8%	32,2%
	MII2DOX	159	7	152	108	4,4%	95,6%	93,9%	6,1%
	MII3BOX	139	11	128	104	7,9%	92,1%	90,4%	9,6%
La Loma	MLI1BOX	958	20	938	95	2,1%	97,9%	82,6%	17,4%
	MLI1DOX	181	8	173	107	4,4%	95,6%	93,0%	7,0%
	MLI1OX	107	2	105	113	1,9%	98,1%	98,3%	1,7%
	MLI2BOX	191	12	179	103	6,3%	93,7%	89,6%	10,4%
	MLI2DOX	99	2	97	113	2,0%	98,0%	98,3%	1,7%
	MLI3BOX	470	15	455	100	3,2%	96,8%	87,0%	13,0%
	MLI3DOX	62	4	58	111	6,5%	93,5%	96,5%	3,5%
<b>Promedio Total</b>		273	12	261	103	5,2%	94,8%	89,8%	10,2%

<b>TÉCNICA MULTIUMBRAJIZACIÓN AUTOMÁTICA 8 VENTANAS IMÁGENES DE INTERIOR</b>									
<b>Imágenes Estandarizadas OC=115</b>		<b>NS</b>	<b>VP</b>	<b>FP</b>	<b>FN</b>	<b>P (%)</b>	<b>RF (%)</b>	<b>RFN (%)</b>	<b>S (%)</b>
La Isla	MII1SOX	59	3	56	112	5,1%	94,9%	97,4%	2,6%
	MII1DOX	168	13	155	102	7,7%	92,3%	88,7%	11,3%
	MII2BOX	672	39	633	76	5,8%	94,2%	66,1%	33,9%
	MII2DOX	162	6	156	109	3,7%	96,3%	94,8%	5,2%
	MII3BOX	77	4	73	111	5,2%	94,8%	96,5%	3,5%
La Loma	MLI1BOX	683	25	658	90	3,7%	96,3%	78,3%	21,7%
	MLI1DOX	187	8	179	107	4,3%	95,7%	93,0%	7,0%
	MLI1OX	113	1	112	114	0,9%	99,1%	99,1%	0,9%
	MLI2BOX	205	12	193	103	5,9%	94,1%	89,6%	10,4%
	MLI2DOX	111	2	109	113	1,8%	98,2%	98,3%	1,7%
	MLI3BOX	480	17	463	98	3,5%	96,5%	85,2%	14,8%
	MLI3DOX	69	5	64	110	7,2%	92,8%	95,7%	4,3%
<b>Promedio Total</b>		<b>249</b>	<b>11</b>	<b>238</b>	<b>104</b>	<b>4,6%</b>	<b>95,4%</b>	<b>90,2%</b>	<b>9,8%</b>
<b>TÉCNICA MULTIUMBRAJIZACIÓN AUTOMÁTICA 16 VENTANAS IMÁGENES DE INTERIOR</b>									
<b>Imágenes Estandarizadas OC=115</b>		<b>NS</b>	<b>VP</b>	<b>FP</b>	<b>FN</b>	<b>P (%)</b>	<b>RF (%)</b>	<b>RFN (%)</b>	<b>S (%)</b>
La Isla	MII1SOX	67	2	65	113	3,0%	97,0%	98,3%	1,7%
	MII1DOX	126	14	112	101	11,1%	88,9%	87,8%	12,2%
	MII2BOX	932	37	895	78	4,0%	96,0%	67,8%	32,2%
	MII2DOX	172	7	165	108	4,1%	95,9%	93,9%	6,1%
	MII3BOX	66	6	60	109	9,1%	90,9%	94,8%	5,2%
La Loma	MLI1BOX	1279	28	1251	87	2,2%	97,8%	75,7%	24,3%
	MLI1DOX	177	8	169	107	4,5%	95,5%	93,0%	7,0%
	MLI1OX	85	2	83	113	2,4%	97,6%	98,3%	1,7%
	MLI2BOX	213	10	203	105	4,7%	95,3%	91,3%	8,7%
	MLI2DOX	117	2	115	113	1,7%	98,3%	98,3%	1,7%
	MLI3BOX	745	8	737	107	1,1%	98,9%	93,0%	7,0%
	MLI3DOX	87	5	82	110	5,7%	94,3%	95,7%	4,3%
<b>Promedio Total</b>		<b>339</b>	<b>11</b>	<b>328</b>	<b>104</b>	<b>4,5%</b>	<b>95,5%</b>	<b>90,7%</b>	<b>9,3%</b>

PROMEDIOS POR IMAGEN DE INTERIOR									
Imágenes Estandarizadas OC=115		NS	VP	FP	FN	P (%)	RF (%)	RFN (%)	S (%)
La Isla	MII1SOX	70	4	67	111	5,1%	94,9%	96,7%	3,3%
	MII1DOX	156	15	141	101	9,7%	90,3%	87,4%	12,6%
	MII2BOX	939	34	906	81	4,1%	95,9%	70,7%	<b>29,3%</b>
	MII2DOX	174	8	166	107	4,6%	95,4%	92,8%	7,2%
	MII3BOX	99	8	92	107	7,7%	92,3%	93,3%	6,7%
La Loma	MLI1BOX	1354	26	1328	90	2,3%	97,7%	77,8%	<b>22,2%</b>
	MLI1DOX	187	8	179	108	4,0%	96,0%	93,5%	6,5%
	MLI1OX	115	2	112	113	1,9%	98,1%	98,0%	2,0%
	MLI2BOX	218	11	208	104	5,1%	94,9%	90,7%	9,3%
	MLI2DOX	111	2	108	113	2,0%	98,0%	98,0%	2,0%
	MLI3BOX	781	15	766	101	2,3%	97,7%	87,4%	12,6%
	MLI3DOX	86	6	80	109	7,1%	92,9%	94,6%	5,4%
<b>Promedio Total</b>		<b>357</b>	<b>11</b>	<b>346</b>	<b>104</b>	<b>4,7%</b>	<b>95,3%</b>	<b>90,1%</b>	<b>9,9%</b>

Resumen de comparación en técnica umbralización imágenes de interior

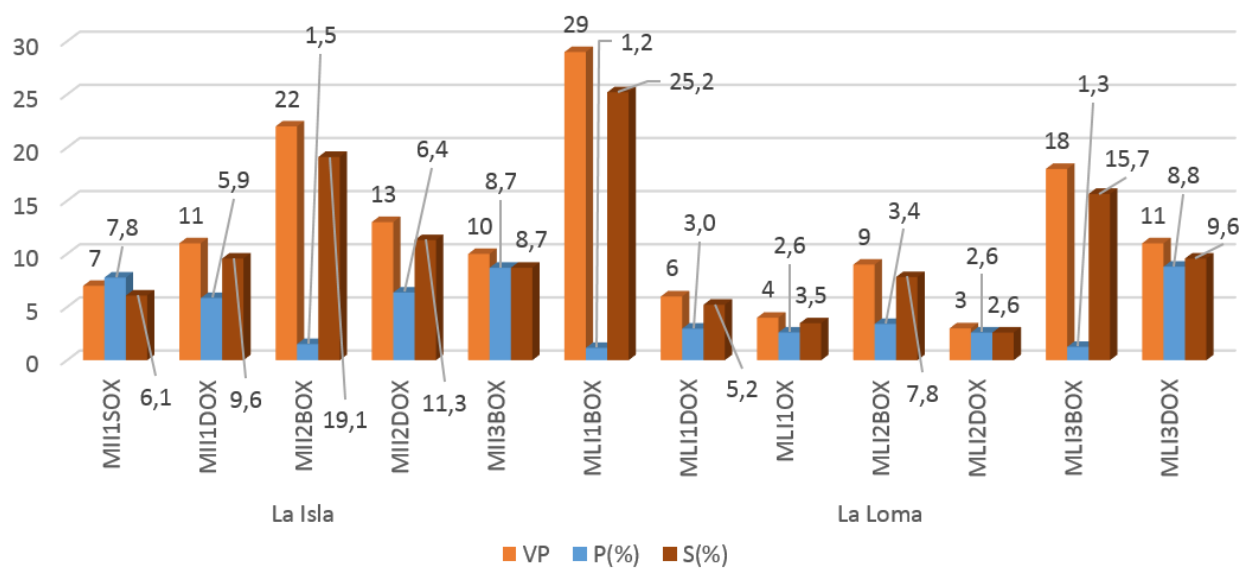


Figura 4-11: Relaciones de proteínas conocidas, precisión y sensibilidad en técnica umbralización automática imágenes de interior Tabla 4-7

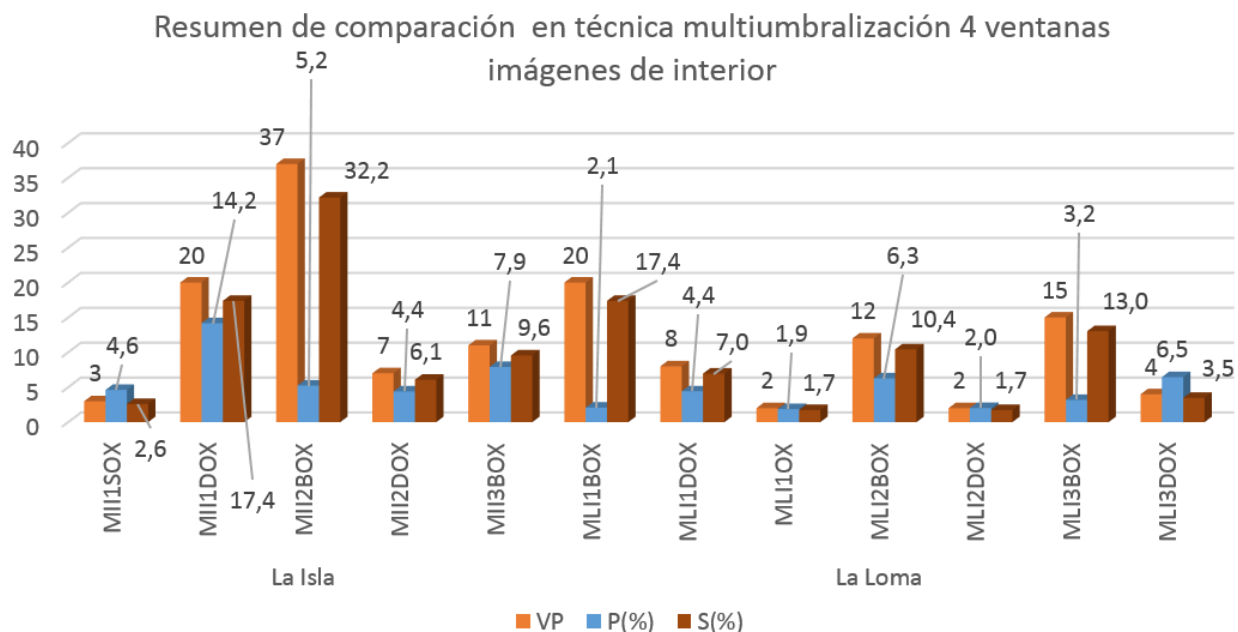


Figura 4-12: Relaciones de proteínas conocidas, precisión y sensibilidad en técnica multiumbralización automática 4 ventanas imágenes de interior Tabla 4-7

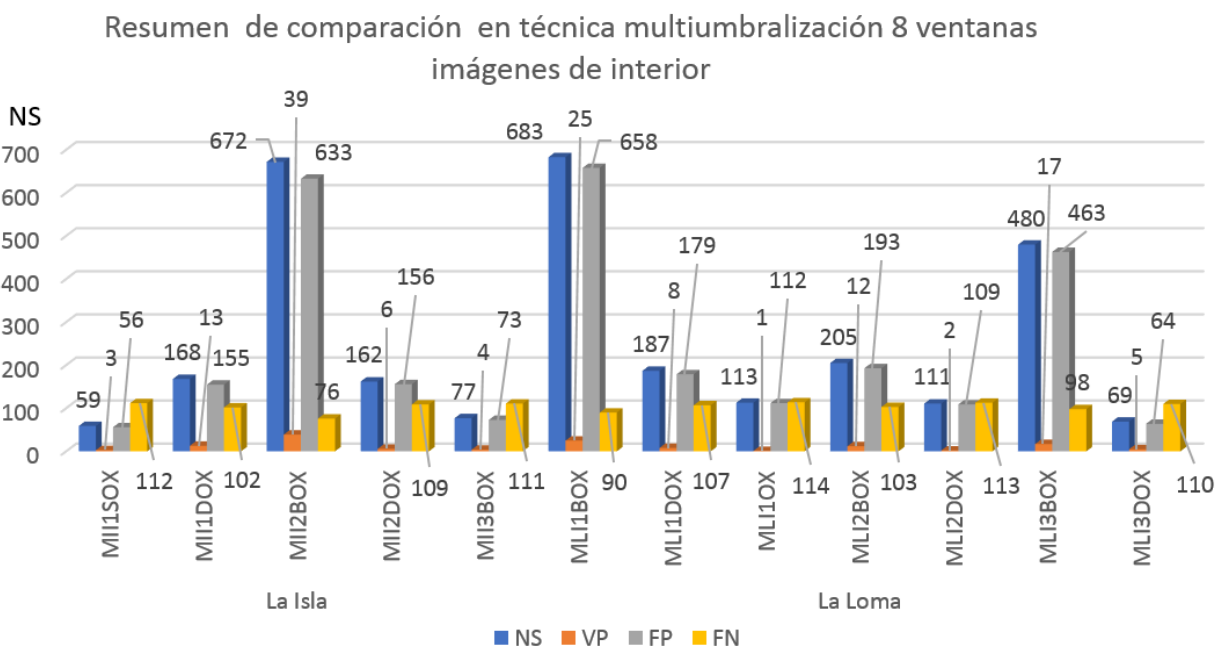


Figura 4-13: Relaciones de objetos segmentados (NS), proteínas de patrón detectadas (VP), objetos segmentados no reconocidos (FP), proteínas de patrón no detectadas (FN) en técnica multiumbralización automática 8 ventanas imágenes de interior Tabla 4-7

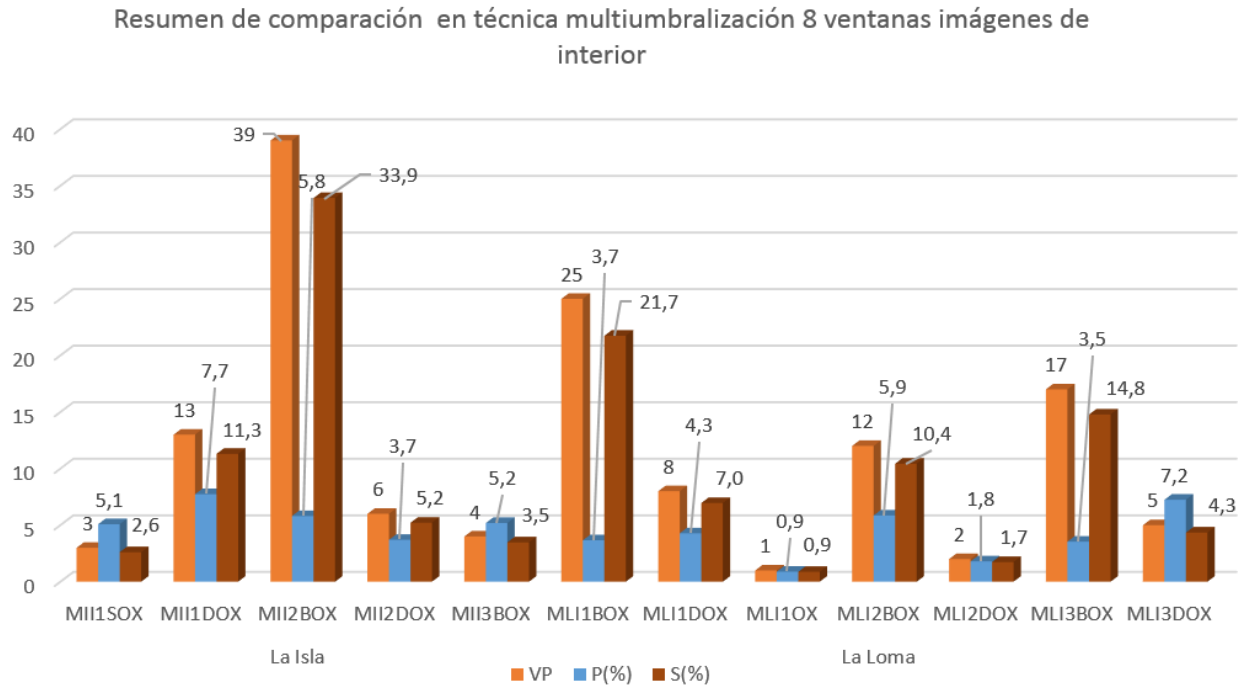


Figura 4-14: Relaciones de proteínas conocidas, precisión y sensibilidad en técnica multiumbralización automática 8 ventanas imágenes de interior Tabla 4-7

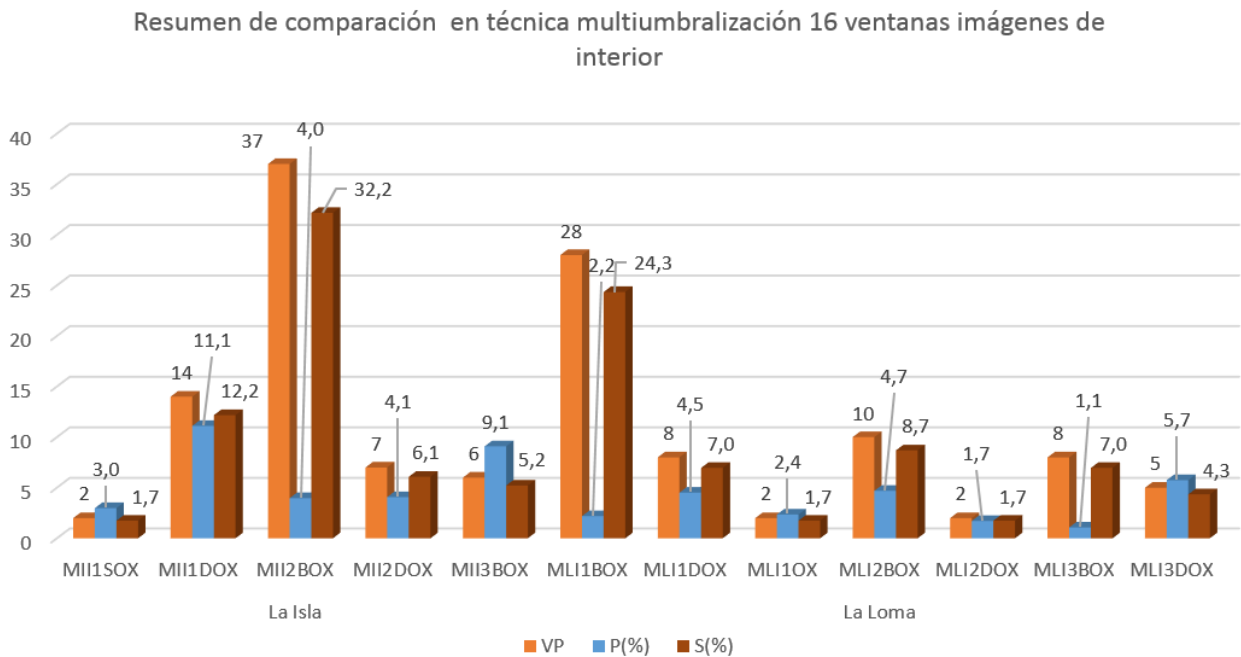


Figura 4-15: Relaciones de proteínas conocidas, precisión y sensibilidad en técnica multiumbralización automática 16 ventanas imágenes de interior Tabla 4-7

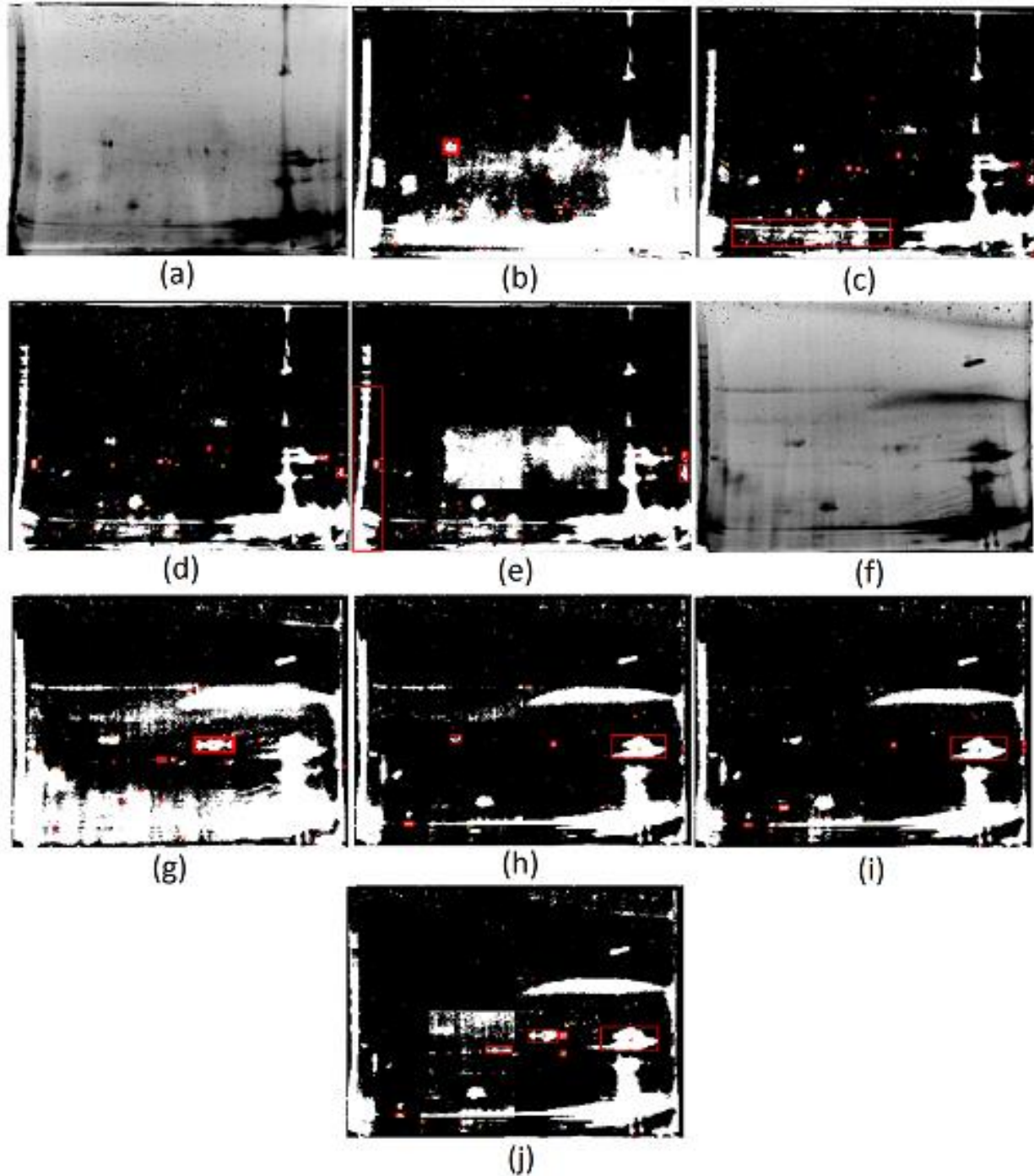


La Figura 4-10 muestra la imagen Isla interior 2 bastante oscura (MII2BO) dividida en 4 ventanas y su complemento también dividido en 4 ventanas.

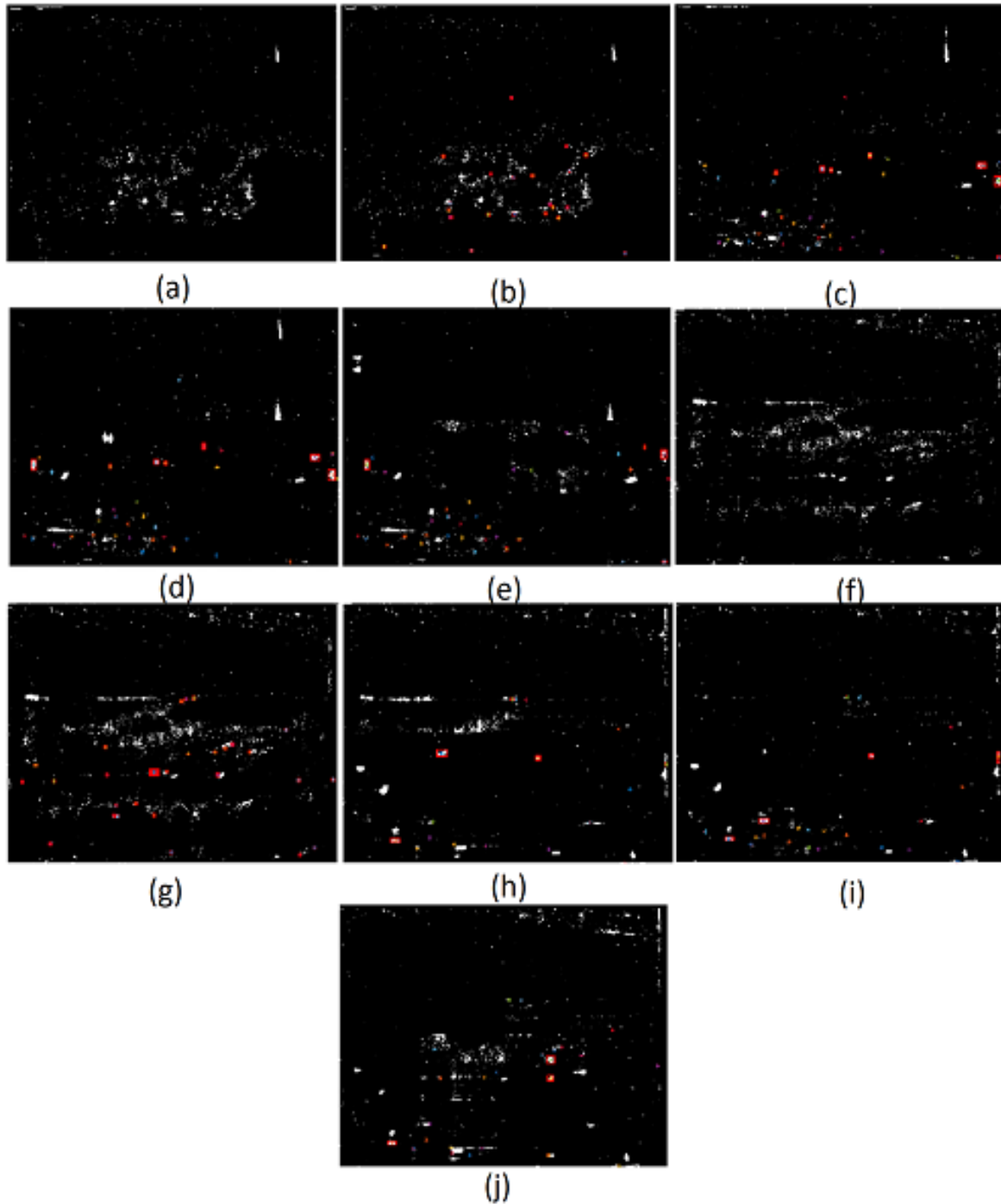
Los resultados obtenidos para cada una de las técnicas de comparación para las muestras de interior y de exterior son mostrados en las Tablas 4-7, 4-8 y Figuras 4-11, 4-12, 4-13, 4-14, 4-15, 4-18, 4-19, 4-20, 4-21, 4-22. Las imágenes procesadas cambian su nombre, agregándole una X al final, para identificarlas como imágenes originales a las cuales se les ha aplicado la estandarización en el tamaño ( $1020 \times 1364$  *pixeles*).

En la Figura 4-16 se realiza la presentación de las muestras de las imágenes MII2BOX (muestra Isla Interior 2 bastante oscuro) y MLI1BOX (muestra Loma interior 1 bastante oscuro), también se observan las proteínas conocidas marcadas para cada una de las técnicas de segmentación descritas. Según la Tabla 4-7 estas dos muestras presentan los resultados más relevantes en los promedios de sensibilidad por imagen (29.35% y 22.17% respectivamente) y en Figuras 4-11, 4-12, 4-14, 4-15 se visualizan éstos resultados, presentándose el mayor valor de proteínas conocidas (VP) y porcentaje de sensibilidad (S%) en la técnica de multiumbralización automática 8 ventanas con valores de 39 proteínas y 33.9% respectivamente (Figura 4-14). Las Figuras 4-13 y 4-20 muestran los resultados de NS, VP, FP y FN, para el mejor comportamiento en técnicas de segmentación empleadas en el análisis de las muestras de imágenes de interior y de exterior.

Debido a la existencia de objetos con áreas muy grandes en las imágenes, lo cual dificulta la observación y detección de objetos con áreas más pequeñas que se encuentren en sus proximidades, se eliminan los objetos con áreas mayores a 1000 *pixeles*. Esta eliminación se observa en la Figura 4-17, en la cual se muestran las mismas imágenes de la Figura 4-16, pero sin áreas mayores a 1000 *pixeles*.



**Figura 4-16:** Imágenes resultantes de los procesos de umbralización automática y multiumbralización para 4 – 8 y 16 ventanas para las muestras Isla interior-2 bastante oscuro (MII2BO) y Loma interior1 bastante oscuro (MLI1BO): (a) MII2BO; (b) Umbral automático MII2BO; (c) Multiumbral 4 Ventanas MII2BO; (d) Multiumbral 8 Ventanas MII2BO; (e) Multiumbral 16 Ventanas MII2BO; (f) MLI1BO; (g) Umbral automático MLI1BO; (h) Multiumbral 4 Ventanas MLI1BO; (i) Multiumbral 8 Ventanas LI1BO; (j) Multiumbral 16 Ventanas MLI1BO.



**Figura 4- 17:** Imágenes resultantes de los procesos de umbralización automática y multiumbralización para 4 – 8 y 16 ventanas para las muestras Isla interior-2 bastante oscuro y Loma interior1 bastante oscuro, habiendo eliminado los objetos con áreas mayores a 1000 píxeles. (a) MII2BO; (b) Umbral automático MII2BO; (c) Multiumbral 4 Ventanas MII2BO; (d) Multiumbral 8 Ventanas MII2BO; (e) Multiumbral 16 Ventanas MII2BO; (f) MLI1BO; (g) Umbral automático MLI1BO; (h) Multiumbral 4 Ventanas MLI1BO; (i) Multiumbral 8 Ventanas MLI1BO; (j) Multiumbral 16 Ventanas MLI1BO.

**Tabla 4-8:** Resumen general de comparación técnicas Umbralización automática, Multiumbralización en 4 - 8 y 16 ventanas para las imágenes de exterior muestras La Esmeralda, La Isla y La loma, con respecto al número de objetos conocidos definidos en el patrón de apitoxina mostrados en la Tabla 4-6.

<b>TÉCNICA UMBRALIZACIÓN AUTOMÁTICA IMÁGENES DE EXTERIOR</b>									
<b>Imágenes Estandarizadas OC=115</b>		<b>NS</b>	<b>VP</b>	<b>FP</b>	<b>FN</b>	<b>P (%)</b>	<b>RF (%)</b>	<b>RFN (%)</b>	<b>S (%)</b>
La Esmeralda	MEE1DOX	174	9	165	106	5,2%	94,8%	92,2%	7,8%
	MEE2BOX	365	5	360	110	1,4%	98,6%	95,7%	4,3%
	MEE2DOX	266	11	255	104	4,1%	95,9%	90,4%	9,6%
La Isla	MIE2OX	154	8	146	107	5,2%	94,8%	93,0%	7,0%
	MIE3OX	126	4	122	111	3,2%	96,8%	96,5%	3,5%
La Loma	MLE1OX	77	6	71	109	7,8%	92,2%	94,8%	5,2%
	MLE1BOX1	1573	21	1552	94	1,3%	98,7%	81,7%	18,3%
	MLE2BOX	251	9	242	106	3,6%	96,4%	92,2%	7,8%
	MLE3DEEX	42	1	41	114	2,4%	97,6%	99,1%	0,9%
	MLE4BOX	228	8	220	107	3,5%	96,5%	93,0%	7,0%
	MLE4DOX	210	7	203	108	3,3%	96,7%	93,9%	6,1%
<b>Promedio Total</b>		315	8	307	107	3,7%	96,3%	93,0%	7,0%
<b>TÉCNICA MULTIUMBRAIZACIÓN AUTOMÁTICA 4 VENTANAS IMÁGENES DE EXTERIOR</b>									
<b>Imágenes Estandarizadas OC=115</b>		<b>NS</b>	<b>VP</b>	<b>FP</b>	<b>FN</b>	<b>P (%)</b>	<b>RF (%)</b>	<b>RFN (%)</b>	<b>S (%)</b>
La Esmeralda	MEE1DOX	83	9	74	106	10,8%	89,2%	92,2%	7,8%
	MEE2BOX	437	16	421	99	3,7%	96,3%	86,1%	13,9%
	MEE2DOX	166	6	160	109	3,6%	96,4%	94,8%	5,2%
La Isla	MIE2OX	136	3	133	112	2,2%	97,8%	97,4%	2,6%
	MIE3OX	133	5	128	110	3,8%	96,2%	95,7%	4,3%
La Loma	MLE1OX	72	1	71	114	1,4%	98,6%	99,1%	0,9%
	MLE1BOX1	1270	24	1246	91	1,9%	98,1%	79,1%	20,9%
	MLE2BOX	214	8	206	107	3,7%	96,3%	93,0%	7,0%
	MLE3DEEX	44	0	44	115	0,0%	100,0%	100,0%	0,0%
	MLE4BOX	241	3	238	112	1,2%	98,8%	97,4%	2,6%
	MLE4DOX	237	3	234	112	1,3%	98,7%	97,4%	2,6%
<b>Promedio Total</b>		276	7	269	108	3,1%	96,9%	93,8%	6,2%

<b>TÉCNICA MULTIUMBRALIZACIÓN AUTOMÁTICA 8 VENTANAS IMÁGENES DE EXTERIOR</b>									
<b>Imágenes Estandarizadas OC=115</b>		<b>NS</b>	<b>VP</b>	<b>FP</b>	<b>FN</b>	<b>P (%)</b>	<b>RF (%)</b>	<b>RFN (%)</b>	<b>S (%)</b>
La Esmeralda	MEE1DOX	57	2	55	113	3,5%	96,5%	98,3%	1,7%
	MEE2BOX	381	7	374	108	1,8%	98,2%	93,9%	6,1%
	MEE2DOX	227	10	217	105	4,4%	95,6%	91,3%	8,7%
La Isla	MIE2OX	149	4	145	111	2,7%	97,3%	96,5%	3,5%
	MIE3OX	142	6	136	109	4,2%	95,8%	94,8%	5,2%
La Loma	MLE1OX	63	1	62	114	1,6%	98,4%	99,1%	0,9%
	MLE1BOX1	104 4	29	1015	86	2,8%	97,2%	74,8%	25,2%
	MLE2BOX	254	7	247	108	2,8%	97,2%	93,9%	6,1%
	MLE3DEEX	42	0	42	115	0,0%	100,0 %	100,0%	0,0%
	MLE4BOX	252	3	249	112	1,2%	98,8%	97,4%	2,6%
	MLE4DOX	256	3	253	112	1,2%	98,8%	97,4%	2,6%
<b>Promedio Total</b>		261	7	254	108	2,4%	97,6%	94,3%	5,7%
<b>TÉCNICA MULTIUMBRALIZACIÓN AUTOMÁTICA 16 VENTANAS IMÁGENES DE EXTERIOR</b>									
<b>Imágenes Estandarizadas OC=115</b>		<b>NS</b>	<b>VP</b>	<b>FP</b>	<b>FN</b>	<b>P (%)</b>	<b>RF (%)</b>	<b>RFN (%)</b>	<b>S (%)</b>
La Esmeralda	MEE1DOX	114	8	106	107	7,0%	93,0%	93,0%	7,0%
	MEE2BOX	332	8	324	107	2,4%	97,6%	93,0%	7,0%
	MEE2DOX	198	9	189	106	4,5%	95,5%	92,2%	7,8%
La Isla	MIE2OX	141	4	137	111	2,8%	97,2%	96,5%	3,5%
	MIE3OX	145	7	138	108	4,8%	95,2%	93,9%	6,1%
La Loma	MLE1OX	59	2	57	113	3,4%	96,6%	98,3%	1,7%
	MLE1BOX1	114 0	23	1117	92	2,0%	98,0%	80,0%	20,0%
	MLE2BOX	245	7	238	108	2,9%	97,1%	93,9%	6,1%
	MLE3DEEX	47	2	45	113	4,3%	95,7%	98,3%	1,7%
	MLE4BOX	295	3	292	112	1,0%	99,0%	97,4%	2,6%
	MLE4DOX	295	3	292	112	1,0%	99,0%	97,4%	2,6%
<b>Promedio Total</b>		274	7	267	108	3,3%	96,7%	94,0%	6,0%

PROMEDIOS POR IMAGEN DE EXTERIOR									
Imágenes Estandarizadas OC=115		NS	VP	FP	FN	P (%)	RF (%)	RFN (%)	S (%)
La Esmeralda	MEE1DOX	107	7	100	108	6,6%	93,4%	93,9%	6,1%
	MEE2BOX	378,8	9	370	106	2,3%	97,7%	92,2%	7,8%
	MEE2DOX	214,3	9	205	106	4,2%	95,8%	92,2%	7,8%
La Isla	MIE2OX	145	5	140	110	3,2%	96,8%	95,9%	4,1%
	MIE3OX	136,5	6	131	110	4,0%	96,0%	95,2%	4,8%
La Loma	MLE1OX	67,75	3	65,3	113	3,5%	96,5%	97,8%	2,2%
	MLE1BOX1	1257	24	1233	91	2,0%	98,0%	78,9%	21,1%
	MLE2BOX	241	8	233	107	3,2%	96,8%	93,3%	6,7%
	MLE3DEEX	43,75	1	43	114	1,7%	98,3%	99,3%	0,7%
	MLE4BOX	254	4	250	111	1,7%	98,3%	96,3%	3,7%
	MLE4DOX	249,5	4	246	111	1,7%	98,3%	96,5%	3,5%
<b>Promedio Total</b>		<b>281,3</b>	<b>7</b>	<b>274</b>	<b>108</b>	<b>3,1%</b>	<b>96,9%</b>	<b>93,8%</b>	<b>6,2%</b>

Resumen de comparación en técnica umbralización imágenes de exterior

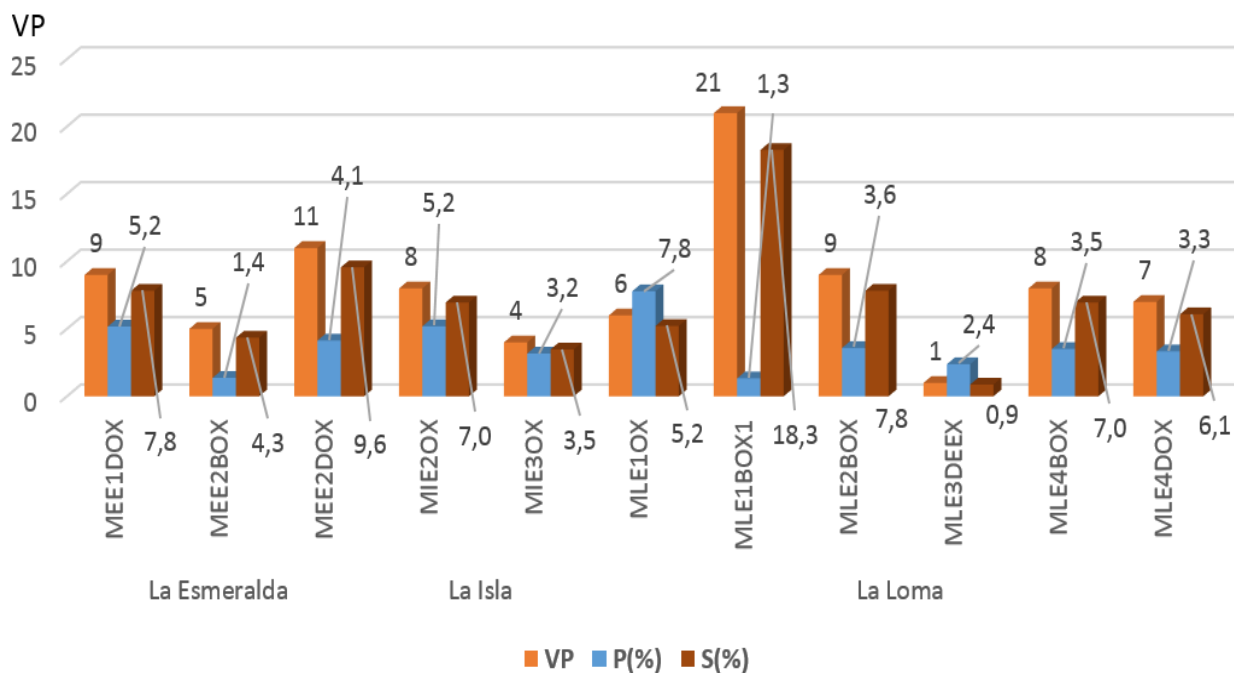


Figura 4-18: Relaciones de proteínas conocidas, precisión y sensibilidad en técnica umbralización automática imágenes de exterior Tabla 4-8

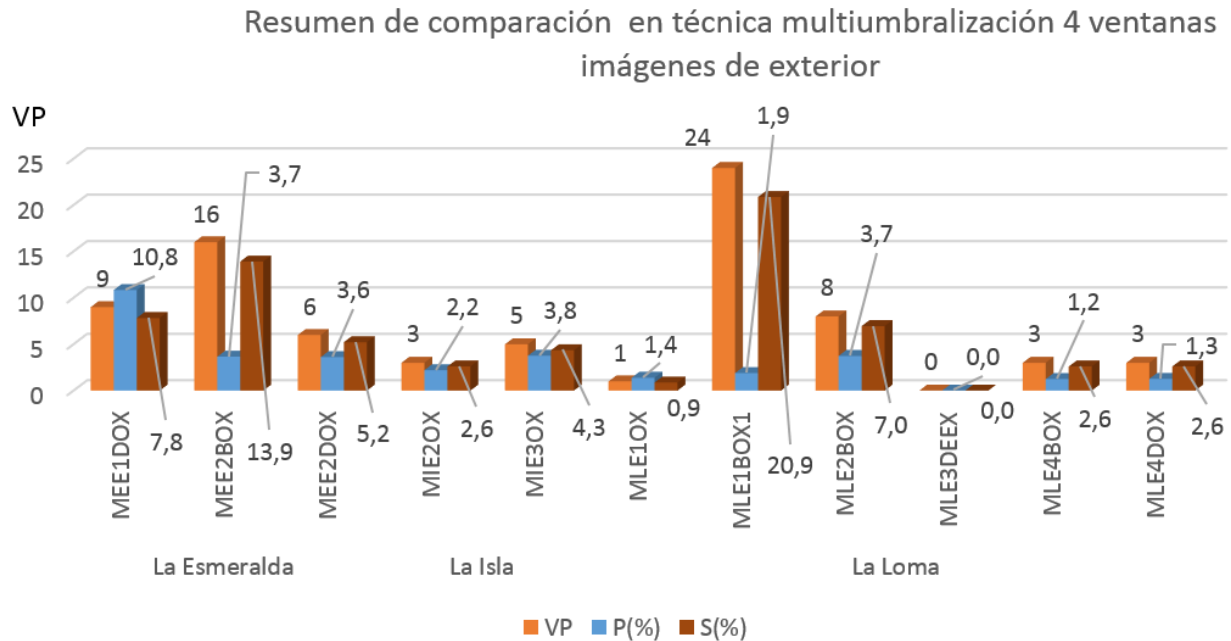


Figura 4- 19: Relaciones de proteínas conocidas, precisión y sensibilidad en técnica multiumbralización automática 4 ventanas imágenes de exterior Tabla 4-8

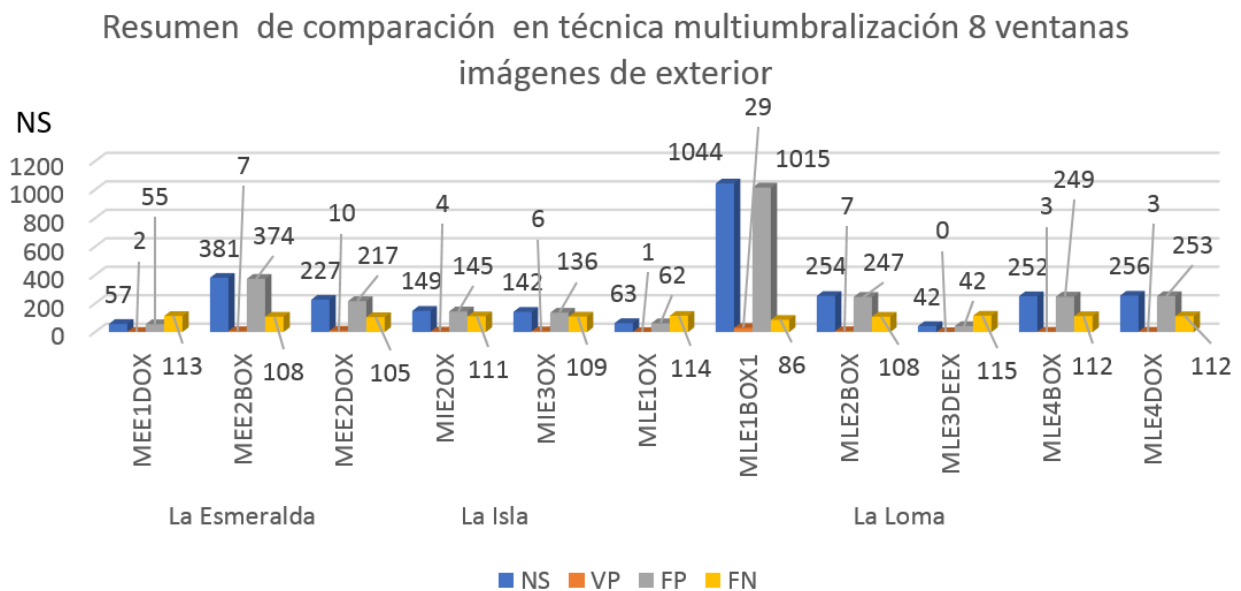


Figura 4-20: Relaciones de objetos segmentados (NS), proteínas de patrón detectadas (VP), objetos segmentados no reconocidos (FP), proteínas de patrón no detectada (FN) en técnica multiumbralización automática 8 ventanas imágenes de exterior Tabla 4-8

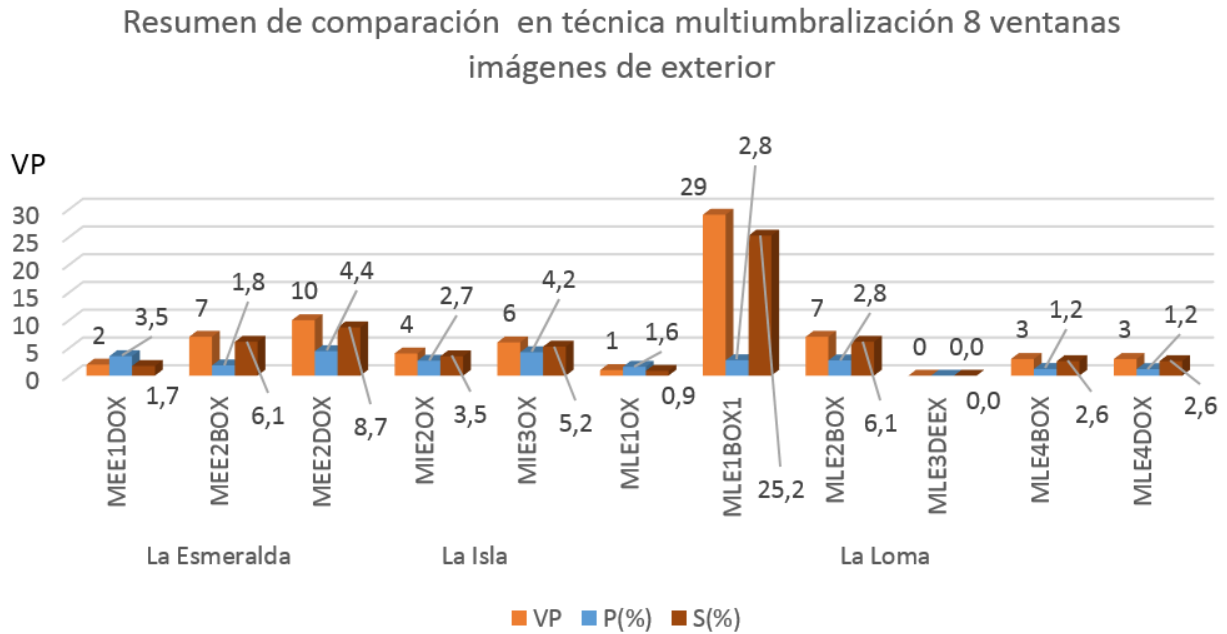


Figura 4-21: Relaciones de proteínas conocidas, precisión y sensibilidad en técnica multiumbralización automática 8 ventanas imágenes de exterior Tabla 4-8

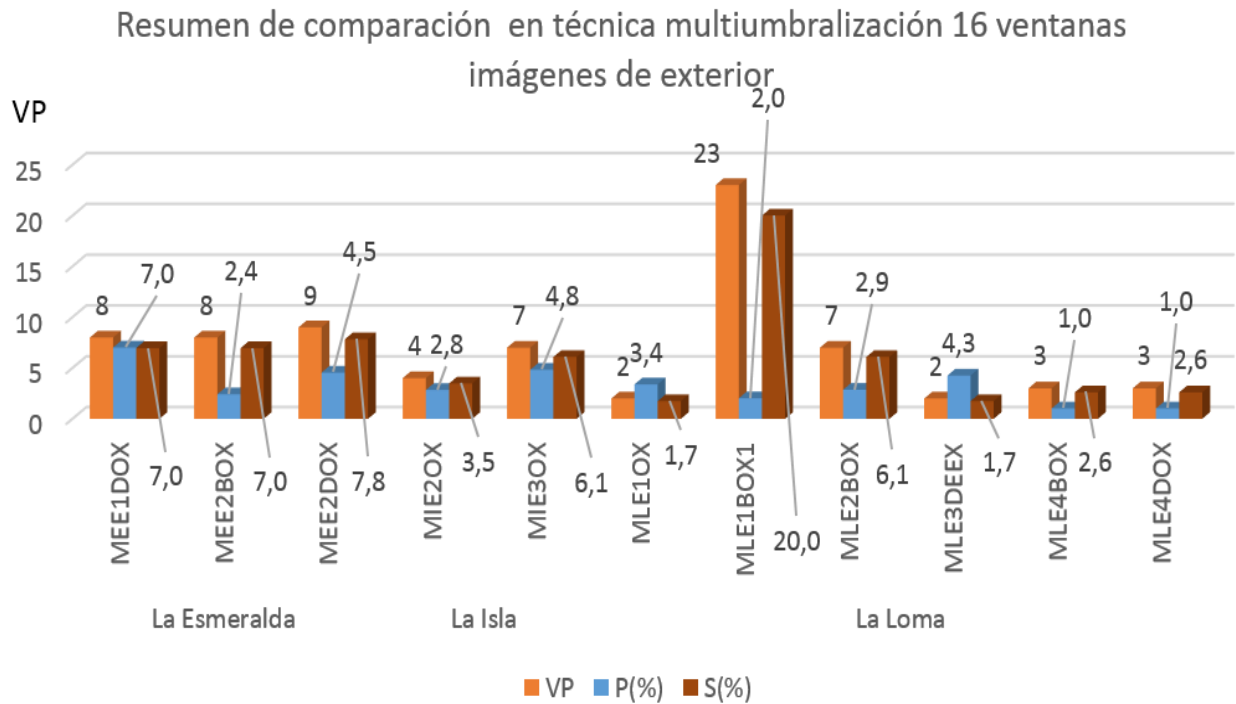


Figura 4-22: Relaciones de proteínas conocidas, precisión y sensibilidad en técnica multiumbralización automática 16 ventanas imágenes de exterior Tabla 4-8



### **4.2.1 Discusión**

En la comparación de las técnicas para las imágenes de apitoxina, no se realizaron procesos con las técnicas Detector de bordes Sobel y Transformada Watershed, ya que, para este tipo de imágenes, según lo observado en el procesamiento para la imagen Dr. Dirk de Graaf, los resultados no eran los más satisfactorios, en el primero por la gran cantidad de objetos que encontraba, llegándose a confundir algunos bordes como objetos múltiples y el segundo por la poca detección de objetos y por los problemas que presenta este método cuando no es combinado con otro método adicional para evitar una sobre segmentación. En los datos de comparación de la Tabla 4-7 se observa que el promedio general mayor para la relación de objetos detectados con respecto a los objetos conocidos se encuentra representado en los resultados de la técnica umbralización automática con un 10.4%, relación que también depende de la mayor cantidad de objetos totales de la imagen, así como lo revela el promedio de 569 objetos. Los promedios para el resto de los métodos están muy próximos al mayor, siendo de 10.2%, 9.8% y 9.3% para las técnicas de multimumbralización automática para 4, 8 y 16 ventanas respectivamente, pero con una cantidad promedio de objetos totales con propiedades casi igual a la mitad de los de la técnica de umbralización automática. Los mayores promedios por imagen fueron dados por las muestras que contenían mayor cantidad de objetos como los fueron la imagen Isla Interior-2 bastante oscura (MII2BOX) con un 29.35% de objetos detectados con respecto a los conocidos de una cantidad total de objetos con propiedades de 939 y la muestra de la Loma Interior-1 bastante oscuro (MLI1BOX) con un 22,17% de un total de 1354 objetos totales con propiedades. Los datos de éstas dos imágenes nos revelan que los resultados dependen del grado de intensidad de luminosidad al que sea sometido la muestra, cómo también del proceso de tratamiento de la imagen y escaneo realizado en el laboratorio, dato muy importante para ser tenido en cuenta por el personal que realiza dichos procedimientos, teniendo en cuenta que para las mismas muestras con diferentes intensidades tal como la muestra Isla-2 demasiado oscura y muestra Loma -1 demasiado oscura, los cálculos resultantes fueron mucho menores que los de las muestras inicialmente referidas. Los resultados dados para imágenes iguales con diferentes grados de luminosidad no dependen de la técnica que se emplee sino del grado de estandarización que se logre implementar en el laboratorio para obtener imágenes con una buena definición del contenido de proteínas u objetos de análisis.

Los resultados de los datos de comparación de la Tabla 4-8 para las imágenes de exterior, son muy similares a los de la Tabla 4-7 de las imágenes de interior, con la diferencia que los promedios resultantes son un poco menores que los de la Tabla 4-7 y que la imagen con mejores resultados fue la muestra Loma Exterior-1 bastante oscuro (MLE1BOX) con una relación promedia de objetos detectados sobre objetos conocidos del 21.1% de un total promedio de objetos con propiedades de la imagen de 1257. En resumen la técnica multiumbralización automática 8 ventanas fue la que mostró los mejores rendimientos en cuanto a las variables de sensibilidad (S) y detección de proteínas conocidas (VP), tal como se muestra en los valores dados en las Tablas 4-7, 4-8 y en las Figuras 4-11, 4-12, 4-13, 4-14, 4-18, 4-19, 4-20, 4-21 y 4-22.

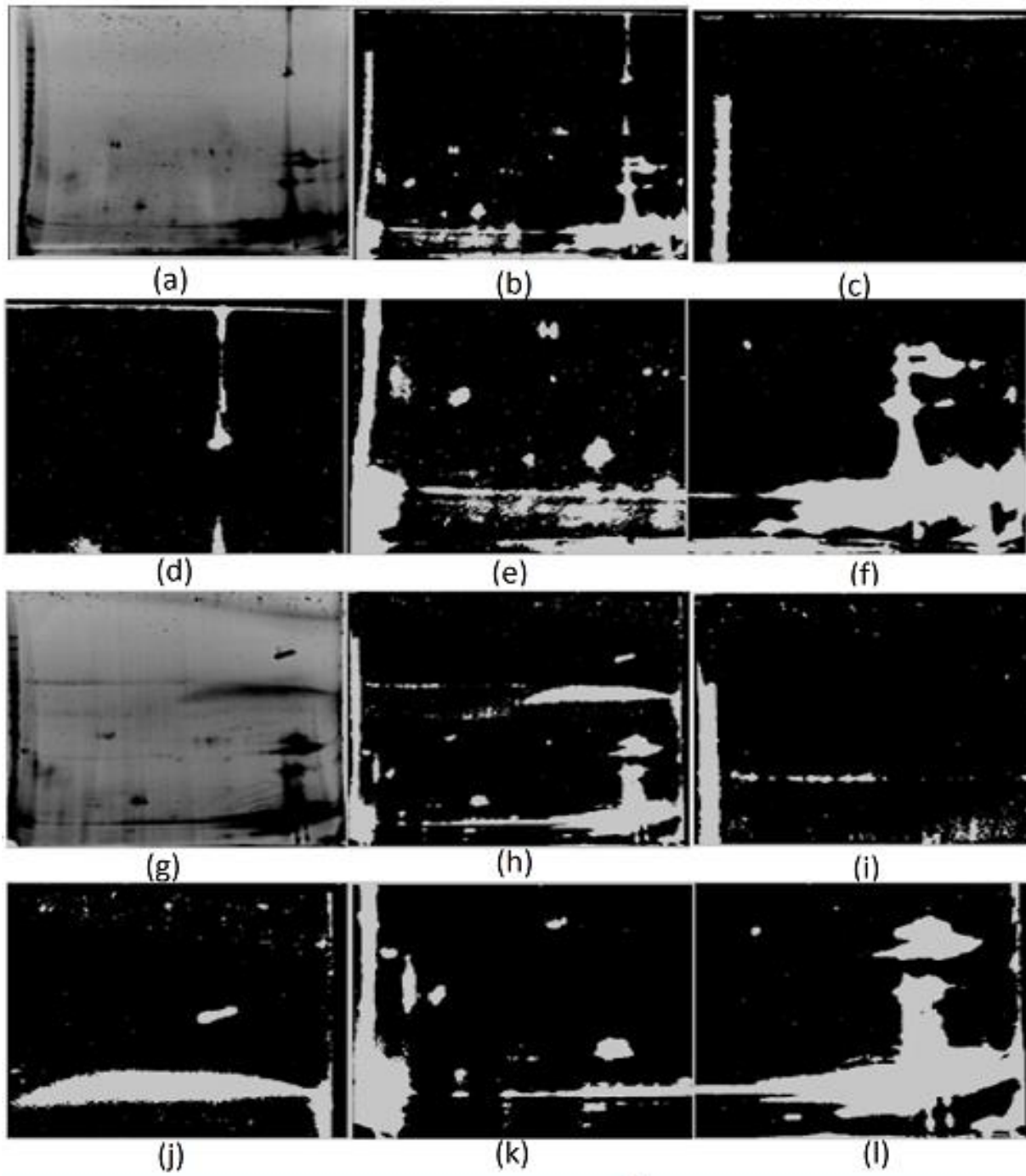
### **4.3 Resultados comparación imágenes apitoxina fusionadas.**

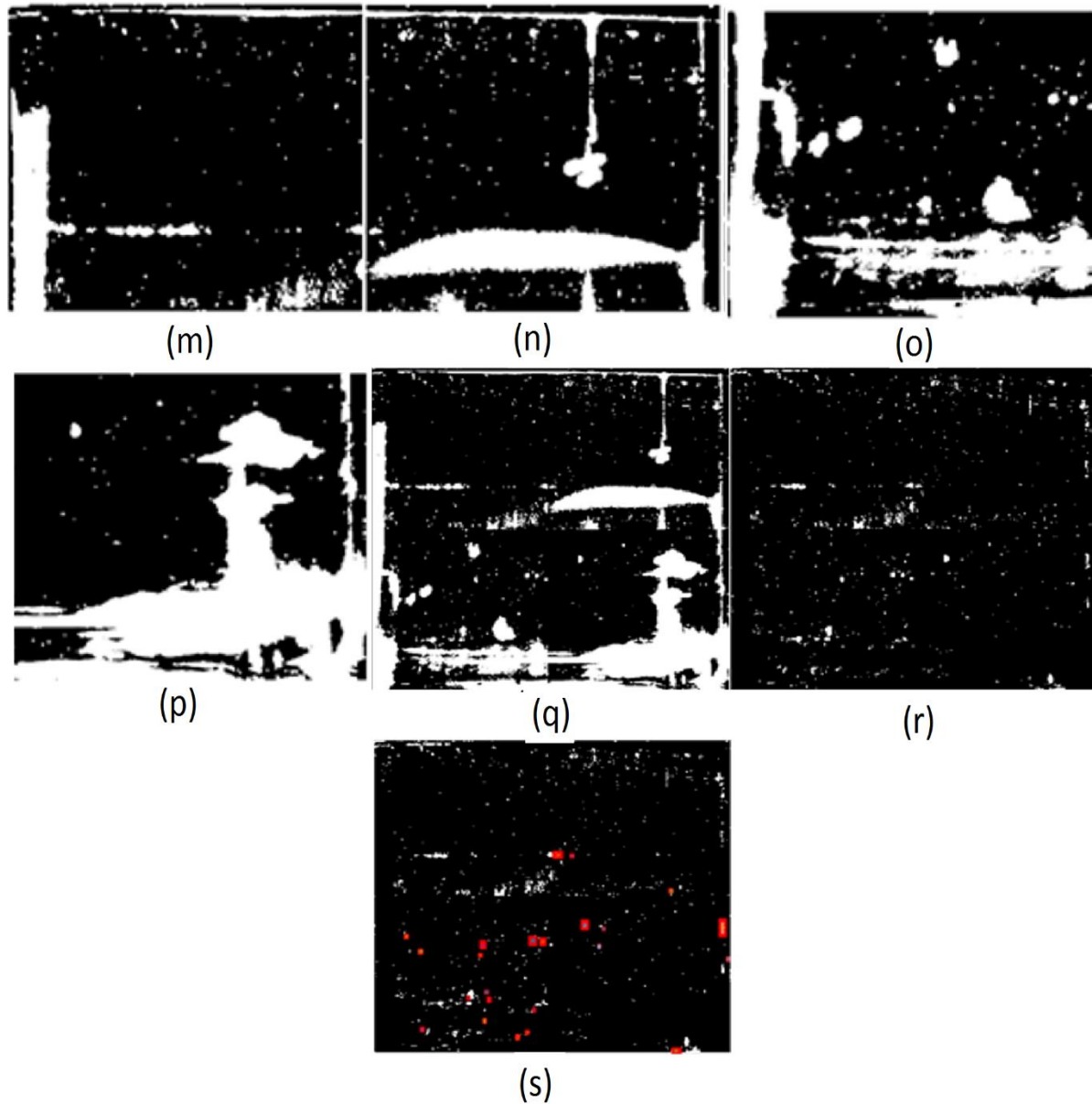
Las 23 imágenes consolidadas en los procesos de comparación de las técnicas de segmentación y relacionadas en la Tabla 4-5 y en las Figuras 4-4; 4-5, fueron fusionadas por pares y por grupos de interior y de exterior y de acuerdo con el lugar de procedencia, como lo son las veredas La Esmeralda, La Isla y La Loma. La Figura 4-23 muestra el proceso de fusión de las imágenes MII2BOX y MLI1BOX, que fueron las muestras en donde se encontró mayor cantidad de proteínas comparadas con el patrón definido en la Tabla 4-6. Inicialmente se muestran las imágenes originales MII2BOX y MLI1BOX (imágenes (a) y (g)) sometidas al proceso de fusión, seguidas de las imágenes resultantes después de ejecutar el proceso de multiumbralización 4 ventanas (imágenes (b); (h)), así mismo se muestran las 4 ventanas en que se dividen las dos imágenes (imágenes (c); (d); (e); (f); (i); (j), (k); (l)) seguidas de la fusión de cada ventana correspondiente de ambas imágenes procesadas (imágenes (m); (n); (o); (p)) y por último las dos imágenes fusionadas (imagen (q)), seguida de la misma imagen eliminando áreas mayores a 1000 pixeles (imagen (r)) y por último la imagen en donde se observan los objetos detectados resaltados con color rojo, resultado de comparar el total de objetos de la imagen fusionada con los objetos del patrón referenciado en la Tabla 4-6 (imagen (s)).

Las Tablas 4-9 y 4-10 muestran los resultados obtenidos en la fusión de las imágenes. En la Tabla 4-9 que presenta las imágenes de las muestras de interior de La Isla y La Loma fusionadas, el valor mayor promedio de sensibilidad está dado en el proceso de la técnica umbralización automática

con un promedio de relación de objetos conocidos detectados con respecto a los objetos conocidos de la matriz patrón de 24.7%, seguida de la técnica multiumbralización 8 ventanas con un 23.6%; multiumbralización 4 ventanas con un 22.6% y técnica multiumbralización 16 ventanas con un 22.2%. Los mejores porcentajes promedios de fusión a nivel de imágenes individuales los tuvo la imagen muestra MII123SOBODOXR, que corresponde a las imágenes fusionadas muestra Isla Interior 1 súper oscura, muestra Isla Interior 1 demasiado oscura, muestra Isla Interior 2 bastante oscura, muestra Isla Interior 2 demasiado oscura y muestra Isla Interior 3 bastante oscura, cuyo promedio total de sensibilidad fue del 44.1% en la relación establecida entre los objetos conocidos detectados y los objetos conocidos del patrón.

Los promedios de la Tabla 4-10 que presentan las imágenes de las muestras de exterior de La Esmeralda, La Isla y La Loma, varían un poco con respecto a los promedios dados para las imágenes de interior, donde el promedio mayor de sensibilidad está dado en el proceso de fusión llevado a cabo por la técnica de multiumbralización automática 4 ventanas con un 17.2% de relación de objetos conocidos detectados con respecto a los objetos conocidos de la matriz patrón, seguida de la técnica multiumbralización automática 8 ventanas con un 16.4%; multiumbralización automática 16 ventanas con un 16% y técnica umbralización automática con un 15.7%.





**Figura 4-23:** Imágenes resultantes del proceso de fusión en dos muestras seleccionadas.(a) MII2BOX; (b) MII2BOX multiumbral 4 ventanas; (c) MII2BOX ventana 1; (d) MII2BOX ventana 2; (e) MII2BOX ventana 3; (f) MII2BOX ventana 4; (g) MLI1BOX; (h) MLI1BOX multiumbral 4 ventanas; (i) MLI1BOX ventana 1; (j) MLI1BOX ventana 2; (k) MLI1BOX ventana 3; (l) MLI1BOX ventana 4; (m) Fusión MII2BOX-MLI1BOX ventanas 1; (n) Fusión MII2BOX-MLI1BOX ventanas 2; (o) Fusión MII2BOX-MLI1BOX ventanas 3;(p) Fusión MII2BOX-MLI1BOX ventanas 4; (q) Fusión MII2BOX-MLI1BOX; (r) Fusión MII2BOX-MLI1BOX sin áreas mayores a 100 píxeles; (s) Detección de objetos en imágenes fusionadas

**Tabla 4-9:** Resumen general de comparación técnicas Umbralización automática, Multiumbralización en 4 - 8 y 16 ventanas para las imágenes fusionadas de interior muestras La Isla y La loma, con respecto al número de objetos conocidos definidos en el patrón de apitoxina mostrados en la Tabla 4-6.

<b>TÉCNICA UMBRALIZACIÓN AUTOMÁTICA IMÁGENES DE INTERIOR</b>									
<b>Imágenes Fusionadas OC=115</b>		<b>NS</b>	<b>VP</b>	<b>FP</b>	<b>FN</b>	<b>P (%)</b>	<b>RF (%)</b>	<b>RFN (%)</b>	<b>S (%)</b>
La Isla	MII1SODOXR	265	19	246	96	7,2%	92,8%	83,5%	16,5%
	MII2BODOXR	1570	33	1537	82	2,1%	97,9%	71,3%	28,7%
	MII13SOBODOXR	345	29	316	86	8,4%	91,6%	74,8%	25,2%
	MII123SOBODOXR	1830	45	1785	70	2,5%	97,5%	60,9%	39,1%
La Loma	MLI1BODOXR	2642	35	2607	80	1,3%	98,7%	69,6%	30,4%
	MLI1ODOX1R	236	9	227	106	3,8%	96,2%	92,2%	7,8%
	MLI2BODOXR	281	9	272	106	3,2%	96,8%	92,2%	7,8%
	MLI3BODOXR	1522	26	1496	89	1,7%	98,3%	77,4%	22,6%
	MLI1OBODOXR	2738	38	2700	77	1,4%	98,6%	67,0%	33,0%
	MLI23BODOXR	1725	28	1697	87	1,6%	98,4%	75,7%	24,3%
	MLI123OBODOXR	3889	42	3847	73	1,1%	98,9%	63,5%	36,5%
<b>Promedio Total</b>		1549	28	1521	87	3,1%	96,9%	75,3%	24,7%
<b>TÉCNICA MULTIUMBRAIZACIÓN AUTOMÁTICA 4 VENTANAS IMÁGENES DE INTERIOR</b>									
<b>Imágenes Fusionadas OC=115</b>		<b>NS</b>	<b>VP</b>	<b>FP</b>	<b>FN</b>	<b>P (%)</b>	<b>RF (%)</b>	<b>RFN (%)</b>	<b>S (%)</b>
La Isla	MII1SODOXR	191	19	172	96	9,9%	90,1%	83,5%	16,5%
	MII2BODOXR	842	39	803	76	4,6%	95,4%	66,1%	33,9%
	MII13SOBODOXR	314	25	289	90	8,0%	92,0%	78,3%	21,7%
	MII123SOBODOXR	1110	53	1057	62	4,8%	95,2%	53,9%	46,1%
La Loma	MLI1BODOXR	1114	26	1088	89	2,3%	97,7%	77,4%	22,6%
	MLI1ODOX1R	236	8	228	107	3,4%	96,6%	93,0%	7,0%
	MLI2BODOXR	227	10	217	105	4,4%	95,6%	91,3%	8,7%
	MLI3BODOXR	519	16	503	99	3,1%	96,9%	86,1%	13,9%
	MLI1OBODOXR	1242	28	1214	87	2,3%	97,7%	75,7%	24,3%
	MLI23BODOXR	703	24	679	91	3,4%	96,6%	79,1%	20,9%
	MLI123OBODOXR	1762	38	1724	77	2,2%	97,8%	67,0%	33,0%
<b>Promedio Total</b>		751	26	725	89	4,4%	95,6%	77,4%	22,6%

<b>TÉCNICA MULTIUMBRAIZACIÓN AUTOMÁTICA 8 VENTANAS IMÁGENES DE INTERIOR</b>									
<b>Imágenes Fusionadas OC=115</b>		<b>NS</b>	<b>VP</b>	<b>FP</b>	<b>FN</b>	<b>P (%)</b>	<b>RF (%)</b>	<b>RFN (%)</b>	<b>S (%)</b>
La Isla	MII1SODOXR	204	12	192	103	5,9%	94,1%	89,6%	10,4%
	MII2BODOXR	815	42	773	73	5,2%	94,8%	63,5%	36,5%
	MII13SOBODOXR	271	16	255	99	5,9%	94,1%	86,1%	13,9%
	MII123SOBODOXR	1055	53	1002	62	5,0%	95,0%	53,9%	46,1%
La Loma	MLI1BODOXR	846	31	815	84	3,7%	96,3%	73,0%	27,0%
	MLI1ODOX1R	251	8	243	107	3,2%	96,8%	93,0%	7,0%
	MLI2BODOXR	265	9	256	106	3,4%	96,6%	92,2%	7,8%
	MLI3BODOXR	533	19	514	96	3,6%	96,4%	83,5%	16,5%
	MLI1OBODOXR	987	35	952	80	3,5%	96,5%	69,6%	30,4%
	MLI23BODOXR	757	27	730	88	3,6%	96,4%	76,5%	23,5%
	MLI123OBODOXR	1553	47	1506	68	3,0%	97,0%	59,1%	40,9%
<b>Promedio Total</b>		685	27	658	88	4,2%	95,8%	76,4%	23,6%
<b>TÉCNICA MULTIUMBRAIZACIÓN AUTOMÁTICA 16 VENTANAS IMÁGENES DE INTERIOR</b>									
<b>Imágenes Fusionadas OC=115</b>		<b>NS</b>	<b>VP</b>	<b>FP</b>	<b>FN</b>	<b>P (%)</b>	<b>RF (%)</b>	<b>RFN (%)</b>	<b>S (%)</b>
La Isla	MII1SODOXR	171	13	158	102	7,6%	92,4%	88,7%	11,3%
	MII2BODOXR	1070	40	1030	75	3,7%	96,3%	65,2%	34,8%
	MII13SOBODOXR	227	19	208	96	8,4%	91,6%	83,5%	16,5%
	MII123SOBODOXR	1271	52	1219	63	4,1%	95,9%	54,8%	45,2%
La Loma	MLI1BODOXR	1418	32	1386	83	2,3%	97,7%	72,2%	27,8%
	MLI1ODOX1R	211	8	203	107	3,8%	96,2%	93,0%	7,0%
	MLI2BODOXR	255	8	247	107	3,1%	96,9%	93,0%	7,0%
	MLI3BODOXR	818	12	806	103	1,5%	98,5%	89,6%	10,4%
	MLI1OBODOXR	1504	35	1469	80	2,3%	97,7%	69,6%	30,4%
	MLI23BODOXR	1027	19	1008	96	1,9%	98,1%	83,5%	16,5%
	MLI123OBODOXR	2317	43	2274	72	1,9%	98,1%	62,6%	37,4%
<b>Promedio Total</b>		935	26	910	89	3,7%	96,3%	77,8%	22,2%

PROMEDIOS POR IMAGEN DE INTERIOR									
Imágenes Fusionadas OC=115		NS	VP	FP	FN	P (%)	RF (%)	RFN (%)	S (%)
La Isla	MII1SODOXR	208	16	192	99	7,7%	92,3%	86,3%	13,7%
	MII2BODOXR	1074	39	1036	77	3,9%	96,1%	66,5%	33,5%
	MII13SOBODOXR	289	22	267	93	7,7%	92,3%	80,7%	19,3%
	MII123SOBODOXR	1317	51	1266	64	4,1%	95,9%	55,9%	44,1%
La Loma	MLI1BODOXR	1505	31	1474	84	2,4%	97,6%	73,0%	27,0%
	MLI1ODOX1R	234	8,3	225	107	3,5%	96,5%	92,8%	7,2%
	MLI2BODOXR	257	9	248	106	3,5%	96,5%	92,2%	7,8%
	MLI3BODOXR	848	18	830	97	2,5%	97,5%	84,1%	15,9%
	MLI10BODOXR	1618	34	1584	81	2,4%	97,6%	70,4%	29,6%
	MLI23BODOXR	1053	25	1029	91	2,6%	97,4%	78,7%	21,3%
	MLI123OBODOXR	2380	43	2338	73	2,0%	98,0%	63,0%	37,0%
<b>Promedio Total</b>		980	27	953	88	3,8%	96,2%	76,7%	23,3%



**Tabla 4-10:** Resumen general de comparación técnicas Umbralización automática, Multiumbralización en 4 - 8 y 16 ventanas para las imágenes fusionadas de Exterior muestras La Esmeralda, La Isla y La loma, con respecto al número de objetos conocidos definidos en el patrón de apitoxina mostrados en la Tabla 4-6.

<b>TÉCNICA UMBRALIZACIÓN AUTOMÁTICA IMÁGENES DE EXTERIOR</b>									
<b>Imágenes Fusionadas OC=115</b>		<b>NS</b>	<b>VP</b>	<b>FP</b>	<b>FN</b>	<b>P (%)</b>	<b>RF (%)</b>	<b>RFN (%)</b>	<b>S (%)</b>
La Esmeralda	MEE12DOBOXR	525	12	513	103	2,3%	97,7%	89,6%	10,4%
	MEE122DOBOXR	762	20	742	95	2,6%	97,4%	82,6%	17,4%
La Isla	MIE23OXR	221	9	212	106	4,1%	95,9%	92,2%	7,8%
La Loma	MLE1OXBOX1R	1075	27	1048	88	2,5%	97,5%	76,5%	23,5%
	MLE23BODEEXR	244	9	235	106	3,7%	96,3%	92,2%	7,8%
	MLE4BODOXR	199	7	192	108	3,5%	96,5%	93,9%	6,1%
	MLE123OBODEEXR	937	29	908	86	3,1%	96,9%	74,8%	25,2%
	MLE1234OBODODEEXR	1029	31	998	84	3,0%	97,0%	73,0%	27,0%
<b>Promedio Total</b>		624	18	606	97	3,1%	96,9%	84,3%	15,7%
<b>TÉCNICA MULTIUMBRAIZACIÓN AUTOMÁTICA 4 VENTANAS IMÁGENES DE EXTERIOR</b>									
<b>Imágenes Fusionadas OC=115</b>		<b>NS</b>	<b>VP</b>	<b>FP</b>	<b>FN</b>	<b>P (%)</b>	<b>RF (%)</b>	<b>RFN (%)</b>	<b>S (%)</b>
La Esmeralda	MEE12DOBOXR	498	19	479	96	3,8%	96,2%	83,5%	16,5%
	MEE122DOBOXR	644	23	621	92	3,6%	96,4%	80,0%	20,0%
La Isla	MIE23OXR	255	8	247	107	3,1%	96,9%	93,0%	7,0%
La Loma	MLE1OXBOX1R	1330	26	1304	89	2,0%	98,0%	77,4%	22,6%
	MLE23BODEEXR	250	10	240	105	4,0%	96,0%	91,3%	8,7%
	MLE4BODOXR	268	3	265	112	1,1%	98,9%	97,4%	2,6%
	MLE123OBODEEXR	1564	34	1530	81	2,2%	97,8%	70,4%	29,6%
	MLE1234OBODODEEXR	1754	35	1719	80	2,0%	98,0%	69,6%	30,4%
<b>Promedio Total</b>		820	20	801	95	2,7%	97,3%	82,8%	17,2%
<b>TÉCNICA MULTIUMBRAIZACIÓN AUTOMÁTICA 8 VENTANAS IMÁGENES DE EXTERIOR</b>									
<b>Imágenes Fusionadas OC=115</b>		<b>NS</b>	<b>VP</b>	<b>FP</b>	<b>FN</b>	<b>P (%)</b>	<b>RF (%)</b>	<b>RFN (%)</b>	<b>S (%)</b>
La Esmeralda	MEE12DOBOXR	422	8	414	107	1,9%	98,1%	93,0%	7,0%
	MEE122DOBOXR	629	17	612	98	2,7%	97,3%	85,2%	14,8%
La Isla	MIE23OXR	278	11	267	104	4,0%	96,0%	90,4%	9,6%
La Loma	MLE1OXBOX1R	1096	31	1065	84	2,8%	97,2%	73,0%	27,0%
	MLE23BODEEXR	288	9	279	106	3,1%	96,9%	92,2%	7,8%
	MLE4BODOXR	281	3	278	112	1,1%	98,9%	97,4%	2,6%
	MLE123OBODEEXR	1334	35	1299	80	2,6%	97,4%	69,6%	30,4%
	MLE1234OBODODEEXR	1546	37	1509	78	2,4%	97,6%	67,8%	32,2%
<b>Promedio Total</b>		734	19	715	96	2,6%	97,4%	83,6%	16,4%

<b>TÉCNICA MULTIUMBRALIZACIÓN AUTOMÁTICA 16 VENTANAS IMÁGENES DE EXTERIOR</b>									
<b>Imágenes Fusionadas OC=115</b>		<b>NS</b>	<b>VP</b>	<b>FP</b>	<b>FN</b>	<b>P (%)</b>	<b>RF (%)</b>	<b>RFN (%)</b>	<b>S (%)</b>
La Esmeralda	MEE12DOBOXR	431	14	417	101	3,2%	96,8%	87,8%	12,2%
	MEE122DOBOXR	597	20	577	95	3,4%	96,6%	82,6%	17,4%
La Isla	MIE23OXR	272	12	260	103	4,4%	95,6%	89,6%	10,4%
La Loma	MLE1OXBOX1R	1185	25	1160	90	2,1%	97,9%	78,3%	21,7%
	MLE23BODEEXR	281	9	272	106	3,2%	96,8%	92,2%	7,8%
	MLE4BODOXR	281	3	278	112	1,1%	98,9%	97,4%	2,6%
	MLE123OBODEEXR	1434	31	1403	84	2,2%	97,8%	73,0%	27,0%
	MLE1234OBODODEEXR	1675	33	1642	82	2,0%	98,0%	71,3%	28,7%
<b>Promedio Total</b>		<b>770</b>	<b>18</b>	<b>751</b>	<b>97</b>	<b>2,7%</b>	<b>97,3%</b>	<b>84,0%</b>	<b>16,0%</b>
<b>PROMEDIOS POR IMAGEN DE EXTERIOR</b>									
<b>Imágenes Fusionadas OC=115</b>		<b>NS</b>	<b>VP</b>	<b>FP</b>	<b>FN</b>	<b>P (%)</b>	<b>RF (%)</b>	<b>RFN (%)</b>	<b>S (%)</b>
La Esmeralda	MEE12DOBOXR	469	13	456	102	2,8%	97,2%	88,5%	11,5%
	MEE122DOBOXR	658	20	638	95	3,1%	96,9%	82,6%	17,4%
La Isla	MIE23OXR	257	10	247	105	3,9%	96,1%	91,3%	8,7%
La Loma	MLE1OXBOX1R	1172	27	1144	88	2,4%	97,6%	76,3%	23,7%
	MLE23BODEEXR	266	9	257	106	3,5%	96,5%	92,0%	8,0%
	MLE4BODOXR	257	4	253	111	1,7%	98,3%	96,5%	3,5%
	MLE123OBODEEXR	1317	32	1285	83	2,5%	97,5%	72,0%	28,0%
	MLE1234OBODODEEXR	1501	34	1467	81	2,3%	97,7%	70,4%	29,6%
<b>Promedio Total</b>		<b>737</b>	<b>19</b>	<b>718</b>	<b>96</b>	<b>2,8%</b>	<b>97,2%</b>	<b>83,7%</b>	<b>16,3%</b>

Los mejores resultados individuales fueron dados para imagen fusionada MLE1234OBODODEEXR, que es el resultado de la fusión de las imágenes muestra Loma Exterior 1 oscura, muestra Loma Exterior 1 bastante oscura, muestra Loma Exterior 2 bastante oscura, muestra Loma Exterior 3 dee, muestra Loma Exterior 4 bastante oscura y muestra Loma Exterior 4 demasiado oscura con un promedio total de sensibilidad de 29.6% en la relación establecida entre los objetos conocidos detectados y los objetos conocidos del patrón.

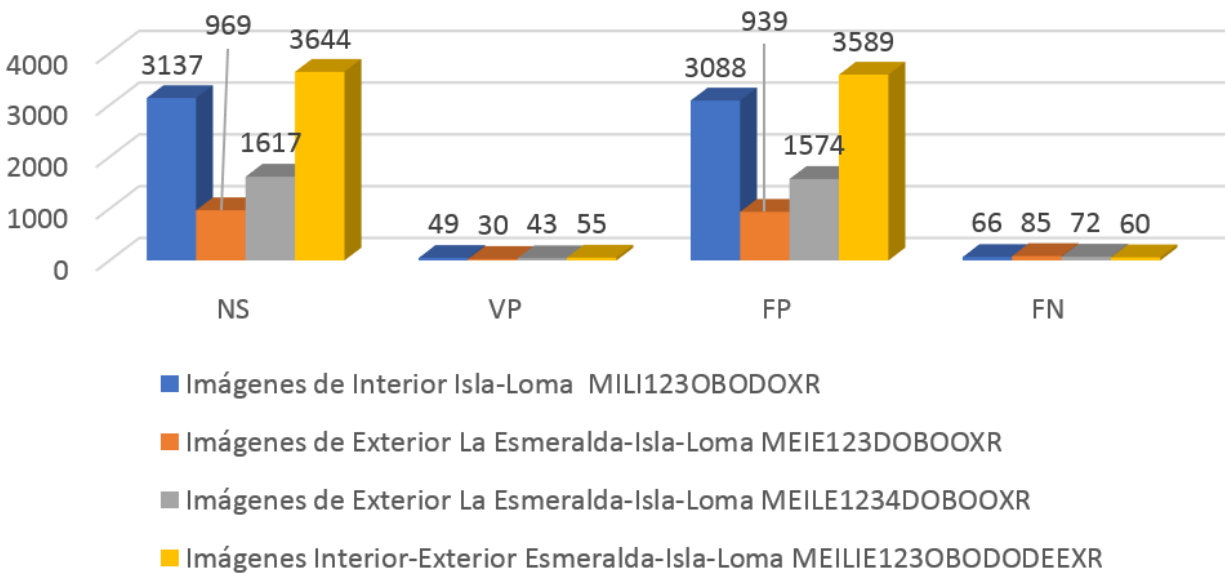
La Tabla 4-11 y las Figuras 4-24, 4-25, 4-26, 4-27, 4-28, 4-29, 4-30 y 4-31 presentan el resumen consolidado de la fusión de las imágenes de interior, de exterior y el total integrado de interior-exterior. La técnica multiumbralización automática 16 ventanas presenta los mejores resultados en la variable de detección de proteínas conocidas (VP) con un valor de 79 proteínas y en la variable sensibilidad con un valor de 68.7%, seguida en orden descendente de las técnicas multiumbralización automática 8 ventanas, 4 ventanas y umbralización automática.

**Tabla 4-11:** Consolidado general de comparación de técnicas de segmentación

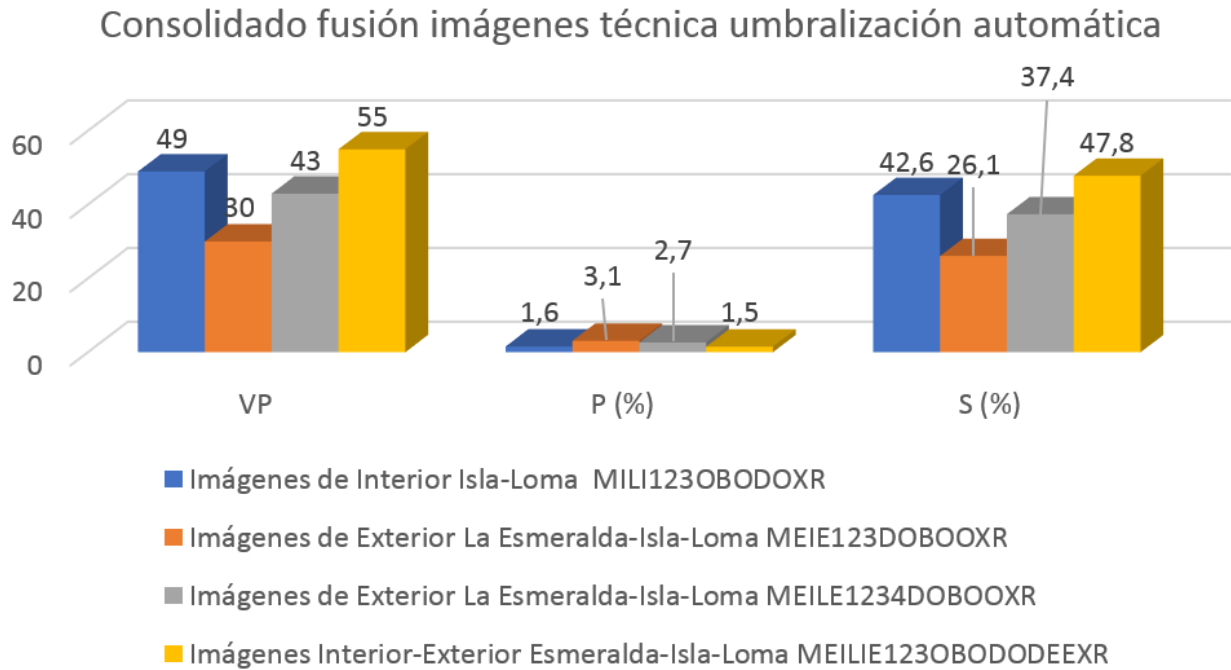
<b>TÉCNICA UMBRALIZACIÓN AUTOMÁTICA</b>									
<b>Imágenes Fusionadas OC=115</b>		<b>NS</b>	<b>VP</b>	<b>FP</b>	<b>FN</b>	<b>P (%)</b>	<b>RF (%)</b>	<b>RFN (%)</b>	<b>S (%)</b>
Imágenes de Interior Isla-Loma	MILI123OBODOXR	3137	49	3088	66	1,6%	98,4%	57,4%	42,6%
Imágenes de Exterior La Esmeralda-Isla-Loma	MEIE123DOBOOXR	969	30	939	85	3,1%	96,9%	73,9%	26,1%
	MEIE1234DOBOOXR	1617	43	1574	72	2,7%	97,3%	62,6%	37,4%
Imágenes Interior-Exterior Esmeralda-Isla-Loma	MEILIE123OBODODEEXR	3644	55	3589	60	1,5%	98,5%	52,2%	47,8%
<b>TÉCNICA MULTIUMBRAIZACIÓN AUTOMÁTICA 4 VENTANAS</b>									
<b>Imágenes Fusionadas OC=115</b>		<b>NS</b>	<b>VP</b>	<b>FP</b>	<b>FN</b>	<b>P (%)</b>	<b>RF (%)</b>	<b>RFN (%)</b>	<b>S (%)</b>
Imágenes de Interior Isla-Loma	MILI123OBODOXR	2650	63	2587	52	2,4%	97,6%	45,2%	54,8%
Imágenes de Exterior La Esmeralda-Isla-Loma	MEIE123DOBOOXR	878	30	848	85	3,4%	96,6%	73,9%	26,1%
	MEIE1234DOBOOXR	2491	53	2438	62	2,1%	97,9%	53,9%	46,1%
Imágenes Interior-Exterior Esmeralda-Isla-Loma fusionada	MEILIE123OBODODEEXR	4277	73	4204	42	1,7%	98,3%	36,5%	63,5%
<b>TÉCNICA MULTIUMBRAIZACIÓN AUTOMÁTICA 8 VENTANAS</b>									
<b>Imágenes Fusionadas OC=115</b>		<b>NS</b>	<b>VP</b>	<b>FP</b>	<b>FN</b>	<b>P (%)</b>	<b>RF (%)</b>	<b>RFN (%)</b>	<b>S (%)</b>
Imágenes de Interior Isla-Loma	MILI123OBODOXR	2350	66	2284	49	2,8%	97,2%	42,6%	57,4%
Imágenes de Exterior La Esmeralda-Isla-Loma	MEIE123DOBOOXR	876	27	849	88	3,1%	96,9%	76,5%	23,5%
	MEIE1234DOBOOXR	2255	49	2206	66	2,2%	97,8%	57,4%	42,6%
Imágenes Interior-Exterior Esmeralda-Isla-Loma fusionada	MEILIE123OBODODEEXR	3957	74	3883	41	1,9%	98,1%	35,7%	64,3%

TÉCNICA MULTIUMBRALIZACIÓN AUTOMÁTICA 16 VENTANAS									
Imágenes Fusionadas OC=115		NS	VP	FP	FN	P (%)	RF (%)	RFN (%)	S (%)
Imágenes de Interior Isla-Loma	MILI123OBODOXR	3310	69	3241	46	2,1%	97,9%	40,0%	60,0%
Imágenes de Exterior La Esmeralda-Isla-Loma	MEIE123DOBOOXR	830	29	801	86	3,5%	96,5%	74,8%	25,2%
	MEIE1234DOBOOXR	2320	46	2274	69	2,0%	98,0%	60,0%	40,0%
Imágenes Interior-Exterior Esmeralda-Isla-Loma fusionada	MEILIE123OBODODEEXR	4843	79	4764	36	1,6%	98,4%	31,3%	68,7%

Consolidado fusión imágenes técnica umbralización automática

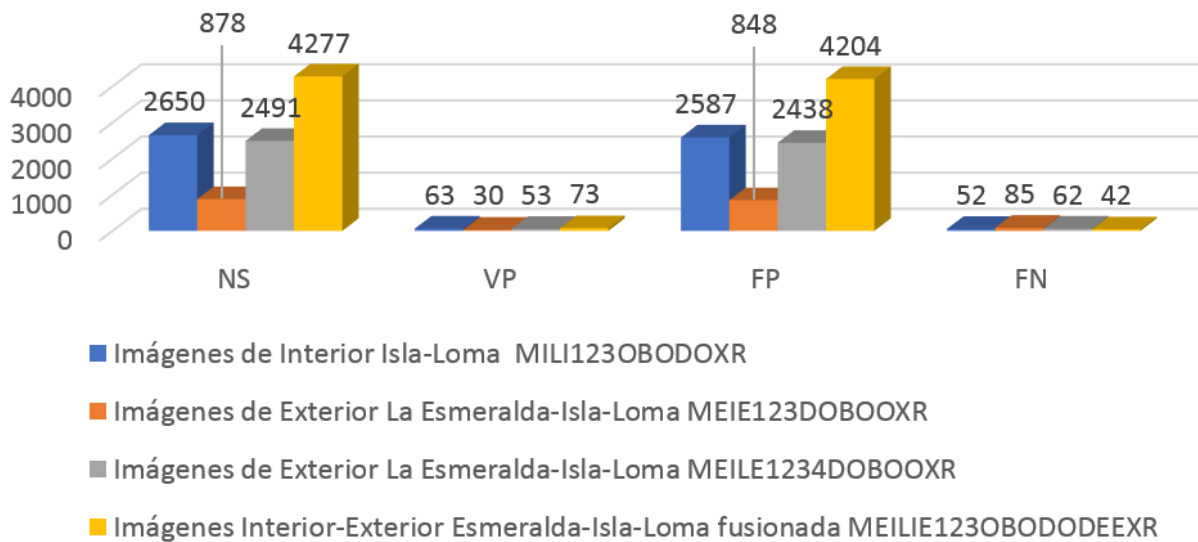


**Figura 4-24:** Relaciones del consolidado total de fusión de objetos segmentados (NS), proteínas de patrón detectadas (VP), objetos segmentados no reconocidos (FP), proteínas de patrón no detectadas (FN) en técnica umbralización automática Tabla 4-11



**Figura 4- 25:** Relaciones del consolidado total de fusión de proteínas de patrón detectadas (VP), precisión (P) y sensibilidad (S) en técnica umbralización automática Tabla 4-11

Consolidado fusión imágenes técnica multiumbralización automática 4 ventanas



**Figura 4-26:** Relaciones del consolidado total de fusión de objetos segmentados (NS), proteínas de patrón detectadas (VP), objetos segmentados no reconocidos (FP), proteínas de patrón no detectadas (FN) en técnica multiumbralización automática 4 ventanas Tabla 4-11

Consolidado fusión imágenes técnica multiumbralización automática 4 ventanas

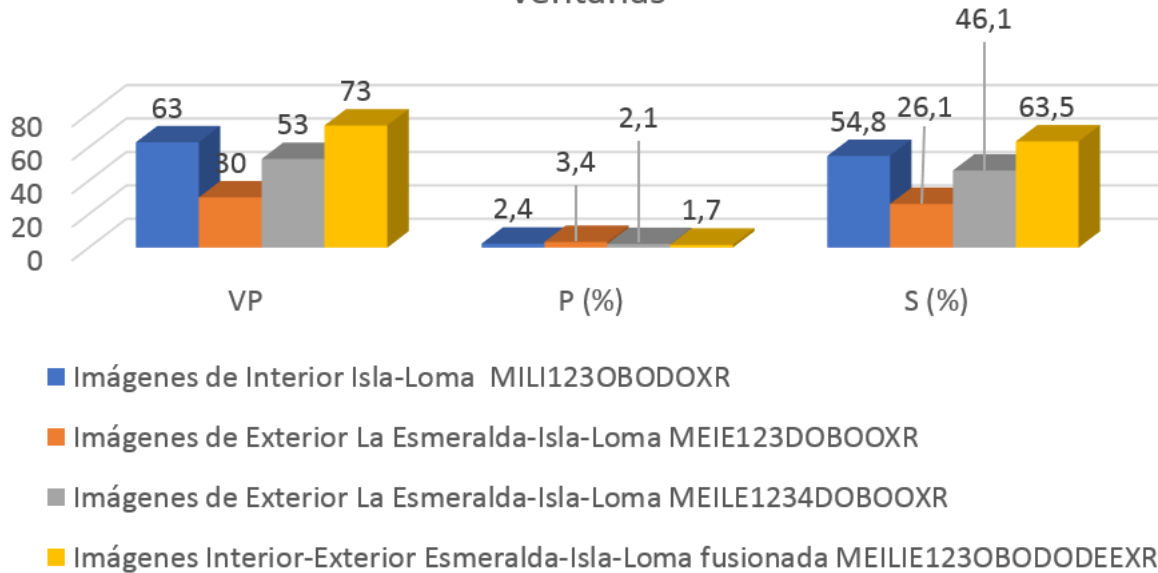


Figura 4-27: Relaciones del consolidado total de fusión de proteínas de patrón detectadas (VP), precisión (P) y sensibilidad (S) en técnica multiumbralización automática 4 ventanas Tabla 4-11

Consolidado fusión imágenes técnica multiumbralización automática 8 ventanas

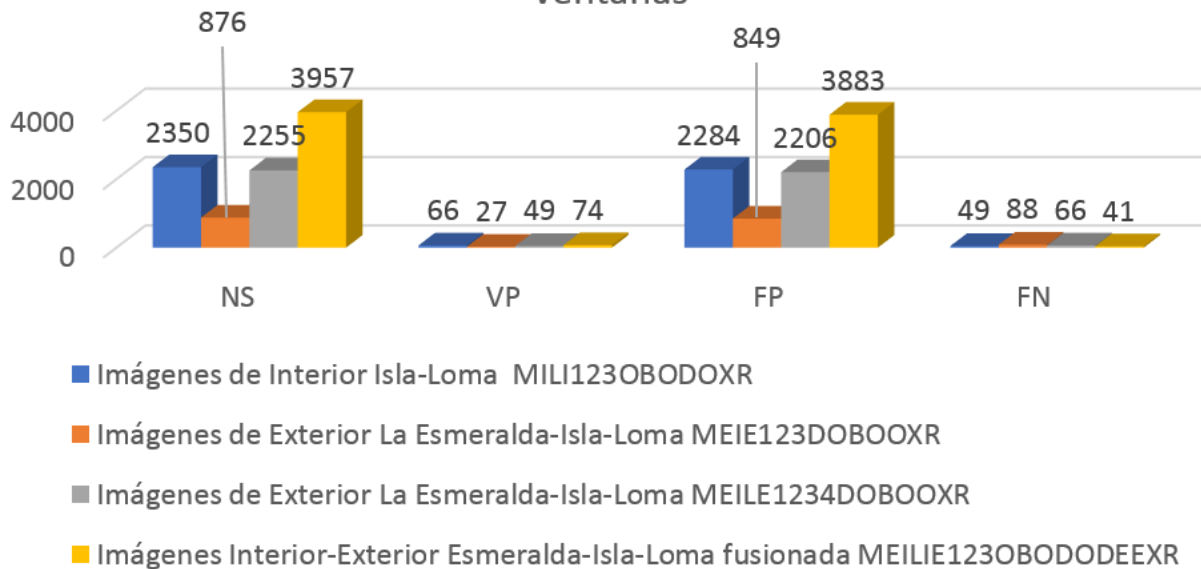


Figura 4-28: Relaciones del consolidado total de fusión de objetos segmentados (NS), proteínas de patrón detectadas (VP), objetos segmentados no reconocidos (FP), proteínas de patrón no detectadas (FN) en técnica multiumbralización automática 8 ventanas Tabla 4-11

Consolidado fusión imágenes técnica multiumbralización automática 8 ventanas

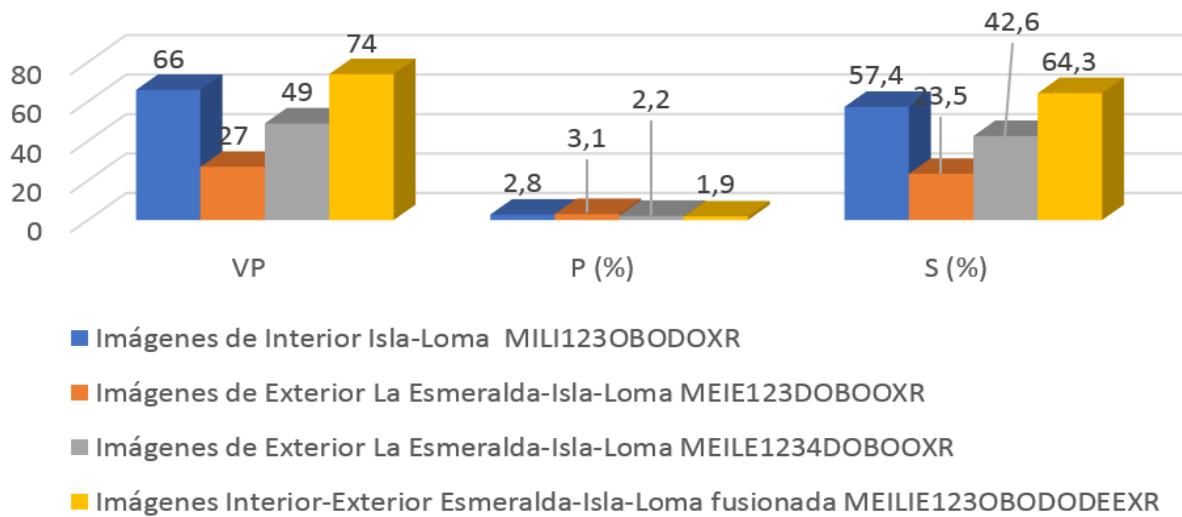


Figura 4-29: Relaciones del consolidado total de fusión de proteínas de patrón detectadas (VP), precisión (P) y sensibilidad (S) en técnica multiumbralización automática 8 ventanas Tabla 4-11

Consolidado fusión imágenes técnica multiumbralización automática 16 ventanas

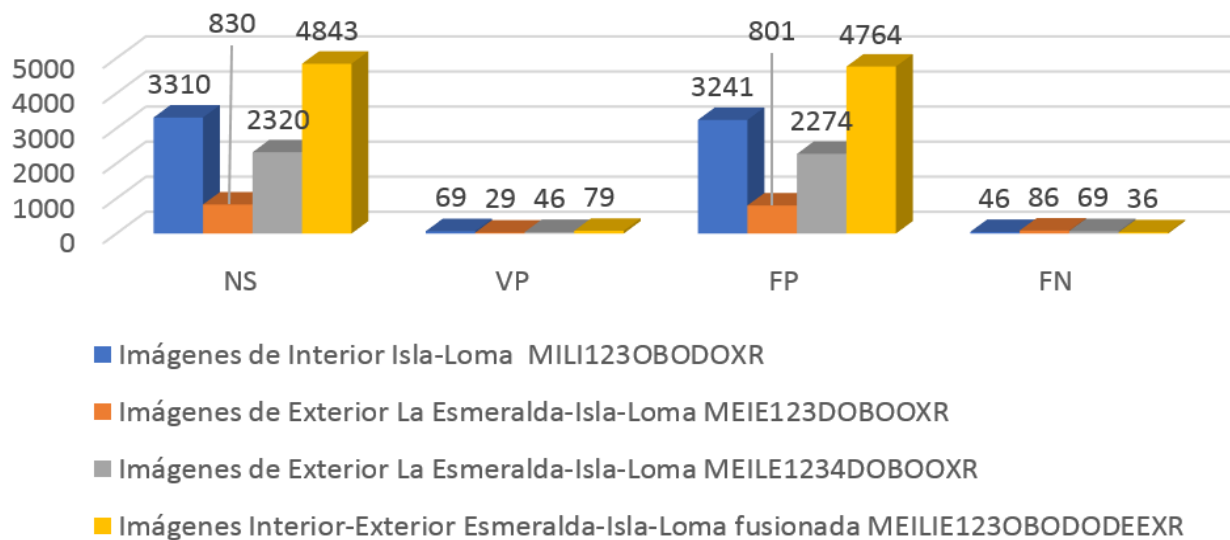
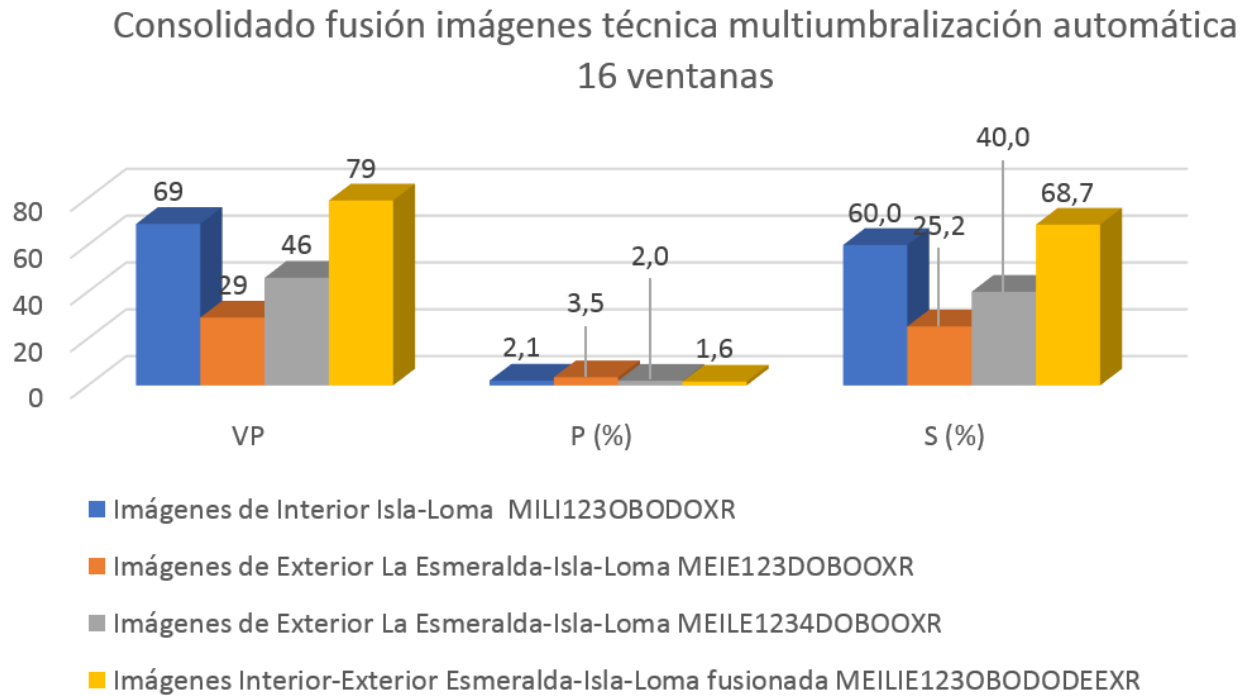


Figura 4-30: Relaciones del consolidado total de fusión de objetos segmentados (NS), proteínas de patrón detectadas (VP), objetos segmentados no reconocidos (FP), proteínas de patrón no detectadas (FN) en técnica multiumbralización automática 16 ventanas Tabla 4-11



**Figura 4-31:** Relaciones del consolidado total de fusión de proteínas de patrón detectadas (VP), precisión (P) y sensibilidad (S) en técnica multiumbralización automática 16 ventanas

En el consolidado total de imágenes fusionadas de interior de la Tabla 4- 11, los promedios mayores fueron dados en la fusión total de las imágenes de interior de La Isla y La Loma con la técnica multiumbralización 16 ventanas con un valor de 60% en la relación de sensibilidad para 69 objetos conocidos detectados y para la fusión de imágenes totales de exterior La Esmeralda, La Isla; La Loma en la técnica de multiumbralización 4 ventanas se calculó un 46.1% en la relación sensibilidad para 53 objetos conocidos detectados. El valor total de sensibilidad en la fusión de las imágenes de interior y de exterior de la Esmeralda, La Isla y La Loma fue de 68.7% en el proceso realizado con la técnica multiumbralización automática 16 ventanas con 79 objetos conocidos.

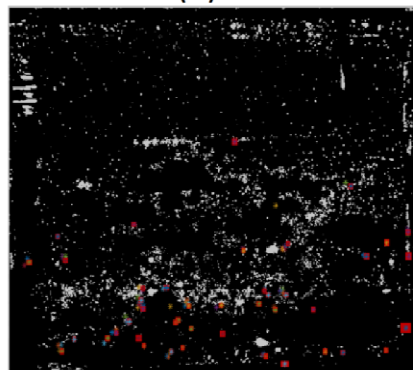




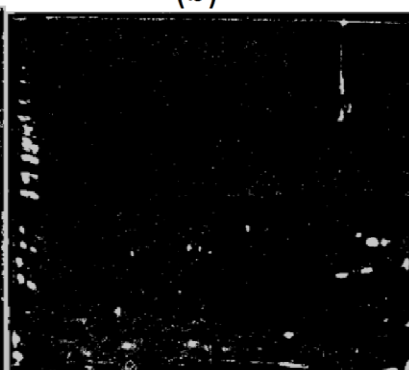
(a)

(b)

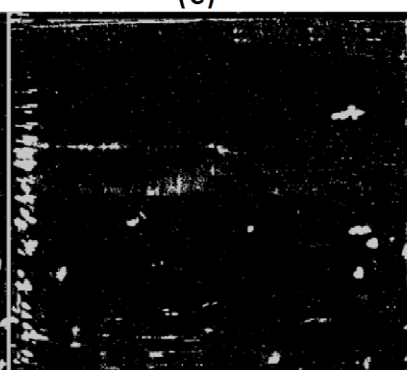
(c)



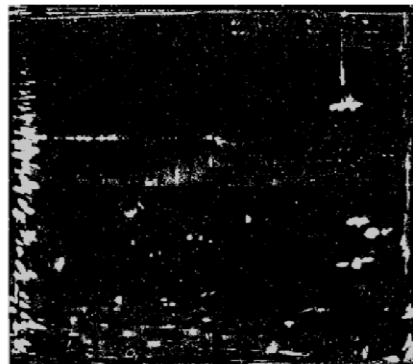
(d)



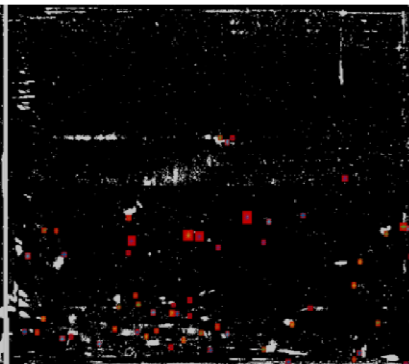
(e)



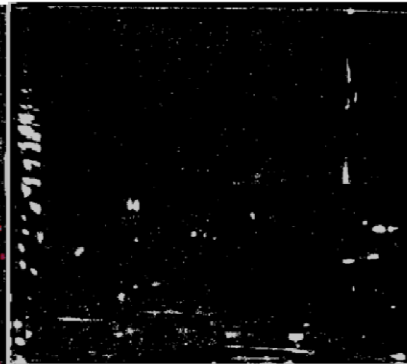
(f)



(g)



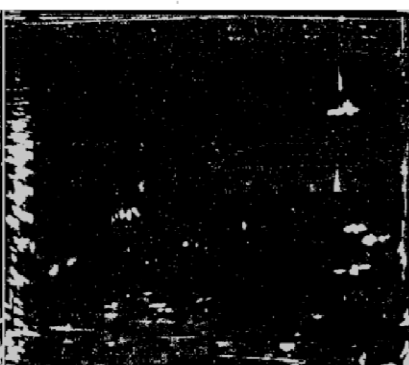
(h)



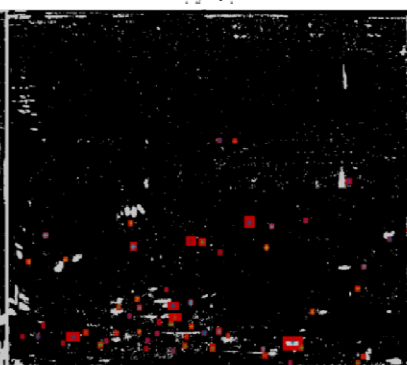
(i)



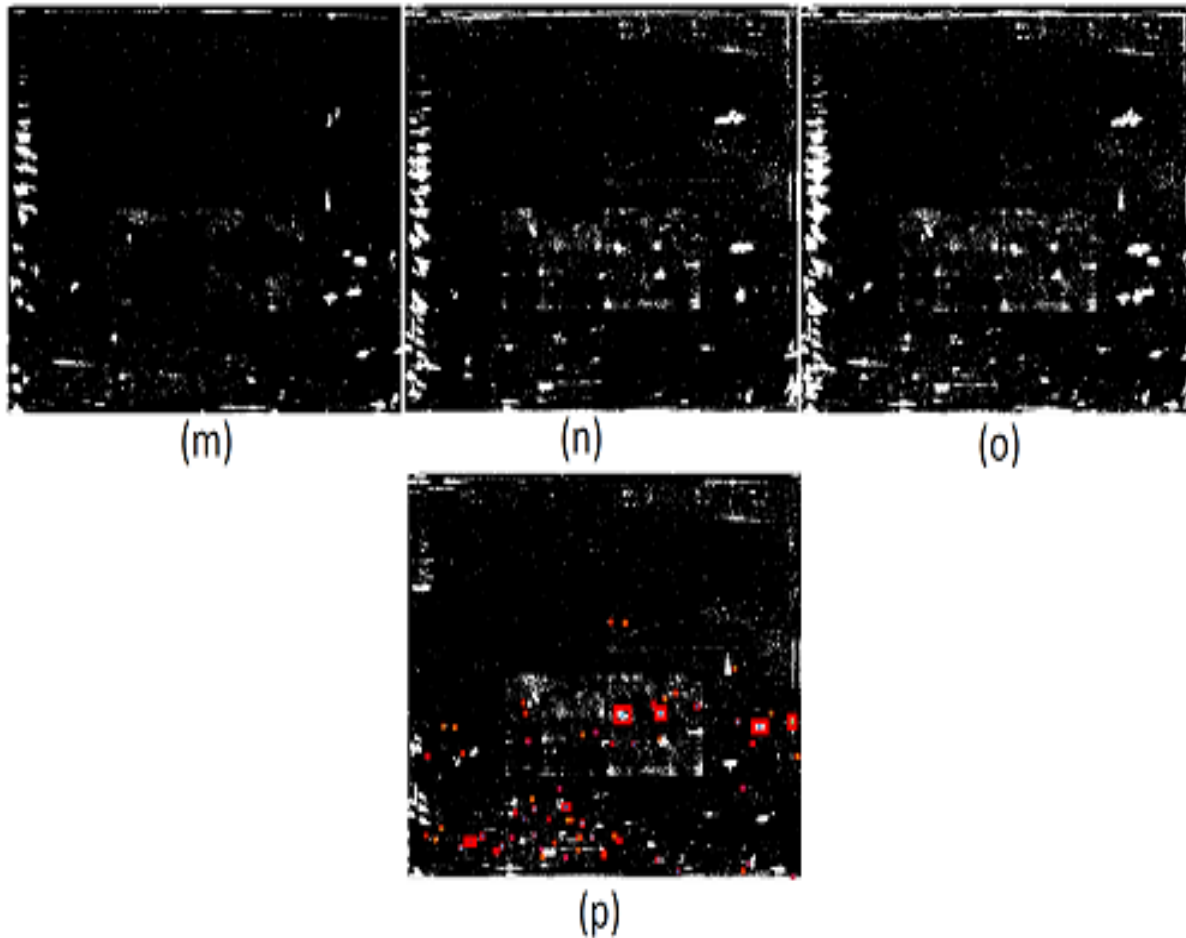
(j)



(k)

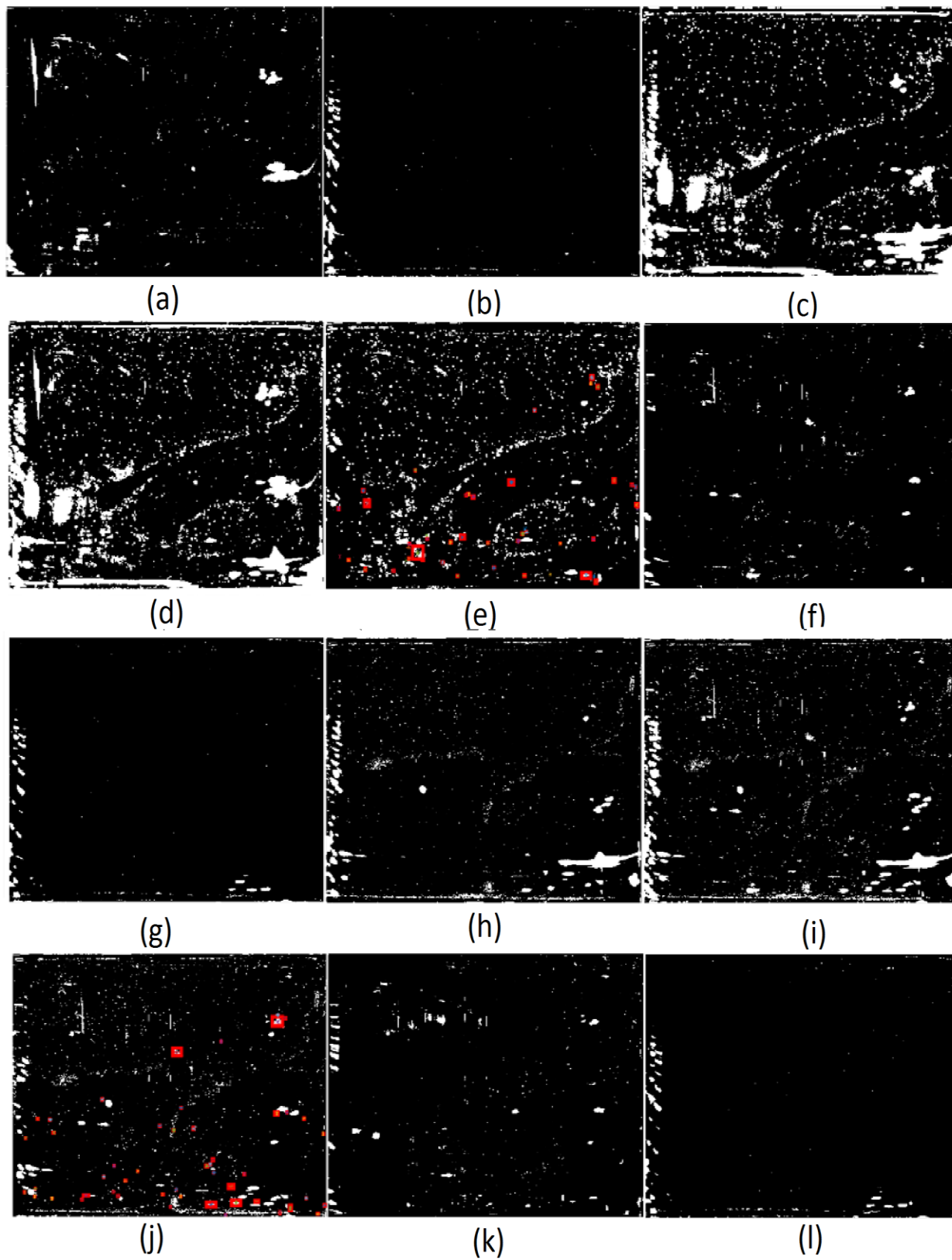


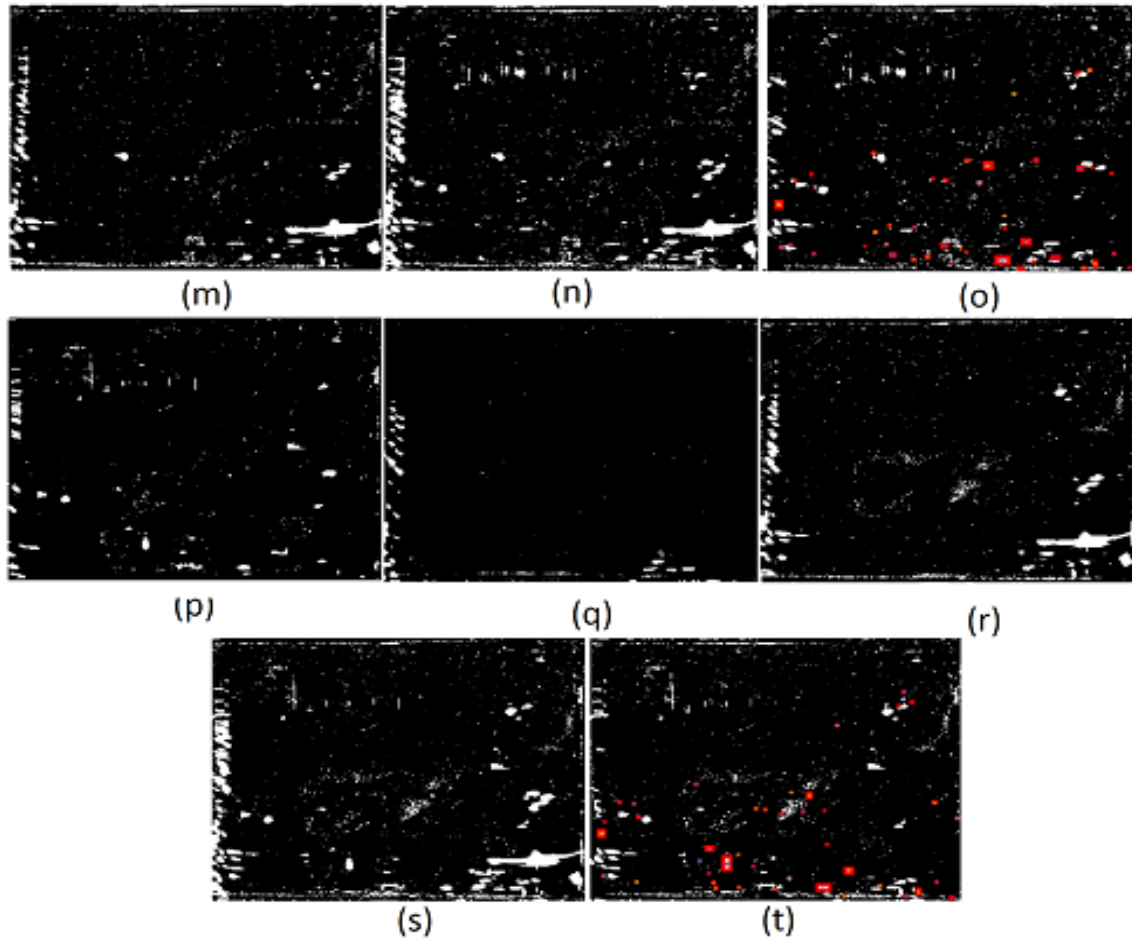
(l)



**Figura 4-32:** Consolidados de la fusión de todas las imágenes de interior. (a) Fusión Isla Interior Umbral automático; (b) Fusión Isla Interior Umbral automático; (c) Fusión Isla- Loma Interior Umbral automático; (d) Detección de objetos fusión Umbral automático; (e) Fusión Isla Interior Multiumbral 4 ventanas; (f) Fusión Loma Interior Multiumbral 4 ventanas; (g) Fusión Isla- Loma Interior Multiumbral 4 ventanas; (h) Detección de objetos fusión multiumbral 4 ventanas; (i) Fusión Isla Interior Multiumbral 8 ventanas; (j) Fusión Loma Interior Multiumbral 8 ventanas; (k) ) Fusión Isla- Loma Interior Multiumbral 8; (l) Detección de objetos fusión multiumbral 8 ventanas; (m) Fusión Isla Interior Multiumbral 16 ventanas; (n) Fusión Loma Interior Multiumbral 16 ventanas; (o) ) Fusión Isla- Loma Interior Multiumbral 16; (p) Detección de objetos fusión multiumbral 16 ventanas

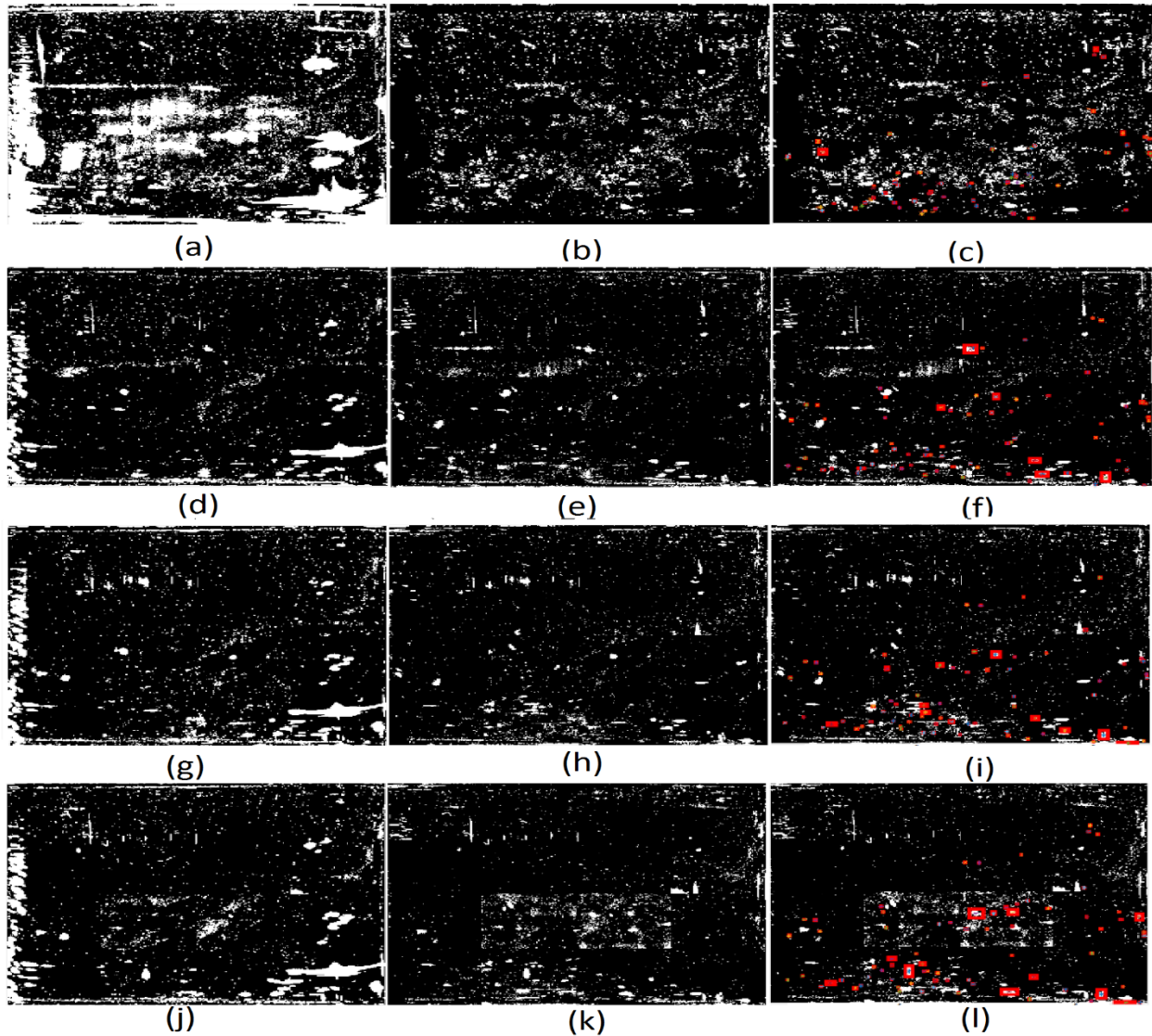
La Figura 4-32 muestra las imágenes de interior La Isla y La Loma fusionadas con las técnicas umbralización automática, multiumbralización 4 - 8 y 16 ventanas, con la cantidad de objetos detectados de 49, 63, 66 y 69 (VP) respectivamente, según datos calculados en la Tabla 4-11.





**Figura 4-33:** Consolidados de la fusión de todas las imágenes de exterior. (a) Fusión Esmeralda Exterior umbral automático; (b) ) Fusión Isla Exterior umbral automático; (c) Fusión Loma Exterior umbral automático; (d) Fusión Total Exterior umbral automático; (e) Detección objetos fusión umbral automático; (f) Fusión Esmeralda Exterior multiumbral 4 ventanas; (g) Fusión Isla Exterior multiumbral 4 ventanas; (h) Fusión Loma Exterior multiumbral 4 ventanas; (i) Fusión Total Exterior multiumbral 4 ventanas; (j) ) Detección objetos fusión multiumbral 4 ventanas; (k) Fusión Esmeralda Exterior multiumbral 8 ventanas; (l) Fusión Isla Exterior multiumbral 8 ventanas; (m) Fusión Loma Exterior multiumbral 8 ventanas; (n) Fusión Total Exterior multiumbral 8 ventanas; (o) ) Detección objetos fusión multiumbral 8 ventanas; (p) Fusión Esmeralda Exterior multiumbral 16 ventanas; (q) Fusión Isla Exterior multiumbral 16 ventanas; (r) Fusión Loma Exterior multiumbral 16 ventanas; (s) Fusión Total Exterior multiumbral 16 ventanas; (t) ) Detección objetos fusión multiumbral 16 ventanas

La Figura 4-33 muestra las imágenes de exterior La Esmeralda, La isla, La Loma fusionadas con las técnicas umbralización automática, multiumbralización 4 - 8 y 16 ventanas, con la cantidad de objetos detectados de 43, 53, 49 y 46 (VP) respectivamente, según datos calculados en la Tabla 4-11.



**Figura 4-34:** Consolidados de la fusión de todas las imágenes de interior y exterior. (a) Fusión Int\_Ext umbral automático todas las áreas; (b) Fusión Int\_Ext umbral automático áreas  $\leq 1000$  píxeles; (c) Detección objetos umbral automático áreas  $\leq 1000$  píxeles; (d) Fusión Int\_Ext multiumbral 4 ventanas todas las áreas; (e) Fusión Int\_Ext multiumbral 4 ventanas áreas  $\leq 1000$  píxeles; (f) Detección objetos multiumbral 4 ventanas áreas  $\leq 1000$  píxeles; (g) ) Fusión Int\_Ext multiumbral 8 ventanas todas las áreas; (h) Fusión Int\_Ext multiumbral 8 ventanas áreas  $\leq 1000$  píxeles; (i) Detección objetos multiumbral 8 ventanas áreas  $\leq 1000$  píxeles; (j) ) Fusión Int\_Ext multiumbral 16 ventanas todas las áreas; (k) Fusión Int\_Ext multiumbral 16 ventanas áreas  $\leq 1000$  píxeles; (l) Detección objetos multiumbral 16 ventanas áreas  $\leq 1000$  píxeles.

La Figura 4-34 muestra las imágenes de interior La Isla, La Loma y de exterior La Esmeralda, La isla, La Loma fusionadas con las técnicas umbralización automática, multiumbralización 4 - 8 y 16 ventanas, con la cantidad de objetos detectados de 55, 73, 74 y 79 (VP) respectivamente, según datos calculados en la

Tabla 4-11. En las primeras imágenes de cada fila (a, d, g, j) se observan las fusiones sin eliminar áreas, en las segundas imágenes de cada fila (b, e, h, k) se han eliminado áreas >1000 píxeles y en las últimas imágenes (c, f, i, l) se observan los objetos conocidos detectados marcados con color rojo, teniendo como referencia los 115 objetos conocidos del patrón.

### 4.3.1 Otra métrica de medición – Coeficiente Sorensen-Dice

El coeficiente ó índice de Sørensen-Dice, también conocido por otros nombres tales como el índice de Sørensen, coeficiente de Dice, es un estadístico utilizado para comparar la similitud de dos muestras. Fue desarrollado independientemente por los botánicos Thorval Sorensen y Lee Raymond Dice, que publicaron en 1957 y 1945 respectivamente (Sørensen, T. (1957), Dice, Lee R. (1945). El coeficiente de similitud DICE (DSC) es un índice de superposición espacial y una métrica de validación de la reproducibilidad. También se llamó la proporción de acuerdo específico por Fleiss. (Fleiss, 1981). El valor de un DSC varía desde 0, lo que indica que no hay superposición espacial entre dos conjuntos de resultados de segmentación binarios, hasta 1, lo que indica una superposición completa.

La fórmula original de Sørensen estaba destinada a ser aplicada a datos de presencia/ausencia, y se define de la siguiente forma:

$$QS = \frac{2C}{A+B} = \frac{2|A \cap B|}{|A|+|B|} ; \quad d = 1 - \frac{2|A \cap B|}{|A|+|B|}$$

Donde A y B son el número de especies en las muestras A y B, respectivamente, y C es el número de especies compartidas por las dos muestras; QS es el cociente de similitud y varía de 0 a 1. Esta expresión se extiende fácilmente a la abundancia en lugar de la presencia / ausencia de especies. Esta versión cuantitativa del índice de Sørensen también se conoce como índice Czekanowski. El índice de Sørensen es idéntico al coeficiente de Dice que siempre está en el rango [0, 1] .

**Tabla 4-12:** Coeficiente Dice para imágenes de interior para técnicas Umbralización y multiumbralización automáticas para 4, 8 y 16 ventanas

RESUMEN GENERAL DE COMPARACION TÉCNICA UMBRALIZACIÓN AUTOMÁTICA IMÁGENES DE INTERIOR								
Imágenes Fusionadas	I1	I1	I2	VP(I1)	VP(I2)	$I1 \cap I2$	Sorensen DICE	Distancia
La Isla	MII1DOX	a	MII123SOBODOXR	11	45	9	0.32	0.68
	MII1SOX	b	MII123SOBODOXR	7	45	4	0.15	0.85
	MII2BOX	c	MII123SOBODOXR	22	45	20	0.60	0.40
	MII2DOX	d	MII123SOBODOXR	13	45	5	0.17	0.83
	MII3BOX	e	MII123SOBODOXR	10	45	6	0.22	0.78
La Loma	MLI1OX	f	MLI123SOBODOXR	4	42	1	0.04	0.96
	MLI1BOX	g	MLI123SOBODOXR	29	42	24	0.68	0.32
	MLI1DOX	h	MLI123SOBODOXR	6	42	3	0.13	0.88
	MLI2BOX	i	MLI123SOBODOXR	9	42	6	0.24	0.76
	MLI2DOX	j	MLI123SOBODOXR	3	42	1	0.04	0.96
	MLI3BOX	k	MLI123SOBODOXR	18	42	9	0.30	0.70
	MLI3DOX	l	MLI123SOBODOXR	11	42	4	0.15	0.85

RESUMEN GENERAL DE COMPARACION TÉCNICA MULTIUMBRAIZACIÓN AUTOMÁTICA 4 VENTANAS IMÁGENES DE INTERIOR								
Imágenes Fusionadas	I1	I1	I2	VP(I1)	VP(I2)	$I1 \cap I2$	Sorensen DICE	Distancia
La Isla	MII1DOX	a	MII123SOBODOXR	20	53	16	0.44	0.56
	MII1SOX	b	MII123SOBODOXR	3	53	0	0.00	1.00
	MII2BOX	c	MII123SOBODOXR	37	53	33	0.73	0.27
	MII2DOX	d	MII123SOBODOXR	7	53	6	0.20	0.80
	MII3BOX	e	MII123SOBODOXR	11	53	9	0.28	0.72
La Loma	MLI1OX	f	MLI123SOBODOXR	2	38	1	0.05	0.95
	MLI1BOX	g	MLI123SOBODOXR	20	38	16	0.55	0.45
	MLI1DOX	h	MLI123SOBODOXR	8	38	6	0.26	0.74
	MLI2BOX	i	MLI123SOBODOXR	12	38	9	0.36	0.64
	MLI2DOX	j	MLI123SOBODOXR	2	38	2	0.10	0.90
	MLI3BOX	k	MLI123SOBODOXR	15	38	11	0.42	0.58
	MLI3DOX	l	MLI123SOBODOXR	4	38	2	0.10	0.90

RESUMEN GENERAL DE COMPARACION TÉCNICA MULTUMBRALIZACIÓN AUTOMÁTICA 8 VENTANAS IMÁGENES DE INTERIOR								
Imágenes Fusionadas	I1	I1	I2	VP(I1)	VP(I2)	$I1 \cap I2$	Sorensen DICE	Distancia
La Isla	MII1DOX	a	MII123SOBODOXR	13	53	10	0.30	0.70
	MII1SOX	b	MII123SOBODOXR	3	53	1	0.04	0.96
	MII2BOX	c	MII123SOBODOXR	39	53	36	0.78	0.22
	MII2DOX	d	MII123SOBODOXR	6	53	5	0.17	0.83
	MII3BOX	e	MII123SOBODOXR	4	53	4	0.14	0.86
La Loma	MLI1OX	f	MLI123OBODOXR	1	47	0	0.00	1.00
	MLI1BOX	g	MLI123OBODOXR	25	47	22	0.61	0.39
	MLI1DOX	h	MLI123OBODOXR	8	47	5	0.18	0.82
	MLI2BOX	i	MLI123OBODOXR	12	47	8	0.27	0.73
	MLI2DOX	j	MLI123OBODOXR	2	47	2	0.08	0.92
	MLI3BOX	k	MLI123OBODOXR	17	47	14	0.44	0.56
	MLI3DOX	l	MLI123OBODOXR	5	47	2	0.08	0.92

RESUMEN GENERAL DE COMPARACION TÉCNICA MULTUMBRALIZACIÓN AUTOMÁTICA 16 VENTANAS IMÁGENES DE INTERIOR								
Imágenes Fusionadas	I1	I1	I2	VP(I1)	VP(I2)	$I1 \cap I2$	Sorensen DICE	Distancia
La Isla	MII1DOX	a	MII123SOBODOXR	14	52	11	0.33	0.67
	MII1SOX	b	MII123SOBODOXR	2	52	2	0.07	0.93
	MII2BOX	c	MII123SOBODOXR	37	52	31	0.70	0.30
	MII2DOX	d	MII123SOBODOXR	7	52	7	0.24	0.76
	MII3BOX	e	MII123SOBODOXR	6	52	5	0.17	0.83
La Loma	MLI1OX	f	MLI123OBODOXR	2	43	1	0.04	0.96
	MLI1BOX	g	MLI123OBODOXR	28	43	25	0.70	0.30
	MLI1DOX	h	MLI123OBODOXR	8	43	5	0.20	0.80
	MLI2BOX	i	MLI123OBODOXR	10	43	7	0.26	0.74
	MLI2DOX	j	MLI123OBODOXR	2	43	1	0.04	0.96
	MLI3BOX	k	MLI123OBODOXR	8	43	6	0.24	0.76
	MLI3DOX	l	MLI123OBODOXR	5	43	1	0.04	0.96



**Tabla 4-13:** Coeficiente Dice para imágenes de Exterior para técnicas Umbralización y multiumbralización automáticas para 4, 8 y 16 ventanas

RESUMEN GENERAL DE COMPARACIÓN TÉCNICA UMBRALIZACIÓN AUTOMÁTICA IMÁGENES DE EXTERIOR								
Imágenes Fusionadas	I1	I1	I2	VP(I1)	VP(I2)	$I1 \cap I2$	Sorensen DICE	Distancia
La Esmeralda	MEE1DOX	m	MEE122DOBOXR	9	20	9	0.62	0.38
	MEE2BOX	n	MEE122DOBOXR	5	20	4	0.32	0.68
	MEE2DOX	o	MEE122DOBOXR	11	20	10	0.65	0.35
La Isla	MIE2OX	p	MIE23OXR	8	9	7	0.82	0.18
	MIE3OX	q	MIE23OXR	4	9	2	0.31	0.69
La Loma	MLE1OX	r	MLE1234OBODODEEXR	6	31	5	0.27	0.73
	MLE1BOX1	s	MLE1234OBODODEEXR	21	31	10	0.38	0.62
	MLE2BOX	t	MLE1234OBODODEEXR	9	31	6	0.30	0.70
	MLE3DEEX	u	MLE1234OBODODEEXR	1	31	1	0.06	0.94
	MLE4BOX	v	MLE1234OBODODEEXR	8	31	8	0.41	0.59
	MLE4DOX	w	MLE1234OBODODEEXR	7	31	7	0.37	0.63

RESUMEN GENERAL DE COMPARACIÓN TÉCNICA MULTIUMBRAIZACIÓN AUTOMÁTICA 4 VENTANAS IMÁGENES DE EXTERIOR								
Imágenes Fusionadas	I1	I1	I2	VP(I1)	VP(I2)	$I1 \cap I2$	Sorensen DICE	Distancia
La Esmeralda	MEE1DOX	m	MEE122DOBOXR	9	23	7	0.44	0.56
	MEE2BOX	n	MEE122DOBOXR	16	23	14	0.72	0.28
	MEE2DOX	o	MEE122DOBOXR	6	23	4	0.28	0.72
La Isla	MIE2OX	p	MIE23OXR	3	8	3	0.55	0.45
	MIE3OX	q	MIE23OXR	5	8	5	0.77	0.23
La Loma	MLE1OX	r	MLE1234OBODODEEXR	1	35	1	0.06	0.94
	MLE1BOX1	s	MLE1234OBODODEEXR	24	35	20	0.68	0.32
	MLE2BOX	t	MLE1234OBODODEEXR	8	35	8	0.37	0.63
	MLE3DEEX	u	MLE1234OBODODEEXR	0	35	0	0.00	1.00
	MLE4BOX	v	MLE1234OBODODEEXR	3	35	3	0.16	0.84
	MLE4DOX	w	MLE1234OBODODEEXR	3	35	2	0.11	0.89

RESUMEN GENERAL DE COMPARACIÓN TÉCNICA MULTIUMBRALIZACIÓN AUTOMÁTICA 8 VENTANAS IMÁGENES DE EXTERIOR								
Imágenes Fusionadas	I1	I1	I2	VP(I1)	VP(I2)	$I1 \cap I2$	Sorensen DICE	Distancia
La Esmeralda	MEE1DOX	m	MEE122DOBOXR	2	17	2	0.21	0.79
	MEE2BOX	n	MEE122DOBOXR	7	17	7	0.58	0.42
	MEE2DOX	o	MEE122DOBOXR	10	17	9	0.67	0.33
La Isla	MIE2OX	p	MIE23OXR	4	11	4	0.53	0.47
	MIE3OX	q	MIE23OXR	6	11	6	0.71	0.29
La Loma	MLE1OX	r	MLE1234OBODODEEXR	1	37	1	0.05	0.95
	MLE1BOX1	s	MLE1234OBODODEEXR	29	37	25	0.76	0.24
	MLE2BOX	t	MLE1234OBODODEEXR	7	37	6	0.27	0.73
	MLE3DEEX	u	MLE1234OBODODEEXR	0	37	0	0.00	1.00
	MLE4BOX	v	MLE1234OBODODEEXR	3	37	3	0.15	0.85
	MLE4DOX	w	MLE1234OBODODEEXR	3	37	3	0.15	0.85

RESUMEN GENERAL DE COMPARACIÓN TÉCNICA MULTIUMBRALIZACIÓN AUTOMÁTICA 16 VENTANAS IMÁGENES DE EXTERIOR								
Imágenes Fusionadas	I1	I1	I2	VP(I1)	VP(I2)	$I1 \cap I2$	Sorensen DICE	Distancia
La Esmeralda	MEE1DOX	m	MEE122DOBOXR	8	20	8	0.57	0.43
	MEE2BOX	n	MEE122DOBOXR	8	20	7	0.50	0.50
	MEE2DOX	o	MEE122DOBOXR	9	20	7	0.48	0.52
La Isla	MIE2OX	p	MIE23OXR	4	12	4	0.50	0.50
	MIE3OX	q	MIE23OXR	7	12	7	0.74	0.26
La Loma	MLE1OX	r	MLE1234OBODODEEXR	2	33	2	0.11	0.89
	MLE1BOX1	s	MLE1234OBODODEEXR	23	33	20	0.71	0.29
	MLE2BOX	t	MLE1234OBODODEEXR	7	33	7	0.35	0.65
	MLE3DEEX	u	MLE1234OBODODEEXR	2	33	2	0.11	0.89
	MLE4BOX	v	MLE1234OBODODEEXR	3	33	3	0.17	0.83
	MLE4DOX	w	MLE1234OBODODEEXR	3	33	3	0.17	0.83

**Tabla 4-14:** Coeficiente Dice consolidado para imágenes de Interior y Exterior para técnicas Umbralización y multiumbralización automáticas para 4, 8 y 16 ventanas

<b>RESUMEN GENERAL DE COMPARACION TÉCNICA UMBRALIZACIÓN AUTOMÁTICA IMÁGENES DE INTERIOR Y EXTERIOR</b>							
<b>Imágenes Fusionadas</b>	<b>I1</b>	<b>I2</b>	<b>VP(I1)</b>	<b>VP(I2)</b>	<b>I1∩I2</b>	<b>SorensenDICE</b>	<b>d</b>
Total Isla Int-Total Interior	MII123SOBODOXR	MILI123SOBODOXR	45	49	33	0.70	0.30
Total Loma Int-Total Interior	MLI123SOBODOXR	MILI123SOBODOXR	42	49	21	0.46	0.54
Total Esmeralda Ext-Total Exterior	MEE122DOBOXR	MEILE123OBODODEEXR	20	42	15	0.48	0.52
Total Isla Ext-Total Exterior	MIE23OXR	MEILE123OBODODEEXR	9	42	6	0.24	0.76
Total Loma Exterior-Total Exterior	MLE1234OBODODEEXR	MEILE123OBODODEEXR	31	42	21	0.58	0.42
Total Interior-Total Int_Ext	MILI123SOBODOXR	MEILE123OBODODEEXR	49	55	41	0.79	0.21
Total Exterior-Total Int_Ext	MEILE123OBODODEEXR	MEILE123OBODODEEXR	42	55	27	0.56	0.44

<b>RESUMEN GENERAL DE COMPARACION TÉCNICA MULTIUMBRAIZACIÓN AUTOMÁTICA 4 VENTANAS IMÁGENES DE INTERIOR Y EXTERIOR</b>							
<b>Imágenes Fusionadas</b>	<b>I1</b>	<b>I2</b>	<b>VP(I1)</b>	<b>VP(I2)</b>	<b>I1∩I2</b>	<b>SorensenDICE</b>	<b>d</b>
Total Isla Int-Total Interior	MII123SOBODOXR	MILI123SOBODOXR	53	63	42	0.72	0.28
Total Loma Int-Total Interior	MLI123SOBODOXR	MILI123SOBODOXR	38	63	34	0.67	0.33
Total Esmeralda Ext-Total Exterior	MEE122DOBOXR	MEILE123OBODODEEXR	23	53	23	0.61	0.39
Total Isla Ext-Total Exterior	MIE23OXR	MEILE123OBODODEEXR	8	53	5	0.16	0.84
Total Loma Exterior-Total Exterior	MLE1234OBODODEEXR	MEILE123OBODODEEXR	35	53	32	0.73	0.27
Total Interior-Total Int_Ext	MILI123SOBODOXR	MEILE123OBODODEEXR	63	73	49	0.72	0.28
Total Exterior-Total Int_Ext	MEILE123OBODODEEXR	MEILE123OBODODEEXR	53	73	33	0.52	0.48

RESUMEN GENERAL DE COMPARACION TÉCNICA MULTITHRESHOLIZACIÓN AUTOMÁTICA 8 VENTANAS IMÁGENES DE INTERIOR Y EXTERIOR							
Imágenes Fusionadas	I1	I2	VP(I1)	VP(I2)	$I1 \cap I2$	SorensenDICE	d
Total Isla Int-Total Interior	MII123SOBODOXR	MILI123SOBODOXR	53	66	41	0.69	0.31
Total Loma Int-Total Interior	MLI123SOBODOXR	MILI123SOBODOXR	47	66	39	0.69	0.31
Total Esmeralda Ext-Total Exterior	MEE122DOBOXR	MEILE123OBODODEEXR	17	49	15	0.45	0.55
Total Isla Ext-Total Exterior	MIE23OXR	MEILE123OBODODEEXR	11	49	8	0.27	0.73
Total Loma Exterior-Total Exterior	MLE1234OBODODEEXR	MEILE123OBODODEEXR	37	49	33	0.77	0.23
Total Interior-Total Int_Ext	MILI123SOBODOXR	MEILE123SOBODODEEXR	66	74	41	0.59	0.41
Total Exterior-Total Int_Ext	MEILE123OBODODEEXR	MEILE123SOBODODEEXR	49	74	35	0.57	0.43

RESUMEN GENERAL DE COMPARACION TÉCNICA MULTITHRESHOLIZACIÓN AUTOMÁTICA 16 VENTANAS IMÁGENES DE INTERIOR Y EXTERIOR							
Imágenes Fusionadas	I1	I2	VP(I1)	VP(I2)	$I1 \cap I2$	SorensenDICE	d
Total Isla Int-Total Interior	MII123SOBODOXR	MILI123SOBODOXR	52	69	45	0.74	0.26
Total Loma Int-Total Interior	MLI123SOBODOXR	MILI123SOBODOXR	43	69	27	0.48	0.52
Total Esmeralda Ext-Total Exterior	MEE122DOBOXR	MEILE123OBODODEEXR	20	46	17	0.52	0.48
Total Isla Ext-Total Exterior	MIE23OXR	MEILE123OBODODEEXR	12	46	9	0.31	0.69
Total Loma Exterior-Total Exterior	MLE1234OBODODEEXR	MEILE123OBODODEEXR	33	46	28	0.71	0.29
Total Interior-Total Int_Ext	MILI123SOBODOXR	MEILE123SOBODODEEXR	69	79	63	0.85	0.15
Total Exterior-Total Int_Ext	MEILE123OBODODEEXR	MEILE123SOBODODEEXR	46	79	33	0.53	0.47

Según lo recomendado por Zijdenbos et al.,(1994) en la literatura de validación de imágenes, se produce una buena superposición cuando  $DSC > 0.7$ .

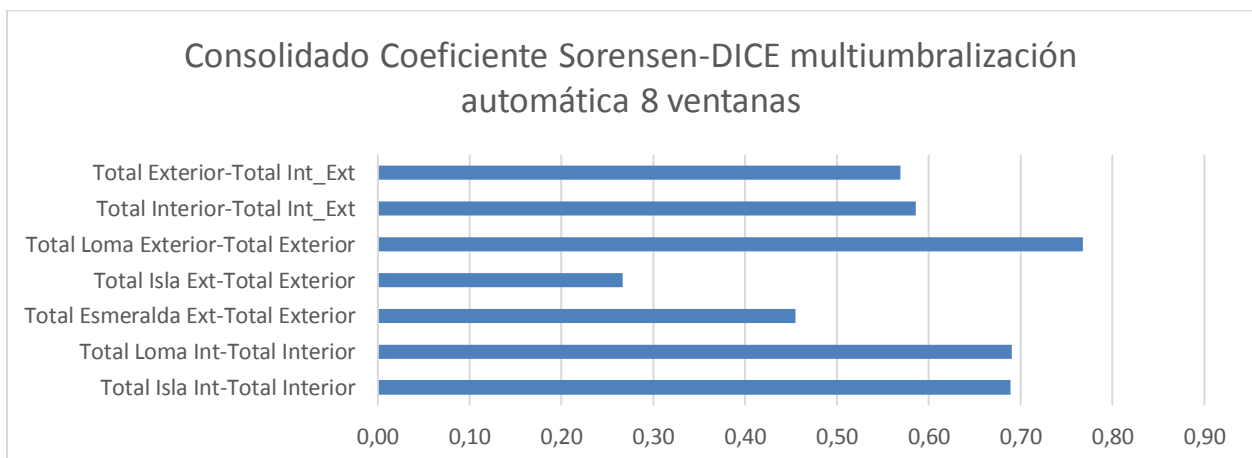
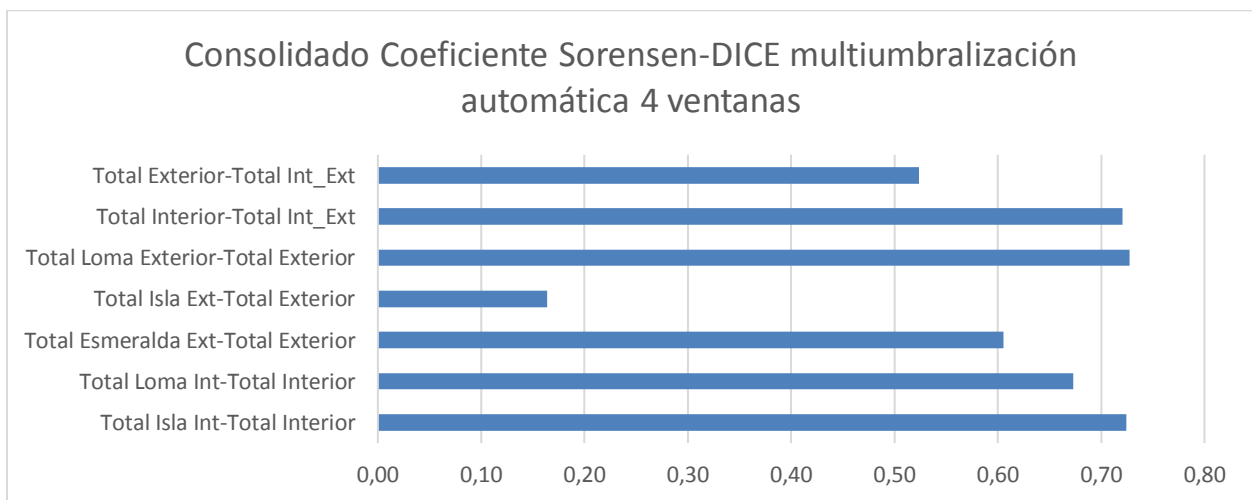
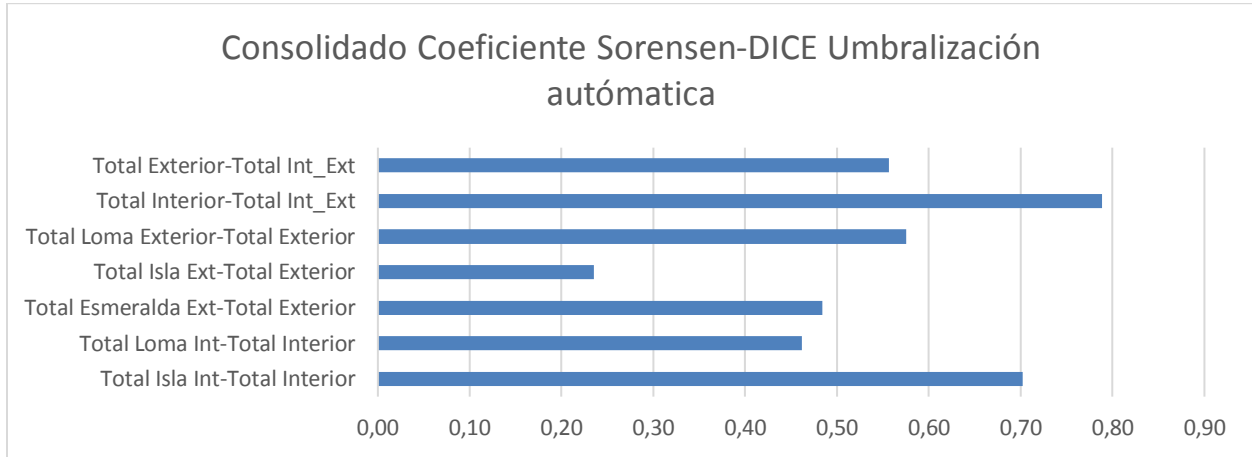
En la Tabla 4-12, se muestran los valores calculados para el coeficiente Dice, para las imágenes de interior aplicando las técnicas de umbralización y multiumbralización automáticas para 4, 8 y 16

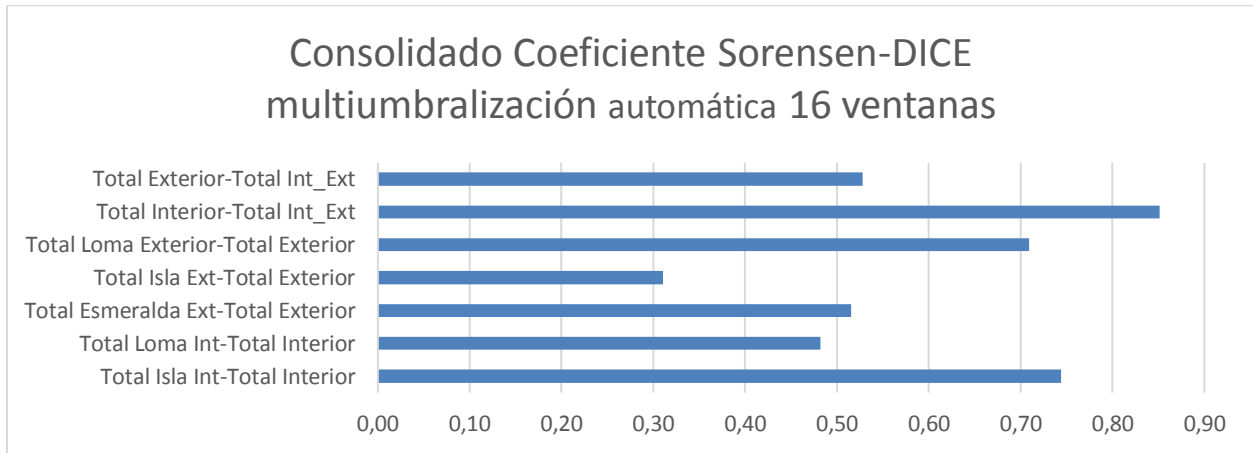
yentanas, el cual resulta de tomar la cantidad de proteínas detectadas del patrón para cada una de las imágenes de interior y compararlas espacialmente con la cantidad de proteínas detectadas en la fusión de todas las imágenes de los grupos de interior (La Isla y La Loma) en cada una de las técnicas de segmentación procesadas para las 12 imágenes (de la a hasta la l). La relación de las proteínas comunes encontradas en los dos conjuntos  $I_1$  e  $I_2$  con respecto a la sumatorias de las proteínas encontradas para los dos conjuntos, da como resultado el coeficiente de similitud existente en los grupos de imágenes procesados. Por ejemplo para el grupo de la Isla, técnica umbralización automática, se toman las imágenes originales estandarizadas  $I_1$  (a,b,c,d,e), en las cuales se detectaron por proceso en la técnica umbralización automática 11, 7, 22 y 13 proteínas respectivamente y se comparan espacialmente con el total de proteínas de las imágenes fusionadas (a,b,c,d,e) detectadas del patrón y que corresponde a la  $I_2$  fusionada y común del grupo de la Isla de Interior que tiene 45 proteínas detectadas del patrón de 113 proteínas. Se hallan las proteínas espacialmente comunes de  $|I_1 \cap I_2|$  que para el caso del grupo de la Isla analizado son 9, 4, 20 y 5 respectivamente y con este valor, relacionado con la sumatoria de  $|I_1| + |I_2|$  se halla el valor del coeficiente.

Los anteriores cálculos también aplican para la Tabla 4-13 que corresponde al cálculo de los coeficientes para imágenes de exterior, que en total son 11 imágenes (desde la m hasta la w) y que corresponden a los grupos de exterior de las veredas La Esmeralda, La Isla y la Loma.

La Tabla 4-14 solamente difiere de las anteriores, en que la relación ya no es establecida entre cada imagen individual con la imagen fusionada del grupo, sino que se relacionan los grupos individuales de interior y de exterior con el total fusionado de su grupo respectivo y el proceso final es la relación de los totales de los dos grupos: de interior y de exterior.

La Figura 4-35 muestra los valores consolidados de los coeficientes Sorensen-Dice para cada una de las técnicas, en donde los valores son muy similares, pero la técnica multiumbralización para 16 ventanas da los mejores resultados con el mayor índice de 0.85 en la relación del total de proteínas fusionadas de imágenes de Interior con el total de proteínas fusionadas en los dos grupos estudiados, además de mostrar un  $DSC > 0.7$  para la relación de las imágenes fusionadas del Total de la Loma exterior con el total de la fusión de las imágenes de exterior (0.71) y el total de las imágenes fusionadas de la Isla interior con el total de la fusión de las imágenes de interior (0.74).





**Figura 4-35:** Coeficiente Sorensen-Dice para el consolidado de las fusiones de imágenes de los grupos de interior y de exterior, en las técnicas de umbralización y multiumbralización para 4-8 y 16 ventanas, según datos de la Tabla 4-14

### 4.3.2 Discusión

La fusión de imágenes nos permite agrupar en una sola imagen todos los objetos comunes y no comunes para un análisis general del comportamiento de grupos. En el caso experimental, al unir dos o más proteínas se puede presentar una concentración de ellas difícil de descifrar, razón por la cual se hace necesario antes del proceso de fusión tratar de eliminar objetos con áreas muy grandes, para poder tener en los resultados finales una desagregación de objetos más fáciles de identificar. Para el caso se eliminaron objetos con áreas mayores a 1000 píxeles y con este proceso inicial sobre las muestras se procedió a realizar la fusión por grupos y total. En el proceso de fusión, todos los pares de imágenes normales procesadas resultantes de la aplicación de las técnicas de segmentación fueron sometidas al cálculo del promedio de píxeles correspondientes, ya que en este proceso se obtenían mayor cantidad de objetos conocidos detectados comparados con los métodos de cálculos promediados de los gradientes dilatación, erosión y morfológico (Los cálculos mencionados se encuentran como rutinas en el proceso de fusión, pero solamente son mostradas en este trabajo los cálculos de píxeles promediados). El consolidado general del proceso de fusión

es mostrado en la Tabla 4-11, en donde se observa que la sensibilidad mayor (cantidad de objetos conocidos detectados con respecto al patrón de apitoxina) se encuentran en la técnica multiumbralización automática para 16 ventanas, encontrándose 69 objetos en la fusión de todas las imágenes de interior y 79 objetos en la integración total de imágenes de exterior y de interior. Para el caso de la fusión total de las imágenes de exterior, la mejor detección fue dada para la técnica multiumbralización automática para 4 ventanas con una sensibilidad de 46.1% y para un VP de 53 objetos conocidos detectados. Para el caso experimental se muestra que entre más objetos se encuentren en la imagen, mayor es la cantidad de objetos detectados, siendo el NS de 4843 objetos para la fusión total; igual sucede con la elaboración de los patrones, ya que entre más objetos tenga reconocidos para el patrón (ver Tabla 4-6), mayor será la probabilidad de encontrar objetos de éste en el conjunto o base de imágenes que sean sometidas a los procesos de análisis y comparación, que para el caso corresponde a una base de 23 imágenes mostradas en las Figuras 4-4 y 4-5. En el consolidado general de fusión la técnica multiumbralización 16 ventanas presentó el mejor comportamiento en cuanto a objetos detectados comparados con el patrón (Tabla 4-6) con un VP de 79 proteínas y con un porcentaje de sensibilidad del 68.7%, tal como lo muestra los valores dados en la Tabla 11 y en la figuras 4-31, 4-32.

El mejor coeficiente de similitud Sorensen-Dice fue dado para la técnica multiumbralización automática para 16 ventanas, tal como se observa en los resultados dados en la Tabla 4-14 y Figura 4-35, en donde se obtuvieron 3 valores del coeficiente DSC  $> 0.7$ .

#### **4.4 Comparación de resultados de imágenes de apitoxina con datos de informe FGN008.**

De las Figuras 4-6, 4-7, 4-8 y 4-9, de las cuales se extrajo el patrón de apitoxina para realizar la comparación de las diferentes técnicas empleadas en este proyecto y de los resultados obtenidos en los procesos llevados a cabo con las diferentes técnicas de segmentación, se establecen las comparaciones dadas en la Tabla 4-15. Esta muestra en la lista de los 115 proteínas propias del patrón, las coordenadas (x, y) de la ubicación de las proteínas, los números indicadores de las proteínas encontradas y referidas en el informe FGN008, para cada uno de los grupos de muestras de imágenes de interior y de exterior (Esmeralda Exterior (EE); Isla Exterior (IE); Loma Exterior



**Tabla 4-15:** Tabla comparativa de spots detectados en las imágenes de apitoxina de interior y de exterior relacionadas con las proteínas registradas en el informe FGN008

COMPARATIVO TÉCNICAS DE SEGMENTACIÓN VS PROTEINAS REGISTRADOS EN EL INFORME FGN008												
#	(x, y)	EE	IE	LE	TE	II	LI	TI	TIE	Gr	Figuras	Ob
1	(942,943)					1-51				IE	15a-16a-16b	CE
2	(1175,985)						2	2	2	I	16b	CE
3	(413,599)					3-20	3-20	3-20	3-20	I	14a-16b	CE
4	(609,663)		4-24	4-24		4-24		4-24	4-24	EI	14a-14b-16b	CE
5	(650,665)		5	5	5	5		5	5	I	16b	CE
6	(41,768)		6	6						E	16a	
7	(117,638)				5-7	5-7		5-7	5-7	E	14b-16a	
8	(175,954)			8	8	8		8	8	E	16a	
9	(219,939)						9	9	9	E	16a	
10	(259,915)					10		10	10	E	16a	
11	(918,349)			11	11				11	E	16a	
12	(186,702)			3-12-49	3-12-49	3-12-49		3-12-49	3-12-49	E	14b-15a-16a	
13	(161,637)			4-13	4-13			4-13	4-13	E	14b-16a	
14	(456,943)	14		14	14	14	14	14	14	E	16a	
15	(429,919)	15		15	15	15	15	15	15	E	16a	
16	(442,842)					16		16	16	E	16a	
17	(494,881)	17			17	17	17	17	17	E	16a	
18	(551,845)		18		18	18	18	18	18	E	16a	
19	(560,869)					19		19	19	I	16a	
20	(435,824)	20			20	20		20	20	E	16a	
21	(439,792)									E	16a	
22	(519,793)					22-50		22-50	22-50	E	15a-16a	
23	(583,962)					24	24	24	24	I	16a	
24	(697,380)					10-26	10-26	10-26	10-26	E	14b-16a	
25	(700,688)				27		27	27	27	I	16a	
26	(780,680)			28	28			28	28	I	16a	
27	(713,378)	11-29			11-29		11-29	11-29	11-29	E	14b-16a	
28	(731,382)									E	14b-16a	
29	(758,382)						13-31	13-31	13-31	E	14b-16a	
30	(945,916)									EI	15a-16a	
31	(883,840)	33			33					I	16a	
32	(921,798)									I	16a	
33	(1015,850)	35			35					E	16a	
34	(1155,789)						22-36	22-36	22-36	E	14b-16a	
35	(1139,491)						37	37	37	E	16a	
36	(1146,619)			6-35-38	6-35-38			6-35-38	6-35-38	E	14a-14b-16a	

#	(x, y)	EE	IE	LE	TE	II	LI	TI	TIE	Gr	Figuras	Ob
37	(1145,270)	27-39		27-39	27-39	27-39				E	14b-16a	
38	(1148,138)									E	14b-16a	
39	(1208,992)	41			41				41	E	16a	
40	(1266,955)			42	42				42	E	16a	
41	(1162,908)									E	16a	
42	(1236,899)									E	16a	
43	(297,952)	1			1		1	1	1	E	14a	
44	(601,831)						2-17			E	14a-14b	
45	(737,603)			9-27	9-27			9-27	9-27	I	14a-14b	
46	(58,924)			2	2	2	2	2	2	I	14a	
47	(97,933)		3		3	3	3	3		I	14a	
48	(351,914)			4		4	4	4	4	E	14a	
49	(869,674)					17-26		17-26	17-26	I	14a-14b	
50	(321,934)	5			5	5	5	5	5	E	14a	
51	(58,891)					6				I	14a	
52	(93,905)			7	7					I	14a	
53	(365,853)					8	8	8	8	I	14a	
54	(95,665)			9		9				I	14a	
55	(873,608)					32		32	32	EI	14a	
56	(394,574)	10			10			10	10	I	14a	
57	(460,963)	11			11		11	11	11	I	14a	
58	(539,970)		12			12	12	12	12	I	14a	
59	(586,929)					13		13	13	I	14a	
60	(432,921)	14		14	14	14	14	14	14	E	14a	
61	(1190,690)							31-39	31-39	EI	14a-14b	
62	(1222,641)					32-40	32-40	32-40		EI	14a-14b	
63	(499,929)					15		15	15	E	14a	
64	(1347,967)	35-43			35-43		35-43	35-43		EI	14a-14b	
65	(1168,640)									EI	14a-14b	
66	(553,885)					16	16	16	16	E	14a	
67	(415,679)					18	18	18	18	I	14a	
68	(607,667)		19	19		19	19	19	19	I	14a	
69	(643,928)	21		21	21	21	21	21	21	E	14a	
70	(731,934)	22	22		22	22		22	22	E	14a	
71	(704,841)									E	14a	
72	(1013,853)	28			28		28			E	14a	
73	(840,832)	29			29					I	14a	
74	(864,814)	30		30	30				30	I	14a	
75	(866,794)	31			31					I	14a	
76	(997,596)			33			33	33	33	I	14a	

#	(x, y)	EE	IE	LE	TE	II	LI	TI	TIE	Gr	Figuras	Ob
77	(1153,1013)			34	34	34	34	34	34	I	14a	
78	(1227,887)									I	14a	
79	(1190,736)						38			I	14a	
80	(1213,658)									I	14a	
81	(1291,1029)								42	I	14a	
82	(1336,902)									I	14a	
83	(1347,706)			45	45	45		45	45	I	14a	
84	(1265,649)			46	46	46	46		46	I	14a	
85	(1342,628)					47	47	47	47	I	14a	
86	(1260,606)							48	48	I	14a	
87	(602,893)					1	1	1	1	E	14b	
88	(846,595)							8	8	I	14b	
89	(933,1008)		14	14	14	14	14	14	14	E	14b	
90	(927,666)									I	14b	
91	(921,557)							18	18	E	14b	
92	(888,565)							19	19	E	14b	
93	(971,969)		15	15	15		15			E	14b	
94	(1144,231)	22			22				22	E	14b	
95	(1158,832)									E	14b	
96	(1175,252)			28	28				28	I	14b	
97	(1146,220)								29	E	14b	
98	(857,972)			52	52		52	52	52	E	15a	I
99	(110,900)					36	36	36	36	E	15b	L
100	(806,608)					37	37		37	E	15b	L
101	(859,974)					38	38	38	38	E	15b	L
102	(1058,971)		39	39	39				39	E	15b	L
103	(1343,697)			40	40	40			40	E	15b	L
104	(1194,607)			0	0					NA	13	f13
105	(62,713)				0	0		0	0	NA	13	f13
106	(397,876)	0			0		0	0	0	NA	13	f13
107	(401,694)									NA	13	f13
108	(684,973)			0	0	0		0	0	NA	13	f13
109	(710,921)			0	0	0	0	0	0	NA	13	f13
110	(945,902)			0	0				0	NA	13	f13
111	(1129,963)									NA	13	f13
112	(1133,937)									NA	13	f13
113	(1253,867)									NA	13	f13
114	(1334,1010)		0	0	0	0	0	0	0	NA	13	f13
115	(1326,622)					0	0	0	0	NA	13	f13

**Tabla 4-16:** Tabla relacional de la cantidad de proteínas totales detectadas en cada muestra y en la fusión de imágenes y la cantidad de proteínas iguales contenidas en las otras muestras relacionales

<b>Spots relacionales</b>	<b>Esmeralda Ext</b>	<b>Isla Ext</b>	<b>Loma Ext</b>	<b>Tot Ext</b>	<b>Isla Int</b>	<b>Loma Int</b>	<b>Tot Int</b>	<b>Tot Int_Ext</b>
Esmeralda Exterior	23	1	6	23	9	12	14	16
Isla Exterior	1	12	8	8	9	7	9	9
Loma Exterior	6	8	37	31	21	15	23	31
Total, Exterior	23	8	31	53	25	23	34	42
Isla Interior	9	9	21	25	53	30	46	47
Loma Interior	12	7	15	23	30	47	41	40
Total, Interior	14	9	23	34	46	41	69	66
Total, Interior-Exterior	16	9	31	42	47	40	66	79
Spots totales detectados	23	12	37	53	53	47	69	79
Porcentaje spots totales detectados relacionados con total spots de patrón	20%	10%	32%	46%	46%	41%	60%	69%

(LE); Isla Interior (II); Loma Interior (LI) , imágenes fusionadas (Total Exterior (TE), Total Interior (TI) y Total Interior – Exterior (TIE) , grupos de imágenes (Gr) para las proteínas detectadas (Interior (I), Exterior (E) o combinaciones (IE)), figuras del informe FGN008 de las cuales fueron extraídos las proteínas detectados (13, 14a, 14b, 15a, 15b, 16a,16b) y la observación (Ob) (Común expresada (CE), Isla (I), Loma (L), Figura 13 (f13)), todos los anteriores ítems relacionados con cada una de las 115 proteínas conocidos del patrón.

En el resumen de la Tabla 4- 13 se puede observar para cada agrupación de imágenes (de interior y de exterior), el total de proteínas conocidas detectadas y su porcentaje relacionado con el número total de proteínas del patrón (115). Se observa que las muestras a nivel individual de sitios, en donde se encontraron la mayoría de los spots están en la imagen la Isla Interior con 53 proteínas, para un porcentaje del 46.1% siendo el menor valor de 12 proteínas para la Isla Exterior. En la fusión de imágenes el mayor valor está dado para el total interior con 69 proteínas detectadas para un 60% y el valor total encontrado en la fusión de imágenes de interior y de exterior fue de 79 proteínas para un 68.7% relacionado con el total de proteínas conocidas del patrón. También se observa para el total de proteínas detectadas a nivel individual de grupos de muestras de imágenes y de fusión, la relación de objetos idénticos que se encuentran en el resto de las agrupaciones, por

ejemplo, de la muestra fusionada de la Esmeralda Exterior se encuentran en la Isla Exterior, Loma Exterior, Total Exterior, Isla Interior, Loma Interior, Total Interior y Total Interior-Exterior 1, 6, 23, 9, 12, 14, 16 proteínas iguales (primera fila). Acá se observa que las 23 proteínas conocidas encontradas en la fusión de las muestras de La Esmeralda Exterior, se encuentran en el total de proteínas conocidas encontradas en la fusión total de imágenes del exterior. La Tabla 4-17 presenta la relación de proteínas tomadas del Informe FGN008, presentes solo en exterior, presentes solo en Interior y presentes en ambos con expresión diferencial.

**Tabla 4-17:** Proteínas con expresión diferencial entre la apitoxina de obreras de exterior e interior FGN008

Spot #	Nivel de expresión - Interior	Nivel de expresión - Exterior	Fold change	Exp. MW	Exp. pI	Posible Identidad
<b>Proteínas presentes solo en Exterior</b>						
6	0	1213519		30	3	
7	0	34908		55	4	Major royal jelly protein 4
8	0	149078		11	4,5	Prepromelittin
9	0	2039829		13	4,8	Profilin
10	0	738511		15	4,8	
11	0	165582		15	5	
12	0	515768		45	4,5	Acid phosphatase
13	0	147718		55	4	Icarapin variant 2 (Api m10)
14	0	155313		12	5,5	Defensin-1
15	0	158290		15	5,4	New venom protein 1
16	0	904224		25	6	New venom protein 3
18	0	249953		22	6,8	
20	0	240388		25	6	Vitellogenin
21	0	81350		27	6	
22	0	561393		27	6,5	
23	0	22671197		?	?	
25	0	264528		?	?	
26	0	93893		100	7,5	Dipeptidylpeptidase IV
29	0	269208		100	7,6	
30	0	406018		100	7,7	
31	0	65248		100	7,8	
32	0	199504		15	8,4	
35	0	403507		21	8,6	
36	0	5284502		22	9	
37	0	520281		78	9	Venom carboxylesterase-6 (Api m8)
38	0	871782		60	9	
39	0	516417		160	9	
40	0	36737		240	9	
41	0	129040		7	9,2	Api m6
42	0	711867		10	9,4	Preproscapin
45	0	14076654		?	?	
46	0	6325669		22	9,2	
47	0	1136530		15	9,3	

Proteínas presentes solo en Interior						
19	174141	0		20	6,8	Phospholipase A2
24	1371402	0		10	7	Secapin 3
27	43213	0		45	7,5	CUB serine protease (Api m7)
28	73173	0		48	7,8	Major royal jelly protein 9
33	12960	0		22	8,2	Allatostatins
34	14385	0		26	8,4	
Proteínas presentes en ambos - con expresión diferencial						
1	671501	3273276	4,9	14	8,4	
2	7001	675010	9,9	7	9,2	
3	317898	44777	7,1	62	6	Major royal jelly protein 3
4	62091	245224	3,9	50	7,2	
5	43231	182827	4,2	50	7,4	Serpin 5

Tabla 4-18: Proteínas presentes sólo en exterior e interior con técnicas de segmentación

PROTEINAS PRESENTES SÓLO EN EXTERIOR				PROTEINAS PRESENTES SÓLO EN INTERIOR				
#Pat	(x,y)	EE	IE	LE	#Pat	(x,y)	II	LI
4	(609,663)		4-24	4-24	1	(942,943)	1-51	
5	(650,665)		5	5	2	(1175,985)		2
6	(41,768)		6	6	3	(413,599)	3-20	3-20
8	(175,954)			8	9	(219,939)		9
11	(349,918)			11	10	(259,915)	10	
12	(186,702)			3-12-49	16	(442,842)	16	
13	(161,637)			4-13	19	(560,869)	19	
43	(297,952)	1			22	(519,793)	22-50	
50	(321,934)	5			23	(583,962)	24	24
18	(551, 845)		18		24	(697,380)	10-26	10-26
56	(394,574)	10			29	(758,382)		13-31
26	(780,680)			28	34	(1155,789)		22-36
57	(460,963)	11			35	(1139,491)		37
14	(456,943)	14		14	44	(601,831)		2-17
60	(432,921)	14		14	49	(869,674)	17-26	
36	(1146,619)			6-35-38	51	(58,891)	6	
15	(429,919)	15		15	53	(365,853)	8	8
17	(494,881)	17			55	(873,608)	32	
40	(1266,955)			42	59	(586,929)	13	
20	(435,824)	20			62	(1222,641)	32-40	32-40
45	(737,603)			9-27	63	(499,929)	15	
46	(58,924)			2	66	(553,885)	16	16
47	(97,933)		3		67	(415,679)	18	18
48	(351,914)			4	79	(1190,736)		38
69	(643,928)	21		21	85	(1342,628)	47	47
52	(93,905)			7	87	(602,893)	1	1
54	(95,665)			9	99	(110,900)	36	36
70	(731,934)	22	22		100	(806,608)	37	37
94	(1144,231)	22			101	(859,974)	38	38

ROTEINAS PRESENTES SÓLO EN EXTERIOR				PROTEINAS PRESENTES SÓLO EN INTERIOR				
#Pat	(x, y)	EE	IE	LE	#Pat	(x, y)	II	LI
58	(539,970)		12					
72	(1013,853)	28						
73	(840,832)	29						
68	(607,667)		19	19				
74	(864,814)	30		30				
75	(866,794)	31						
31	(883,840)	33						
33	(1015,850)	35						
39	(1208,992)	41						
27	(713,378)	11-29						
76	(997,596)			33				
77	(1153,1013)			34				
83	(1347,706)			45				
84	(1265,649)			46				
89	(933,1008)		14	14				
93	(971,969)		15	15				
37	(1145,270)	27-39		27-39				
96	(1175,252)			28				
98	(857,972)			52				
102	(1058,971)		39	39				
103	(1343,697)			40				
64	(1347,967)	35-43						
<b>TOTAL EXTERIOR</b>				<b>51</b>	<b>TOTAL INTERIOR</b>			<b>29</b>



**Tabla 4-19:** Posible identidad de algunas proteínas y sus características procesadas con técnica de segmentación multiumbralización 8 ventanas.

#	Area	Centroide (x, y)	BoundingBox	#Spot	Grupo	Fig	Posible identidad
1	199	[607,5 656,5]	[597,5 649,5 20 14]	4-24	IE,LE,II	14a-16b	Secapin 3
2	4	[184,5 947,5]	[183,5 946,5 2 2]	8	LE,II	16a	Prepromelittin
3	2	[328 942,5]	[327,5 941,5 1 2]	5	EE,II,LI	14a	Serpin 5
4	4	[473,3 946,3]	[471,5 945,5 3 2]	14	EE,LE,II,LI	16a	Defensin-1
5	4	[439 916]	[438,5 915,5 1 1]	14	EE,LE,II,LI	14a	Defensin-1
6	4	[395,871]	[394,5 870,5 1 1]	15	EE,LE,II,LI	16a	New venom protein 1
7	14	[1277,5 944,7]	[1275,5 942,5 4 4]	42	LE	16a	Preprosecapin
8	3	[431,7 824,3]	[430,5 823,5 2 2]	20	EE,II	16a	Vitellogenin
9	16	[737,1 591,1]	[735,5 588,5 4 5]	9-27	LE	14a-14b	CUB serine protease (Api m7)
10	4	[555,5 968,5]	[554,5 967,5 2 2]	12	IE,II,LI	14a	Add phosphatase
11	199	[607,5 656,5]	[597,5 649,5 20 14]	19	IE,LE,II,LI	14a	Phospholipase A2
12	667	[1197,9 983,2]	[1183,5 961,5 29 39]	41	EE	16a	Api m6
13	4	[991 600,5]	[990,5 599,5 1 2]	33	LE,LI	14a	Allatostatins
14	16	[1183,8 249,4]	[1181,5 247,5 5 4]	28	LE	14b	Major royal jelly protein 9
15	4	[210,5 924,5]	[209,5 923,5 2 2]	9	LI	16a	Profilin
16	159	[555,7 877]	[542,5 871,5 29 10]	16	II	16a	New venom protein 3
17	4	[755,5 378,5]	[754,5 377,5 2 2]	13-31	LI	14b-16a	Icarapin variant 2(Api m10)
18	4	[859 674,5]	[858,5 673,5 1 2]	17-26	II	14a-14b	Dipeptidylpeptidase IV
19	4	[497 921]	[496,5 920,5 1 1]	15	II	14a	New venom protein 1
20	4	[552,5 904]	[551,5 903,5 2 1]	16	II, LI	14a	New venom protein 3
21	542	[807,9 608,4]	[794,5 593,5 30 28]	37	II,LI	15b	Venom carboxylesterase-6(Api m8)

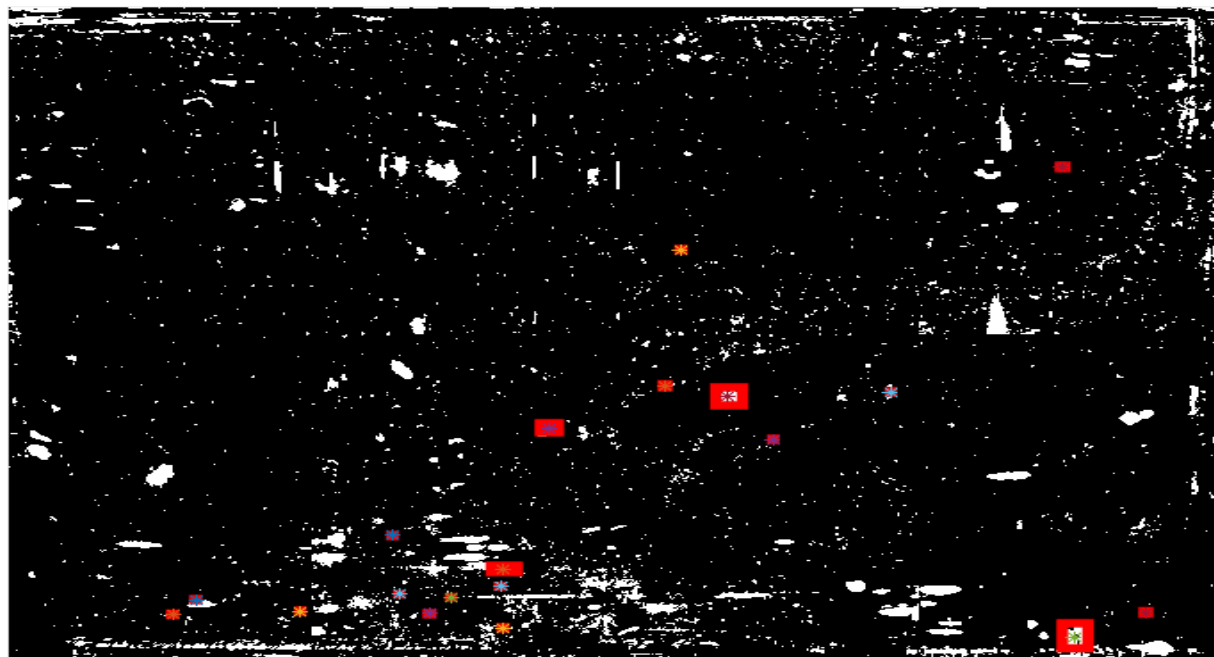


Figura 4-36: Marcación y visualización de algunas proteínas con posible identidad según información dada en la Tabla 4-19



Figura 4-37: Visualización más detallada de características de 3 proteínas según información dada en la Tabla 4-19

Al realizar la comparación de la Tabla 4-17 y la Tabla 4-18 se observa que de las 33 proteínas presentes solo en Exterior extraídas del Informe FGN008, se detectan 26 proteínas iguales equivalentes a un 78.78%. De las 6 proteínas presentes sólo en interior del informe FGN008, se detectaron solamente 2 proteínas iguales (las números 19 y 24 resaltadas), equivalente al 33.33%. En total se encontraron 51 proteínas presentes sólo en exterior empleando las técnicas de segmentación de umbralización automática y multiumbralización automática para 4 – 8 y 16 ventanas, para un porcentaje del 155%, relacionado con las 33 proteínas detectadas en el informe FGN008 y 29 proteínas presentes sólo en interior, equivalente al 483% con respecto a las 6 proteínas encontradas sólo para interior en el informe FGN008.

La Tabla 4-19 es el resultado del proceso de extracción de 21 proteínas de la Tabla 4-17, que tienen una posible identidad y que fue realizado con la técnica de segmentación y detección de proteínas multiumbralización automática 8 ventanas, que fue la que mayor cantidad de proteínas detectó con respecto a las otras técnicas de segmentación empleadas para el análisis de las imágenes. Las proteínas de la Tabla 4-19 se visualizan marcadas en la Figura 4-36 y la Figura 4-37 presenta el detalle de tres proteínas de la misma Tabla con sus características de ubicación en coordenadas (x,y), área y nombre identificador de la proteína.

#### **4.4.1 Discusión**

El reconocimiento de proteínas llevado a cabo en el Informe FGN008 , mediante la información contenida en las Figuras 4-6, 4-7, 4-8, 4-9 fue la base para la creación del patrón de comparación de las proteínas procesadas en las muestras de las 23 imágenes provenientes de los grupos de interior y de exterior de las colmenas de las veredas La Esmeralda, La Isla y La Loma, aplicando en ellas las técnicas de segmentación que más se amoldaban como fueron la umbralización automática y la multiumbralización automática de 4 – 8 y 16 ventanas. La Tabla 4-15 muestra cómo para los 115 spots del patrón se establecen las coordenadas de ubicación (x, y) y los spots relacionales encontrados de acuerdo a la información extraída del informe FGN008, así como la agrupación y las figuras de las cuales se tomó la información. Con toda la información recopilada se pudo elaborar comparaciones y relaciones para las diferentes agrupaciones y fusiones de todas las imágenes, tal como lo muestran también las Tablas 4-16 y 4-18. La Tabla 4-19 representa una

relación abreviada de lo que se puede lograr con la utilización de las técnicas de segmentación trabajadas en el proyecto.

## 5. Conclusiones.

Existen pocos estudios formales relacionados con el análisis proteómico del veneno de las abejas y poco conocimiento de éstos por la comunidad apicultora de Colombia, además que no se han realizado investigaciones que muestren las diferencias proteómicas entre las dos poblaciones de abejas africanizadas estudiadas en este proyecto como lo son las abejas del grupo de interior y las del grupo de exterior, clasificadas por la edad y por las actividades que realizan a nivel interno y externo de la colmena.

Para el desarrollo de este proyecto se utilizó la base de datos de 23 imágenes con las muestras del veneno de abejas apitoxina de los grupos de interior y de exterior generadas en el laboratorio de Biología Celular y Molecular del ITM (Tabla 4-5 y Figuras 4-4 y 4-5) y una imagen de apitoxina facilitada por el Dr. Dirk de Graaf (Figuras 3-3c y 3-4). Se contó con la colaboración de dos grupos interdisciplinarios de co-investigadores y profesionales de diferentes disciplinas del Instituto Tecnológico Metropolitano de Medellín (ITM), se tuvo la herramienta y mano de obra calificada para el pre procesamiento y procesamiento de las imágenes 2DGE, a la vez que para la división e identificación de las muestras de veneno de las colonias de interior y de exterior; trabajo realizado por apicultores expertos, clasificación que no se había realizado en el estado del arte.

Para lograr el objetivo general, se conformaron dos patrones bases para establecer la comparación de las imágenes: el primero conformado con las 62 proteínas de la imagen del Dr. Dirk de Graaf y el segundo constituido con las 115 proteínas reconocidas en el estudio realizado en el informe FGNOO8. La metodología establecida para la identificación de las proteínas se describe en las Figuras 3 – 1 y 3 – 2.

Para la consecución de los objetivos específicos como son la identificación de descriptores morfológicos y el análisis espacial para establecer la comparación y relación proteómica de las imágenes 2DGE, se llevaron a cabo procesos de segmentación empleando métodos de umbralización, multiumbralización automática, detección de bordes y transformada Watershed, técnicas que se aplicaron en todas las imágenes de la base de datos, teniendo como referencia de comparación el patrón de la imagen del Dr. Dirk de Graaf (Tabla 4-1, Figura 3-4) y el patrón de

proteínas conocidas elaborado del análisis realizado a las proteínas detectadas en el Informe FGN008 (Tabla 4-6). Todas las imágenes fueron sometidas a procesos de estandarización y filtrado con el fin de lograr las coincidencias espaciales en los objetos detectados. El proceso se trabajó con una aproximación de  $\pm 10$  píxeles para la ubicación de los puntos de coordenadas  $(x,y)$  referenciados sobre el centro de masa de las proteínas. El grado de aproximación está directamente relacionado con la cantidad de proteínas conocidas detectadas.

Los resultados obtenidos muestran que teniendo patrones de referencia definidos se pueden obtener altos porcentajes en la sensibilidad del detector de puntos de interés, tales como los relacionados en la Tablas 4-2, con valores por encima del 70% en el número de objetos detectados relacionados con los objetos conocidos y los relacionados en las Tablas 4-11 en el consolidado general de la comparación de las técnicas de segmentación en donde los porcentajes de la relación de sensibilidad es mayor del 60%, sobre todo en las técnicas de Multiumbralización en 4 – 8 y 16 ventanas. En los procesos de análisis de imágenes a nivel individual, el proceso que mostró los mejores resultados en cuanto a porcentaje de sensibilidad (S%) y detección de proteínas conocidas (VP) fue la técnica multiumbralización automática 8 ventanas (ver Figura 4-14).

En los procesos de análisis de imágenes a nivel de fusión, el método que mostró los mejores resultados en cuanto a porcentaje de sensibilidad (S%) y detección de cantidad de proteínas conocidas del patrón base (VP) fueron las técnicas de multiumbralización automática, siendo la mejor técnica la de 16 ventanas (ver figuras 4-30 y 4-31), en la cual se detectaron 79 proteínas de las 115 proteínas conocidas del patrón, equivalente a un porcentaje del 68.7%. También esta técnica fue la que mejores resultados dio con respecto al coeficiente de similitud Sorensen-Dice.

La ventaja de dividir la imagen en múltiples ventanas y realizar las comparaciones por ventanas es que se logra mayor precisión en cuanto a cantidad de detecciones, pero la desventaja radica en que, si la región fragmentada posee pocos objetos de interés o en su defecto no tiene, esta región se convierte en el proceso de fusión en una región únicamente iluminada por el color blanco o negro, dependiendo del umbral con el que se este trabajando.

En el proceso de fusión de las imágenes fue necesario eliminar áreas  $> 1000$  píxeles para evitar que las proteínas fusionadas formaran objetos muy densos difíciles de identificar y que le inyectarían mucho ruido al análisis general de fusión. Esta técnica de eliminar objetos con áreas

determinadas ayuda a tener una mejor identificación y visualización de los objetos que deseamos analizar de acuerdo con el área en píxeles. También los resultados denotan que entre más objetos segmentados contenga la imagen, mayor es la cantidad de objetos detectados de un patrón de referencia, pero teniendo el cuidado de poder identificar en las imágenes objetos de interés que no se confundan con manchas u falso objetos pertenecientes a ruidos.

En la comparación realizada de las proteínas presentes sólo en Exterior y sólo en interior con respecto al Informe FGN008, se lograron encontrar unos valores altos de detección, siendo de 51 spots para el total de proteínas sólo del exterior y de 29 spots para el total de proteínas sólo de interior (Ver Tabla 4–15).

Uno de los problemas que se presentan en el procesamiento y análisis automático de imágenes de electroforesis bidimensional en gel 2DGE es la variabilidad implícita de las muestras biológicas, las cuales presentan diversidades de cambios en su preparación y análisis de resultados.

## 6. Recomendaciones.

Con la información que se tiene en los dos patrones básicos de apitoxina con los cuales se realizaron las respectivas comparaciones de las técnicas de segmentación, como los son el patrón del Dr. Dirk de Graaf y el patrón de apitoxina extraído del Informe GN008, además de la información de las características espaciales referenciadas en el centro de masa de las proteínas de las 23 muestras analizadas, se pueden continuar estudios que muestren las diferencias entre el veneno de las abejas africanizadas con respecto al resto de abejas estudiadas por diversos científicos como lo es el estudio publicado hace unos 4 años en donde se demuestra que la apitoxina de las abejas africanizadas contienen 4 proteínas únicas que no se encuentran en otras abejas Europeas estudiadas ((Resende, Vasilj, Santos, Palma, & Shevchenko, 2013). Se puede realizar la comparación de los resultados experimentales obtenidos con el patrón Dr. Dirk de Graaf con los resultados experimentales obtenidos en las 23 muestras de apitoxina procesadas con las diferentes técnicas expuestas en este proyecto.

Se observa en la figuras 4-6, 4-7, 4-8 y 4-9 que en general, la apitoxina de abejas africanizadas está compuesta en su mayor parte por proteínas y péptidos de bajo peso molecular (10 – 60 kDa), lo cual coincide con lo reportado en la literatura. La mayoría de las proteínas en la apitoxina de las abejas europeas comunes (*Apis mellifera ligústica* y *Apis mellifera carnica*) están en el rango entre los 7 y 88 kDa (Li et al., 2011; Peiren et al., 2005). La agrupación más abundante de proteínas se encuentra en la parte inferior derecha de la imagen, correspondiente a un punto isoelectrico básico (pI 7.5 - 9) y a un bajo peso molecular (< 5 kDa). Basado en reportes anteriores, es muy probable que estos spots correspondan a las proteínas Fosfolipasa A2 (MW 16 kDa, pI 8.07) y Api m6 (MW 7 kDa, pI 9.70), que son dos de los principales alérgenos de la apitoxina. Otras proteínas importantes, aunque menos abundantes, son la Serina Proteasa (MW 39 kDa, pI 8.64), Hyaluronidase (MW 40 kDa, pI 8.72), Icarapin variant 2 (MW 50 kDa, pI 4.51), Acid Phosphatase (MW 45 kDa, pI 5.63) y major royal jelly proteins (MW 50 kDa, pI 6.53 and 8.70). Con la información del rango de ubicación en kDa de las proteínas y con la ubicación en coordenadas de estos rangos, se puede en un estudio posterior realizar una discriminación detallada de muchas proteínas conocidas con nombre propio, como las enunciadas en las Tablas 4-17, 4-19 y que se



encuentran presentes en las muestras de las 23 imágenes de la base y en las imágenes presentadas en los procesos de fusión llevados a cabo con estas imágenes procesadas con las técnicas descritas en el proyecto.

Se puede introducir en los algoritmos generados en este proyecto, otras técnicas de filtrado diferentes a la empleada, tales como las técnicas Wavelet, Contourlet y Variación Total, además de experimentar más profundamente con la transformada Watershed en combinación con otras técnicas que me permitan evitar los problemas de sobre segmentación.

## 7. Bibliografía

- Adams, R. (1993). Radial Decomposition of Disks and Spheres. *CVGIP: Graphical Models and Image Processing*, 55(5), 325–332. <http://doi.org/10.1006/cgip.1993.1024>
- Aittokallio, T., Salmi, J., Nyman, T. A., & Nevalainen, O. S. (2005). Geometrical distortions in two-dimensional gels: applicable correction methods. *Journal of Chromatography B*, 815(1), 25–37. <http://doi.org/10.1016/j.jchromb.2004.07.037>
- Alvarez-Fischer, D., Noelker, C., Vulinović, F., Grünewald, A., Chevarin, C., Klein, C., ... Hartmann, A. (2013). Bee Venom and Its Component Apamin as Neuroprotective Agents in a Parkinson Disease Mouse Model. *PLoS ONE*, 8(4), e61700. <http://doi.org/10.1371/journal.pone.0061700>
- Amdam, G. V., Csondes, A., Fondrk, M. K., & Page, R. E. (2006). Complex social behaviour derived from maternal reproductive traits. *Nature*, 439(7072), 76–78. <http://doi.org/10.1038/nature04340>
- Appel, R. D., Vargas, J. R., Palagi, P. M., Walther, D., & Hochstrasser, D. F. (1997). Melanie II - a third-generation software package for analysis of two-dimensional electrophoresis images: II. Algorithms. *Electrophoresis*, 18(15), 2735–2748. <http://doi.org/10.1002/elps.1150181507>
- Bandow, J. E., Baker, J. D., Berth, M., Painter, C., Sepulveda, O. J., Clark, K. A., ... VanBogelen, R. A. (2008). Improved image analysis workflow for 2-D gels enables large-scale 2-D gel-based proteomics studies - COPD biomarker discovery study. *PROTEOMICS*, 8(15), 3030–3041. <http://doi.org/10.1002/pmic.200701184>
- Berth, M., Moser, F. M., Kolbe, M., & Bernhardt, J. (2007). The state of the art in the analysis of two-dimensional gel electrophoresis images. *Applied Microbiology and Biotechnology*, 76(6), 1223–1243. <http://doi.org/10.1007/s00253-007-1128-0>
- Blackstock, W. P., & Weir, M. P. (1999). Proteomics: quantitative and physical mapping of cellular proteins. *Trends in Biotechnology*, 17(3), 121–127. [http://doi.org/10.1016/S0167-7799\(98\)01245-1](http://doi.org/10.1016/S0167-7799(98)01245-1)
- Brandão, A. R., Barbosa, H. S., & Arruda, M. A. Z. (2010). Image analysis of two-dimensional gel electrophoresis for comparative proteomics of transgenic and non-transgenic soybean seeds. *Journal of Proteomics*, 73(8), 1433–1440. <http://doi.org/10.1016/j.jprot.2010.01.009>
- Bray, J. Roger; Curtis, J. T. (1948). «An Ordination of the Upland Forest Communities of Southern Wisconsin». *Ecological Monographs* 27 (4): 326-349
- Calligaris, D., Villard, C., & Lafitte, D. (2011). Advances in top-down proteomics for disease biomarker discovery. *Journal of Proteomics*, 74(7), 920–934. <http://doi.org/10.1016/j.jprot.2011.03.030>

- Caruso, G., Cavaliere, C., Guarino, C., Gubbiotti, R., Foglia, P., & Laganà, A. (2008). Identification of changes in *Triticum durum* L. leaf proteome in response to salt stress by two-dimensional electrophoresis and MALDI-TOF mass spectrometry. *Analytical and Bioanalytical Chemistry*, 391(1), 381–390. <http://doi.org/10.1007/s00216-008-2008-x>
- Cieslak, A., & Ribera, I. (2009). Aplicaciones de proteómica en ecología y evolución. *Revista Ecosistemas*, 18(1). <http://doi.org/10.7818/RE.2014.18-1.00>
- Clark, B. N., & Gutstein, H. B. (2008). The myth of automated, high-throughput two-dimensional gel analysis. *PROTEOMICS*, 8(6), 1197–1203. <http://doi.org/10.1002/pmic.200700709>
- Cutler, P., Heald, G., White, I. R., & Ruan, J. (2003). A novel approach to spot detection for two-dimensional gel electrophoresis images using pixel value collection. *PROTEOMICS*, 3(4), 392–401. <http://doi.org/10.1002/pmic.200390054>
- Daszykowski, M., Bierczynska-Krzsik, A., Silberring, J., & Walczak, B. (2010). Avoiding spots detection in analysis of electrophoretic gel images. *Chemometrics and Intelligent Laboratory Systems*, 104(1), 2–7. <http://doi.org/10.1016/j.chemolab.2009.10.001>
- De Felipe, M., & Vandame, R. (1999). Curso de capacitación sobre control alternativo de Varroa en la apicultura...4/15.
- Daszykowski, M., Færgestad, E. M., Grove, H., Martens, H., & Walczak, B. (2009). Matching 2D gel electrophoresis images with Matlab “Image Processing Toolbox.” *Chemometrics and Intelligent Laboratory Systems*, 96(2), 188–195. <http://doi.org/10.1016/j.chemolab.2009.01.011>
- Dice, Lee R. (1945). «Measures of the Amount of Ecologic Association Between Species». *Ecology* 26 (3): 297-302
- Do, M. N., & Vetterli, M. (2005). The contourlet transform: an efficient directional multiresolution image representation. *IEEE Transactions on Image Processing*, 14(12), 2091–2106. <http://doi.org/10.1109/TIP.2005.859376>
- dos Anjos, A., Moller, A. L. B., Ersboll, B. K., Finnie, C., & Shahbazkia, H. R. (2011). New approach for segmentation and quantification of two-dimensional gel electrophoresis images. *Bioinformatics*, 27(3), 368–375. <http://doi.org/10.1093/bioinformatics/btq666>
- Dowsey, A. W., Dunn, M. J., & Yang, G.-Z. (2003). The role of bioinformatics in two-dimensional gel electrophoresis. *PROTEOMICS*, 3(8), 1567–1596. <http://doi.org/10.1002/pmic.200300459>
- Efrat, A., Hoffmann, F., Kriegel, K., Schultz, C., & Wenk, C. (2002). Geometric Algorithms for the Analysis of 2D-Electrophoresis Gels. *Journal of Computational Biology*, 9(2), 299–315. <http://doi.org/10.1089/10665270252935476>
- Faheem, M. T., Rashwan, S., Sarhan, A., & Youssef, B. A. (2011). De-noising 2D gel images using Wavelet transform based on genetic algorithm. University of Tanta, Tanta, Egypt. Retrieved from

- Fleiss, J. L. (1981). The Measurement of Interrater Agreement. *Statistical Methods for Rates and Proportions*, 2nd edn pp 212–304.  
<https://pdfs.semanticscholar.org/3a7c/58805aaa7f61c40726f483240a03ee4734de.pdf>
- Fujita, T., Kozuka-Hata, H., Uno, Y., Nishikori, K., Morioka, M., Oyama, M., & Kubo, T. (2010). Functional analysis of the honeybee (*Apis mellifera* L.) salivary system using proteomics. *Biochemical and Biophysical Research Communications*, 397(4), 740–744. <http://doi.org/10.1016/j.bbrc.2010.06.023>
- Gonzalez, R. C., Woods, R. E. (2002). *Digital image processing*. Second edition. Beijing: Publishing House of Electronics Industry.
- Görg, A., Weiss, W., & Dunn, M. J. (2005). Current two-dimensional electrophoresis technology for proteomics (vol. 4, Issue 12, pp. 3665-3685). *PROTEOMICS*, 5(3), 826–827. <http://doi.org/10.1002/pmic.200590007>
- Havre, S. L., Singhal, M., Payne, D. A., Weir Lipton, M. S., & Webb-Robertson, B.-J. M. (2005). Enabling proteomics discovery through visual analysis. *IEEE Engineering in Medicine and Biology Magazine*, 24(3), 50–57. <http://doi.org/10.1109/MEMB.2005.1436460>
- Hood, J. L., Jallouk, A. P., Campbell, N., Ratner, L., & Wickline, S. A. (2012). Cytolytic nanoparticles attenuate HIV-1 infectivity. *Antiviral Therapy*, 18(1), 95–103. <http://doi.org/10.3851/IMP2346>
- Iakovidis, D. K., Maroulis, D., Zacharia, E., & Kossida, S. (2006). A genetic approach to spot detection in two-dimensional gel electrophoresis images. In *Proceedings of International Conference on Information Technology in Biomedicine (ITAB)*. Ioannina, Greece. Retrieved from [https://www.researchgate.net/profile/Dimitris\\_Maroulis/publication/230689616\\_A\\_genetic\\_approach\\_to\\_spot\\_detection\\_in\\_two-dimensional\\_gel\\_electrophoresis\\_images/links/0c960518ac355ba629000000.pdf](https://www.researchgate.net/profile/Dimitris_Maroulis/publication/230689616_A_genetic_approach_to_spot_detection_in_two-dimensional_gel_electrophoresis_images/links/0c960518ac355ba629000000.pdf)
- Jacobsen, S., Grove, H., Nedenskov Jensen, K., Sørensen, H. A., Jessen, F., Hollung, K., ... Søndergaard, I. (2007). Multivariate analysis of 2-DE protein patterns – Practical approaches. *ELECTROPHORESIS*, 28(8), 1289–1299. <http://doi.org/10.1002/elps.200600414>
- Jones, R., & Soille, P. (1996). Periodic lines: Definition, cascades, and application to granulometries. *Pattern Recognition Letters*, 17(10), 1057–1063. [http://doi.org/10.1016/0167-8655\(96\)00066-9](http://doi.org/10.1016/0167-8655(96)00066-9)
- Kaczmarek, K., Walczak, B., Jong, S. de, & Vandeginste, B. G. M. (2002). Feature Based Fuzzy Matching of 2D Gel Electrophoresis Images. *Journal of Chemical Information and Modeling*, 42 (6), 1431–1442. <http://doi.org/10.1021/CI020266K>
- Kaczmarek, K., Walczak, B., Jong, S. de, & Vandeginste, B. G. M. (2003). Matching 2D Gel Electrophoresis Images. *Journal of Chemical Information and Modeling*, 43 (3), 978–986.

<http://doi.org/10.1021/CI0256337>

- Kim, Y., Kim, J., Won, Y., & In, Y. (2003). Segmentation of Protein Spots in 2D Gel Electrophoresis Images with Watersheds Using Hierarchical Threshold (pp. 389–396). Springer, Berlin, Heidelberg. [http://doi.org/10.1007/978-3-540-39737-3\\_49](http://doi.org/10.1007/978-3-540-39737-3_49)
- Kostopoulou, E., Katsigiannis, S., & Maroulis, D. (2015). 2D-gel spot detection and segmentation based on modified image-aware grow-cut and regional intensity information. *Computer Methods and Programs in Biomedicine*, 122(1), 26–39. <http://doi.org/10.1016/j.cmpb.2015.06.007>
- Kostopoulou, E., Katsigiannis, S., & Maroulis, D. (2015). A custom grow-cut based scheme for 2D-gel image segmentation. In 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 2407–2410). IEEE. <http://doi.org/10.1109/EMBC.2015.7318879>
- Kostopoulou, E., Katsigiannis, S., Maroulis, D., Pappa, K. I., & Anagnou, N. P. (2015). A novel approach for accurate 2D-gel image segmentation, 10, 9–11. Retrieved from [http://rtsimage.di.uoa.gr/publications/HSCBB15\\_submission.pdf](http://rtsimage.di.uoa.gr/publications/HSCBB15_submission.pdf)
- Kostopoulou, E., Zacharia, E., & Maroulis, D. (2012). Accurate segmentation of 2D-PAGE images. In *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European* (pp. 2258–2262). [IEEE].
- Kostopoulou, E., Zacharia, E., & Maroulis, D. (2014). An effective approach for detection and segmentation of protein spots on 2-d gel images. *IEEE Journal of Biomedical and Health Informatics*, 18(1), 67–76. <http://doi.org/10.1109/JBHI.2013.2259208>
- Lee, J.-D., Park, H.-J., Chae, Y., & Lim, S. (2005). An Overview of Bee Venom Acupuncture in the Treatment of Arthritis. *Evidence-Based Complementary and Alternative Medicine: eCAM*, 2(1), 79–84. <http://doi.org/10.1093/ecam/neh070>
- Lee, J., Han, J. J., Altwerger, G., & Kohn, E. C. (2011). Proteomics and biomarkers in clinical trials for drug development. *Journal of Proteomics*, 74(12), 2632–2641. <http://doi.org/10.1016/j.jprot.2011.04.023>
- Lemkin, P. F., & Lipkin, L. E. (1981). GELLAB: A computer system for 2D gel electrophoresis analysis I. Segmentation of spots and system preliminaries. *Computers and Biomedical Research*, 14(3), 272–297. [http://doi.org/10.1016/0010-4809\(81\)90050-1](http://doi.org/10.1016/0010-4809(81)90050-1)
- Lemkin, P. F., & Lipkin, L. E. (1983). 2-D Electrophoresis gel data base analysis: Aspects of data structures and search strategies in GELLAB. *Electrophoresis*, 4(1), 71–81. <http://doi.org/10.1002/elps.1150040110>
- Li, F., Seillier-Moiseiwitsch, F., & Korostyshevskiy, V. R. (2011). Region-based statistical analysis of 2D PAGE images. *Computational Statistics & Data Analysis*, 55(11), 3059–3072. <http://doi.org/10.1016/j.csda.2011.05.013>
- Lin, C.-Y., Wang, V., Shui, H.-A., Juang, R.-H., Hour, A.-L., Chen, P.-S., ... Chen, H.-M. (2009).

- A comprehensive evaluation of imidazole-zinc reverse stain for current proteomic researches. *PROTEOMICS*, 9(3), 696–709. <http://doi.org/10.1002/pmic.200700470>
- Luhn, S., Berth, M., Hecker, M., & Bernhardt, J. (2003). Using standard positions and image fusion to create proteome maps from collections of two-dimensional gel electrophoresis images. *PROTEOMICS*, 3(7), 1117–1127. <http://doi.org/10.1002/pmic.200300433>
- Matysiak, J., Schmelzer, C. E. H., Neubert, R. H. H., & Kokot, Z. J. (2011). Characterization of honeybee venom by MALDI-TOF and nanoESI-QqTOF mass spectrometry. *Journal of Pharmaceutical and Biomedical Analysis*, 54(2), 273–8. <http://doi.org/10.1016/j.jpba.2010.08.020>
- Mendizabal, F. M. (2005). Abejas. Editorial Albatros, 33-34.
- Mercier, J.-F., & Slater, G. W. (2001). An Exactly Solvable Ogston Model of Gel Electrophoresis. 7. Diffusion and Mobility of Hard Spherical Particles in Three-Dimensional Gels. <http://doi.org/10.1021/MA001544O>
- Mirshafiey, A. (2007). Venom therapy in multiple sclerosis. *Neuropharmacology*, 53(3), 353–361. <http://doi.org/10.1016/j.neuropharm.2007.05.002>
- Morris, J. S., Clark, B. N., & Gutstein, H. B. (2008). Pinnacle: a fast, automatic and accurate method for detecting and quantifying protein spots in 2-dimensional gel electrophoresis data. *Bioinformatics*, 24(4), 529–536. <http://doi.org/10.1093/bioinformatics/btm590>
- Oliveri, G., Anselmi, N., & Massa, A. (2014). Compressive Sensing Imaging of Non-Sparse 2D Scatterers by a Total-Variation Approach Within the Born Approximation. *IEEE Transactions on Antennas and Propagation*, 62(10), 5157–5170. <http://doi.org/10.1109/TAP.2014.2344673>
- Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1), 62–66. Retrieved from [http://web-ext.u-aizu.ac.jp/course/bmclass/documents/Otsu's Thershoding1979.pdf](http://web-ext.u-aizu.ac.jp/course/bmclass/documents/Otsu's%20Thershoding1979.pdf)
- Palagi, P. M., Hernandez, P., Walther, D., & Appel, R. D. (2006). Proteome informatics I: Bioinformatics tools for processing experimental data. *PROTEOMICS*, 6(20), 5435–5444. <http://doi.org/10.1002/pmic.200600273>
- Pando-Robles, R. V., & Lanz-Mendoza, H. (2009). [The significance of proteomics in public health]. *Salud Publica de Mexico*, 51 Suppl 3, S386-94. <http://doi.org/10.1590/S0036-36342009000900004>
- Patnaik, S. and Yang, Y.M. (2012). *Soft Computing Techniques in Vision Science* 395. Springer.
- Peiren, N., de Graaf, D. C., Vanrobaeys, F., Danneels, E. L., Devreese, B., Van Beeumen, J., & Jacobs, F. J. (2008). Proteomic analysis of the honey bee worker venom gland focusing on the mechanisms of protection against tissue damage. *Toxicon*, 52(1), 72–83. <http://doi.org/10.1016/j.toxicon.2008.05.003>

- Peiren, N., Vanrobaeys, F., de Graaf, D. C., Devreese, B., Van Beeumen, J., & Jacobs, F. J. (2005). The protein composition of honeybee venom reconsidered by a proteomic approach. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* (Vol. 1752), <http://doi.org/10.1016/j.bbapap.2005.07.017>
- Prewitt, J. M. (1970). *Object enhancement and extraction* (Vol. 75). Academic Press, New York.
- Raman, B., Cheung, A., & Marte, M. R. (2002). Quantitative evaluation of proteins in one- and two-dimensional polyacrylamide gels using a fluorescent stain. *ELECTROPHORESIS*, 23(14), 2194–2202. [http://doi.org/10.1002/1522-2683\(200207\)23:14<2203::AID-ELPS2203>3.0.CO;2-H](http://doi.org/10.1002/1522-2683(200207)23:14<2203::AID-ELPS2203>3.0.CO;2-H)
- Resende, V. M. F., Vasilj, A., Santos, K. S., Palma, M. S., & Shevchenko, A. (2013). Proteome and phosphoproteome of Africanized and European honeybee venoms. *PROTEOMICS*, 13(17), 2638–2648. <http://doi.org/10.1002/pmic.201300038>
- Rockinger, O. (1996). Pixel -Level Fusion of Image Sequences using Wavelet Frames 1. In *Proceedings of the 16th Leeds applied shape research workshop*, Leeds University Press. Berlin. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.42.6544&rep=rep1&type=pdf>
- Rogers, M., Graham, J., & Tonge, R. P. (2003). Statistical models of shape for the analysis of protein spots in two-dimensional electrophoresis gel images. *PROTEOMICS*, 3(6), 887–896. <http://doi.org/10.1002/pmic.200300421>
- Roti, G., & Stegmaier, K. (2012). Genetic and proteomic approaches to identify cancer drug targets. *British Journal of Cancer*, 106(2), 254–261. <http://doi.org/10.1038/bjc.2011.543>
- Rye, M., & Fargestad, E. M. (2012). Preprocessing of electrophoretic images in 2-DE analysis. *Chemometrics and Intelligent Laboratory Systems*, 117, 70–79. <http://doi.org/10.1016/j.chemolab.2011.09.012>
- Sahoo, P. K., Soltani, S. A. K. C., & Wong, A. K. (1988). A survey of thresholding techniques. *Computer vision, graphics, and image processing*, 41(2), 233-260.
- Sarkar, S., & Das, S. (2013). Multilevel Image Thresholding Based on 2D Histogram and Maximum Tsallis Entropy— A Differential Evolution Approach. *IEEE Transactions on Image Processing*, 22(12), 4788–4797. <http://doi.org/10.1109/TIP.2013.2277832>
- Savelonas, M. A., Mylona, E. A., & Maroulis, D. (2012). Unsupervised 2D gel electrophoresis image segmentation based on active contours. *Pattern Recognition*, 45(2), 720–731. <http://doi.org/10.1016/j.patcog.2011.08.003>
- Schumacher, M. J., & Egen, N. B. (1995). Significance of Africanized Bees for Public Health. *Archives of Internal Medicine*, 155(19), 2038–2043. <http://doi.org/10.1001/archinte.1995.00430190022003>
- Sengar, R. S., Upadhyay, A. K., Singh, M., & Gadre, V. M. (2012). Segmentation of two dimensional electrophoresis gel image using the wavelet transform and the watershed

- transform. In 2012 National Conference on Communications (NCC) (pp. 1–5). IEEE. <http://doi.org/10.1109/NCC.2012.6176861>
- Sengar, R. S., Upadhyay, A. K., Singh, M., & Gadre, V. M. (2016). Analysis of 2D-gel images for detection of protein spots using a novel non-separable wavelet based method. *Biomedical Signal Processing and Control*, 25, 62–75. <http://doi.org/10.1016/j.bspc.2015.10.013>
- Shen, P., Fan, X., Zeng, Z., & Cheng, Y. (2005). Two-dimensional electrophoresis analysis of proteomics based on image feature and mathematical morphology. *Science in China Series B*, 48(4), 361–367. <http://doi.org/10.1360/04yb0128>
- Solomon, J. E., & Harrington, M. G. (1993). A robust high-sensitivity algorithm for automated detection of proteins in two-dimensional electrophoresis gels. *Bioinformatics*, 9(2), 133–139. <http://doi.org/10.1093/bioinformatics/9.2.133>
- Sørensen, T. (1957). «A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons». *Kongelige Danske Videnskabernes Selskab* 5 (4): 1-34.
- Srinark, T., & Kambhamettu, C. (2008). An image analysis suite for spot detection and spot matching in two-dimensional electrophoresis gels. *ELECTROPHORESIS*, 29(3), 706–715. <http://doi.org/10.1002/elps.200700244>
- Sun, C., & Wang, X. (2010). Spot segmentation and verification based on improve marker controlled watershed transform. In 2010 3rd International Conference on Computer Science and Information Technology (pp. 63–66). IEEE. <http://doi.org/10.1109/ICCSIT.2010.5563982>
- Van den Boomgaard, R., & van Balen, R. (1992). Methods for fast morphological image transforms using bitmapped binary images. *CVGIP: Graphical Models and Image Processing*, 54(3), 252–258. [http://doi.org/10.1016/1049-9652\(92\)90055-3](http://doi.org/10.1016/1049-9652(92)90055-3)
- Vincent, L., & Soille, P. (1991). Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 13(6), 583–598. Retrieved from <https://pdfs.semanticscholar.org/a381/9dda9a5f00dbb8cd3413ca7422e37a0d5794.pdf>
- Wheelock, Å. M., & Buckpitt, A. R. (2005). Software-induced variance in two-dimensional gel electrophoresis image analysis. *ELECTROPHORESIS*, 26(23), 4508–4520. <http://doi.org/10.1002/elps.200500253>
- Wu, H.-C., Yen, C.-C., Tsui, W.-H., & Chen, H.-M. (2009). A red line not to cross: Evaluating the limitation and properness of gel image tuning procedures. *Analytical Biochemistry*, 396(1), 42–50. <http://doi.org/10.1016/j.ab.2009.08.041>
- Wu, Y., & Zhang, L. (2011). COMPARISON OF TWO ACADEMIC SOFTWARE PACKAGES FOR ANALYZING TWO-DIMENSIONAL GEL IMAGES. *Journal of Bioinformatics and Computational Biology*, 9(6), 775–794. <http://doi.org/10.1142/S0219720011005665>



- Yiou, P., Shaoli, A., Kebin, L., Tao, W., Kui, F., Hua, Z., ... Jinghui, X. (2013). Evaluation of extraction procedures for 2-DE analysis of aphid proteins. *Journal of Separation Science*, 36(3), 532–539. <http://doi.org/10.1002/jssc.201200642>
- Zacharia, E., Kostopoulou, E., Maroulis, D., Anagnou, N. P., & Pappa, K. I. (2013). 2D-GE spot detection combining multidirectional texture and spatial intensity cues. In *13th IEEE International Conference on BioInformatics and BioEngineering* (pp. 1–4). IEEE. <http://doi.org/10.1109/BIBE.2013.6701555>
- Zhang, L., Fang, Y., Li, R., Feng, M., Han, B., Zhou, T., & Li, J. (2012). Towards posttranslational modification proteome of royal jelly. *Journal of Proteomics*, 75(17), 5327–5341. <http://doi.org/10.1016/j.jprot.2012.06.008>
- Zijdenbos, A. P., Dawant, B. M., Margolin, R. A., & Palmer, A. C. (1994). Morphometric analysis of white matter lesions in MR images: method and validation. *IEEE transactions on medical imaging*, 13(4), 716-724.