

## Herramienta alternativa para la clasificación de células cervicales utilizando solo rasgos del núcleo

### Alternative tool for classification of cervical cells using only features of the nucleus

**MSc. Solangel Rodríguez Vázquez,<sup>I</sup> Ing. Andy Vidal Martínez Borges<sup>II</sup>**

I Universidad de las Ciencias Informáticas. Km 2½ carretera San Antonio de los Baños, Rpto. Torrens, La Lisa, Ciudad de la Habana, Cuba. E-mail: [svazquez@uci.cu](mailto:svazquez@uci.cu)

II Empresa de tecnología para la defensa, XETID. Km 2½ carretera San Antonio de los Baños, Rpto. Torrens, La Lisa, Ciudad de la Habana, Cuba. E-mail: [avmartinez@xetid.cu](mailto:avmartinez@xetid.cu)

---

#### RESUMEN

El cáncer de cérvix uterino representa una de las mayores amenazas de muerte por cáncer entre las mujeres. Con el avance continuo en la medicina y la tecnología, las muertes por esta enfermedad han disminuido significativamente. Las investigaciones referentes a este tema han podido determinar síntomas claves que permiten detectar a tiempo esta enfermedad para darle un tratamiento oportuno. La citología convencional es una de las técnicas más utilizadas, siendo ampliamente aceptada, de bajo costo, y con mecanismos de control. Con el objetivo de aliviar la carga de trabajo a los especialistas, algunos investigadores han propuesto el desarrollo de herramientas de visión computacional para detectar y clasificar las transformaciones en las células de la región del cuello uterino. La presente investigación tiene como objetivo proveer a los investigadores de una herramienta de clasificación automática, aplicable a las condiciones existentes en los centros médicos y de investigación del país. Esta herramienta debe ser capaz de clasificar las células del cuello del útero, basándose solamente en las características extraídas de la región del núcleo y sin utilizar las características del citoplasma, de manera que se reduzca la tasa de falsos negativos en la prueba de Papanicolaou. A partir del estudio realizado, se obtuvo una herramienta haciendo uso de la técnica k-vecinos más cercanos con la distancia manhattan, el cual mostró un alto desempeño manteniendo valores de AUC superiores al 91% y llegando hasta un 97.1% con respecto a los clasificadores SVM y RBF Network, los que también fueron analizados.

**Palabras Clave:** cáncer de cérvix uterino, células del cuello uterino, clasificación de células, kNN, núcleos celulares, SVM, distancias.

---

## ABSTRACT

Cervix cancer is one of the biggest threats of cancer death among women. With continued advances in medicine and technology, deaths from the disease have fallen significantly. The investigations concerning this issue have determined key symptoms to detect the disease in time to give timely treatment. Conventional cytology is one of the most widely used techniques, being widely accepted, inexpensive, and with control mechanisms. In order to alleviate the workload of specialists, some researchers have proposed the development of computer vision tools to detect and classify the changes in the cells of the cervical region. This research aims to provide a tool for automatic classification, applicable to medical conditions and research centers of the country. This tool should be able to classify the cells of the cervix, based solely on the features extracted from the core region without using the characteristics of the cytoplasm, so that the rate of false negative Pap test is reduced. From the study, a tool is obtained using the k nearest-neighbors manhattan distance technique, which showed a high performance maintaining AUC values greater than 91% and reaching 97.1% over classifiers SVM and RBF Network, which were also analyzed.

**Key Words:** cervix cancer, cervical cells, cell classification, kNN, cell nucleus, SVM, distances.

---

## INTRODUCCIÓN

El cáncer de cérvix uterino es una de las enfermedades más frecuentes entre las mujeres, aunque más del 80% se produce en países en vías de desarrollo. El cribado de cáncer de cérvix ha demostrado su efectividad para reducir la incidencia y la mortalidad por esta enfermedad, especialmente cuando se realiza a través de programas poblacionales organizados.

Debido al carácter masivo de su aplicación en la población femenina, la prueba de Papanicolaou genera una considerable carga de trabajo para los laboratorios que analizan en el microscopio los frotis resultantes de esta técnica citológica. En un frotis típico se pueden encontrar hasta 300000 células, lo que limita la productividad a no más de unas 60-80 citologías por día de trabajo y observador.<sup>1</sup> Además de esto, a partir de la experiencia de la aplicación de esta prueba durante muchos años, se ha podido determinar que existen diferentes factores que afectan la calidad de los resultados. Estos son, principalmente, los errores en la toma de las muestras, en su procesamiento y en su lectura e interpretación. Sobre este último caso, la necesidad de analizar una gran cantidad de muestras con muy baja tasa de casos positivos tiende a sesgar el resultado de la evaluación, y además provoca errores debidos a la rutina y a la fatiga de los analistas.

"Si tenemos en cuenta que la esperanza de vida promedio de la mujer en nuestro país es de 80 años, con cada fallecimiento se pierden 30 años de vida potenciales,

---

por lo que es la principal causa de muerte prematura en la población femenina cubana. La carga para la sociedad es entonces muy alta, además de la discapacidad que genera en aquellas que no fallecen".<sup>2</sup>

La presente investigación tiene como objetivo proveer a los investigadores de una herramienta de clasificación automática, aplicable a las condiciones existentes en los centros médicos y de investigación del país.

## CONTENIDO

La necesidad de agilizar y aumentar la calidad de las pruebas citológicas, sirve de motivación para la construcción de una herramienta de apoyo a la detección temprana de posibles enfermedades en el cérvix uterino. Mediante el uso de algoritmos de clasificación, y comparando una base de entrenamiento con una matriz de rasgos extraídos de las imágenes a evaluar, se puede obtener una clasificación de las células obtenidas a través de la prueba citológica.

Una muestra recogida en dichas pruebas contiene varias células en sí misma, las cuales se solapan unas a otras interfiriendo en la delimitación de la región del citoplasma perteneciente a una célula u otra, no siendo así en el caso del núcleo, que, como se muestra en las figuras 1(a) y 1(b) es posible identificarlo. De hecho, la región del núcleo celular resulta más factible de ser segmentada con precisión mediante algoritmos de procesamiento digital de imágenes, siendo este pre-procesamiento un tema no abordado dentro de esta investigación. Atendiendo a investigaciones realizadas por<sup>3,4</sup> y a lo anteriormente expuesto las pruebas y muestras a utilizar en la validación de la herramienta serán con los rasgos extraídos de la región del núcleo.

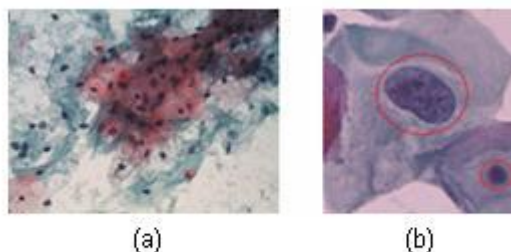
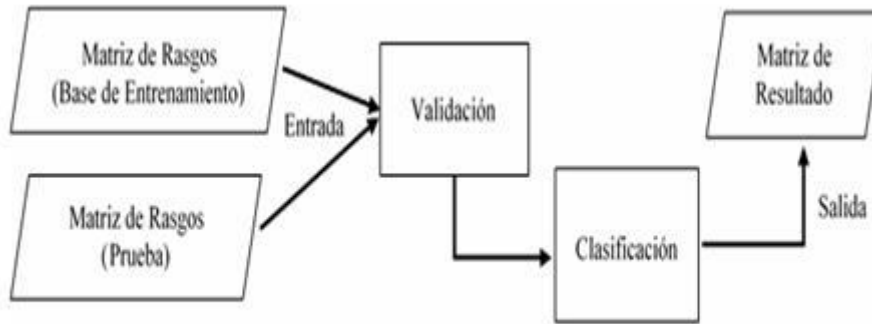


Fig. 1. Células solapadas con identificación del núcleo

### Caracterización de las fases del proceso

En la figura 2 se muestra el diagrama de flujo de la herramienta desarrollada. Dicha solución está compuesta por 2 fases en las cuales las matrices de rasgos (base de entrenamiento ( $m \times n$ )) y (prueba ( $m \times n$ )) son validadas (estructuralmente) para que puedan ser utilizadas por el clasificador, donde son los casos y los rasgos asociados a los casos. Luego de validada la estructura interna de dichas matrices de rasgos, comienza la fase de clasificación la cual brinda como resultado/salida una matriz de resultado ( $m \times 1$ ).



**Fig. 2.** Diagrama de flujo de la solución propuesta

En la fase de validación se propone como entrada que las matrices de rasgos utilizadas tanto como base de entrenamiento, así como los ficheros de prueba, posean una estructura interna de la siguiente manera (Fig. 3):

```

1 1:1410 2:139.882251 3:38.49138889 4:46.8209946 5:0.905531895
1 1:1948 2:166.6101731 3:42.96818515 4:57.88582281 5:0.881852165
1 1:1487 2:143.882251 3:39.32615747 4:48.27557167 5:0.902622978
1 1:1623 2:148.1248917 3:42.35424047 4:48.86739614 5:0.929549009
-1 1:4524 2:269.8650071 3:58.63828838 4:98.82672345 5:0.78061935
-1 1:3078 2:275.6639969 3:58.97125237 4:68.24491694 5:0.509000703
-1 1:2267 2:206.7523087 3:43.91737894 4:66.93235594 5:0.666439375
-1 1:3499 2:231.8650071 3:60.70398481 4:74.32903026 5:0.81786832
  
```

**Fig. 3.** Estructura de la matriz de rasgos (Base de Entrenamiento)

El primer valor representa la clasificación del caso (para los casos normales y para los casos anómalos), posteriormente se observa una estructura, donde representa el rasgo que puede variar desde siempre de forma ascendente, los actúan como separador y representa el valor del rasgo extraído de la célula. Así mismo se evidenciaría en el caso del fichero de prueba (Fig. 4), con la diferencia de que el valor de la clasificación del caso siempre tomaría el valor.

```

0 1:7656.0 2:349.1614716 3:90.73562787 4:108.0589282 5:0.789149269
0 1:4450.0 2:253.2792206 3:66.72134228 4:85.16087607 5:0.87170744
0 1:3496.0 2:240.2081528 3:53.59108946 4:83.52040503 5:0.761387602
0 1:4979.0 2:268.6934342 3:67.66345652 4:93.94496702 5:0.866638308
0 1:2614.0 2:209.3797257 3:44.62966431 4:75.39764428 5:0.749283544
  
```

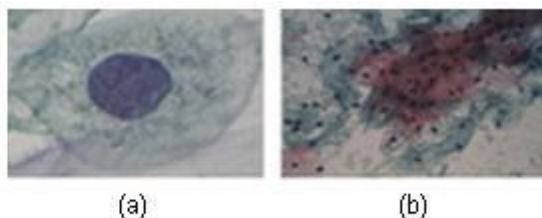
**Fig. 4.** Estructura de la matriz de rasgos (Fichero de Prueba)

Para llevar a cabo la fase de clasificación es necesario haber obtenido con anterioridad las matrices de rasgos (base de conocimiento) y (fichero de prueba) a clasificar. Posteriormente se utiliza la funcionalidad de clasificar y la herramienta devolverá un fichero que contiene la clasificación (enferma o sana) de las células en el mismo orden en el que se encuentran en el fichero de prueba.

## Descripción y metodología de uso de los algoritmos propuestos

Varios métodos han sido propuestos para la clasificación de las células en las imágenes de la prueba de Papanicolaou y que se refieren a las técnicas tales como clasificadores bayesianos,<sup>5</sup> redes neuronales artificiales,<sup>6</sup> máquinas de soporte vectorial (SVM)<sup>7</sup> y vecinos más cercanos.<sup>8</sup> Debe tenerse en cuenta que la mayoría de estos métodos utilizan imágenes pre-segmentadas que contienen solo una célula, por lo que la segmentación correcta del núcleo y del citoplasma es factible (Fig. 5 (a)).

En las imágenes que contienen grupos de células (Fig. 5 (b)), la detección de la frontera del citoplasma es un problema difícil, y hasta ahora, no se ha encontrado ninguna técnica en la literatura estudiada que mencione algoritmos con buenos resultados en la delimitación de las fronteras de la región del citoplasma. Sin embargo, la detección y segmentación de los núcleos de las imágenes que contienen células superpuestas y agrupaciones de células ha sido abordado con éxito en varios estudios.<sup>9,10</sup>



**Fig. 5.** (a) Imagen de una célula simple, (b) imagen de células superpuestas

Los métodos que se refieren a la clasificación de imágenes de Papanicolaou se basan en el cálculo de las características (Tabla 1) extraídas tanto de la región del núcleo como del citoplasma. En el caso de la presente investigación es importante este análisis, debido a que la entrada para el clasificador es una matriz (donde los casos son cada una de las células) con cada uno de los rasgos extraídos de las imágenes.<sup>11</sup>

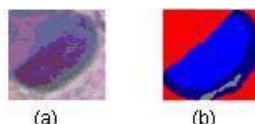
En algunas investigaciones como es el caso de<sup>3</sup> se hace uso de los 9 rasgos del núcleo y técnicas como spectral clustering y fuzzy C-means con reducción de la dimensionalidad, a diferencia de la presente investigación que se dirige hacia la selección de un clasificador sin reducción donde solo se utilizan 5 rasgos de los 9 como son: el área, el perímetro, el diámetro corto, diámetro más largo y la redondez.

Esta selección de rasgos persigue demostrar que, a partir de un conjunto simplificado de datos geométricos básicos, es posible realizar de forma efectiva la clasificación binaria de imágenes en la prueba de Papanicolaou.<sup>9</sup>

La extracción de rasgos es uno de los pasos fundamentales en el procesamiento de imágenes debido a que mientras mejor sea la selección de los atributos, más acertada será la clasificación final de las células. Esto hace que la adecuada selección de los rasgos sea una de las limitantes en las investigaciones para la clasificación de imágenes, de no contar con los rasgos apropiados los resultados obtenidos por el clasificador no tendrán la calidad que se necesita.<sup>12</sup> Para el caso de la presente investigación, las matrices de rasgos utilizadas fueron extraídas de las imágenes ya previamente segmentadas (Fig. 6(b)) pertenecientes a la base de datos Herlev.<sup>13</sup>

**Tabla 1.** Características extraídas de cada imagen en la base de datos<sup>13</sup>

RASGOS CITOPLASMA	DEL	RASGOS NÚCLEO	DEL
1. Área		1. Área	
2. Brillo		2. Brillo	
3. Diámetro Corto		3. Diámetro Corto	
4. Diámetro Largo	más	4. Diámetro más Largo	
5. Elongación		5. Elongación	
6. Redondez		6. Redondez	
7. Perímetro		7. Perímetro	
8. Máxima <sup>1</sup>		8. Máxima <sup>1</sup>	
9. Mínima <sup>1</sup>		9. Mínima <sup>1</sup>	
10. Posición del Núcleo			
11. Tamaño Núcleo/Citoplasma			
<sup>1</sup> El número de píxeles con el valor de intensidad máximo / mínimo en una zona de 3x3 de la región específica.			



**Fig. 6.** Células anómalas. (a) Imagen RGB, (b) Imagen Segmentada en la cual la región correspondiente al núcleo aparece resaltada en azul claro y a la cual se le realiza el cálculo de los rasgos propuestos

Para ello se hizo uso de la herramienta Matlab y de las funciones propias de la misma. De esta forma se desarrolló un algoritmo que toma aleatoriamente de la base de datos Herlev un 80% de las imágenes y extrae de ellas (Fig. 6 (b)) los rasgos antes mencionados. A continuación el algoritmo crea una matriz con los rasgos extraídos de las imágenes, la que será utilizada en el entrenamiento del clasificador.

Posteriormente el algoritmo extrae los rasgos del 20% de las imágenes restantes y conforma la matriz de rasgos que se utilizará para realizar las pruebas. Estas matrices tienen como característica fundamental que poseen vectores de rasgos diferentes lo que posibilita una evaluación correcta del funcionamiento del clasificador.

Para dar cumplimiento al objetivo propuesto en la presente investigación se definió utilizar para la comparación y posterior selección, tres algoritmos de clasificación: las máquinas de soporte vectorial (SVM), la búsqueda por el vecino más cercano (kNN) y las redes neurales artificiales (RNA). Con el fin de estimar la capacidad de discriminación de los parámetros internos para el funcionamiento de dichos algoritmos se realizaron una serie de experimentos donde se tuvieron en cuenta los



kernels para SVM, las distancias en el caso de kNN y dos RNA (Perceptrón Multicapa, RBF Network). Para esto se utilizó un esquema experimental, validación cruzada de k particiones (k-fold cross validation). Se realizó un particionamiento de los datos de la forma antes descrita y se realizaron siete ejecuciones. Se comprobaron los mismos conjuntos de datos para cada clasificador, tanto en las comparaciones internas de estos algoritmos como en la posterior comparación entre los mismos. Para las comparaciones de los resultados en cada algoritmo, de los conjuntos de datos obtenidos en el particionamiento, se utilizaron las tres particiones que mejores resultados mostraron en cuanto a las medidas Pn, área bajo la curva ROC (AUC) y las medias armónicas H y F.

A través de una comparación realizada por la autora se obtuvieron las características internas con las cuales los algoritmos muestran los mejores resultados en la clasificación de este tipo de células. El algoritmo SVM con el kernel RBF mostró resultados en la clasificación que se comportan entre un 78% y 86% de predictividad negativa (Pn) además de su bajo costo computacional en comparación del kernel lineal que mostró una Pn entre 79% y 85% con un elevado costo computacional. En el caso de las redes neuronales comparadas se evidencia que la RBF Network es más eficiente en el momento de la clasificación comportándose entre un 81 y 91% de Pn en comparación con la red neuronal MLP que mantuvo un comportamiento entre un 81% y 86%, así como en el caso del algoritmo kNN, la distancia Manhattan evidenció obtener resultados en la clasificación que se comportan entre un 84 y 93% de Pn con respecto a la distancia mahalanobis que mostró un 84 y 90% de Pn. Para la selección de dichas características se tuvieron en cuenta además los valores obtenidos en cuanto a los índices de efectividad que a continuación también se utilizan en esta investigación. Dichos algoritmos arrojaron resultados que demuestran la eficiencia de cada uno en la clasificación de las células cervicales.

En la presente investigación se utilizaron, para evaluar la efectividad de los diferentes clasificadores utilizados, los índices de efectividad, conocidos como sensibilidad, especificidad, predictividad positiva y negativa, así como la tasa de clasificación correcta. Otros índices de efectividad empleados en esta investigación para evaluar las distancias en el clasificador fueron el área bajo la curva ROC, la predictividad negativa y las medidas F y H, dadas por:

$$F - measure = \frac{2 \times (Precision \times S_e)}{Precision + S_e} \quad (1)$$

$$H - mean = \frac{2 \times (S_p \times S_e)}{S_p + S_e} \quad (2)$$

## RESULTADOS Y DISCUSIÓN

A fin de determinar la calidad relativa de los resultados se realizó una comparación entre los clasificadores anteriormente mencionados. Se utilizó una programación propia escrita en lenguaje Java y se utilizaron además las implementaciones disponibles en el paquete de Minería de Datos WEKA.

### Análisis de los resultados de acuerdo a la medida F

El objetivo de este estudio es realizar un análisis referente al comportamiento del índice de efectividad conocido como medida o media armónica (-measure, en la literatura en inglés). La media armónica determina hasta qué punto los resultados obtenidos se asemejan a los que se hubieran logrado con una categorización manual real por cada iteración de prueba. De acuerdo a los datos ofrecidos en la tabla 2 se muestra que el valor de la medida se mantuvo por encima del 86% entre los tres clasificadores donde la media prevaleció por encima del 90%. En la corrida de las iteraciones de prueba el clasificador kNN obtuvo un 93.4% de medida (mayor valor obtenido).

**Tabla. 2.** Comportamiento de la medida  $F$  en las iteraciones de pruebas realizadas por clasificador

	Conjuntos de Datos						
	1	2	3	4	5	6	7
<i>SV</i>	0.8	0.8	0.9	0.9	0.9	0.8	0.9
<i>M</i>	6	6	2	2	1	9	2
<i>kNN</i>	0.8	0.8	0.9	0.8	0.9	0.9	0.9
	8	7	1	9	3	0	2
<i>RNA</i>	0.9	0.8	0.9	0.9	0.9	0.9	0.9
	0	8	2	1	2	0	3

### Análisis de los resultados de acuerdo a la Predictividad Negativa (Pn)

El objetivo de este estudio es realizar un análisis referente al comportamiento de la medida de evaluación predictividad negativa (Pn). La Pn es la probabilidad de que el clasificador detecte la célula como normal siendo en realidad el resultado de la prueba diagnóstica negativo. De acuerdo a los datos que se ofrecen en la tabla 3 se muestra que el valor de Pn se mantuvo por encima del 73% entre los tres clasificadores donde la media prevaleció por encima del 83%. En la corrida de las iteraciones de prueba el clasificador kNN obtuvo un 93% de Pn (mayor valor obtenido).

**Tabla. 3.** Comportamiento de la medida Pn en las iteraciones de pruebas realizadas por clasificador

	Conjuntos de Datos						
	1	2	3	4	5	6	7
<i>SV</i>	0.7	0.7	0.8	0.8	0.8	0.8	0.8
<i>M</i>	4	9	7	5	1	2	9
<i>kNN</i>	0.7	0.7	0.8	0.8	0.8	0.8	0.9
	7	9	5	0	8	3	3
<i>RNA</i>	0.7	0.8	0.8	0.8	0.8	0.8	0.9
	3	0	9	4	6	1	1

### Área bajo la curva ROC (A-ROC o AUC)

El objetivo de este estudio es realizar un análisis referente al comportamiento de la medida de evaluación área bajo la curva ROC (A-ROC o AUC). El AUC según<sup>14</sup> y <sup>15</sup>, se define como la probabilidad de clasificar correctamente un par de individuos



(sano y enfermo), seleccionados al azar de la población, mediante los resultados obtenidos al aplicar la prueba diagnóstica. Los datos ofrecidos en la tabla 4 muestran que el valor de la AUC se mantuvo por encima del 81% entre los tres clasificadores donde la media prevaleció por encima del 91%. En la corrida de las iteraciones de prueba el clasificador kNN y el RBF Network obtuvieron un 97.1% de AUC (mayor valor obtenido). En la tabla 4 se muestra la estabilidad del clasificador kNN por encima de SVM y RBF Network, donde obtuvo los mayores valores del área bajo la curva ROC en todas las iteraciones de pruebas realizadas.

**Tabla. 4.** Comportamiento de la medida de evaluación AUC en las iteraciones de pruebas realizadas por clasificador

	Conjuntos de Datos						
	1	2	3	4	5	6	7
<i>SV</i>	0.8	0.8	0.8	0.9	0.9	0.8	0.8
<i>M</i>	2	1	9	0	0	5	8
<i>kNN</i>	0.9	0.9	0.9	0.9	0.9	0.9	0.9
	1	4	5	4	7	2	5
<i>RNA</i>	0.8	0.9	0.9	0.9	0.9	0.9	0.9
	6	2	4	4	7	2	3

### Análisis comparativo de los clasificadores

Luego del análisis realizado de acuerdo a las medidas de evaluación aplicadas a los clasificadores es necesario evaluar si existe alguna diferencia significativa que demuestre la superioridad de un clasificador sobre los demás. Para ello se siguieron las recomendaciones realizadas por Demšar<sup>16</sup> y las extensiones presentadas por García y Herrera.<sup>17</sup> Se procedió a realizar un análisis estadístico a través de pruebas no paramétricas para k muestras no relacionadas mediante el Test de Friedman<sup>16</sup> con el objetivo de probar la hipótesis nula de que todos los clasificadores alcanzan los mismos resultados en promedio. Como pruebas post-hoc se aplicó primeramente el test Bonferroni-Dunn<sup>16</sup> para definir que existen diferencias significativas entre el clasificador mejor rankeado y el siguiente. Para contrastar los resultados se aplicó el test de Holm.<sup>16</sup>

**Tabla. 5.** Resultados obtenidos para AUC por los clasificadores

<i>Particiones</i>	<i>SVM</i>	<i>kNN</i>	<i>RNA</i>
1	0.82	0.91 3	0.865
2	0.814	0.94 8	0.928
3	0.894	0.95 2	0.94
4	0.901	0.94 5	0.94
5	0.90	0.97 1	0.971
6	0.856	0.92 5	0.92
7	0.888	0.95 1	0.939
<i>Ranking</i>	2.99	1.07	1.92
<i>Posición</i>	3	1	2

**Tabla. 6.** Resultados obtenidos para Pn por los clasificadores

<i>Particiones</i>	<i>SVM</i>	<i>kNN</i>	<i>RNA</i>
1	0.74	0.77	0.73
2	0.79	0.79	0.80
3	0.87	0.85	0.89
4	0.85	0.8	0.84
5	0.81	0.88	0.86
6	0.82	0.83	0.81
7	0.89	0.93	0.91
<i>Ranking</i>	2.21	1.78	2.0
<i>Posición</i>	3	1	2

**Tabla. 7.** Resultados obtenidos para medida F por los clasificadores

<i>Particiones</i>	<i>SVM</i>	<i>kNN</i>	<i>RNA</i>
1	0.863	0.88	0.909
2	0.866	0.87 2	0.884
3	0.923	0.91 8	0.928
4	0.923	0.89 7	0.918
5	0.914	0.93 4	0.929
6	0.895	0.90 1	0.902
7	0.922	0.92 7	0.933
<i>Ranking</i>	2.57	2.14	1.28
<i>Posición</i>	3	2	1

Las tablas 5, 6 y 7 muestran los resultados experimentales de los tres clasificadores para cada una de las particiones realizadas del conjunto de datos. Las dos últimas filas muestran el ranking de cada uno de los clasificadores y su posición. Después de analizar estadísticamente los resultados para detectar si existen diferencias significativas entre los clasificadores para cada una de las medidas utilizadas, se comprobó que el test de Friedman rechazó la hipótesis nula para AUC con un valor de  $p < 0.05$  y para la medida F con un valor de  $p < 0.05$ , mientras que para Pn no se encontraron diferencias significativas. Por esta razón solo se aplicaron los test post-hoc para la medida AUC y medida F. Para ambas medidas los test Bonferroni-Dunn y de Holm rechazaron la hipótesis nula para valores de  $p < 0.05$  y respectivamente, con un valor de confianza (Tabla 8) y (Tabla 9). Los resultados presentados en la tabla 8, demuestran que el clasificador kNN tiene un mejor desempeño que SVM y equivalente que RNA para el particionamiento realizado de los datos. En el caso de la tabla 9, los resultados presentados demuestran que el clasificador RNA tiene un mejor desempeño que SVM y equivalente que kNN para el particionamiento realizado de los datos.

**Tabla 8.** Orden de los clasificadores por el valor de  $p$  y ajustando el valor de  $\alpha$  para el procedimiento Holm con  $\alpha = 0.05$  para la medida AUC

$i$	Clasificador	$\bar{z} = \frac{(R_0 - R_i)}{SE}$	$p$	$\alpha/i$
2	SVM	3.60	$3.08 \times 10^{-4}$	0.025
1	RNA	1.60	0.10	0.05

**Tabla 9.** Orden de los clasificadores por el valor de  $p$  y ajustando el valor de  $\alpha$  para el procedimiento Holm con  $\alpha = 0.05$  para la medida  $F$

$i$	Clasificador	$\bar{z} = \frac{(R_0 - R_i)}{SE}$	$p$	$\alpha/i$
2	SVM	2.40	0.0	0.025
1	RNA	1.60	0.1	0.05

De acuerdo al análisis realizado, se evidencia que existen diferencias significativas entre los clasificadores kNN y RNA con respecto a SVM de acuerdo a la medida AUC y a la medida  $F$ , no así entre kNN y RNA. Por tal motivo, es necesario realizar un análisis del comportamiento de los clasificadores kNN y RNA de acuerdo a los valores obtenidos en ambas medidas en las tablas 2 y 4. Se evidencia que el mayor valor obtenido en ambas medidas entre ambos clasificadores fue obtenido por kNN, con un valor de 93.4% en la medida  $F$  y del 97.1% para la medida AUC.

Luego de evaluadas las medidas en cada clasificador, se evidencia que el clasificador con mejor desempeño en la clasificación de células cervicales, para el conjunto de datos estudiados, es kNN haciendo uso de la distancia Manhattan. Dicho clasificador presenta los mejores resultados para las medidas evaluadas, con valores promedios de medida  $F$  de 90.4%,  $P_n$  de 83% y AUC de 94.3%. Además este clasificador mostró un alto desempeño manteniendo los valores de AUC superiores al 91% y llegando hasta un 97.1% (Tabla 4). Por estas razones se recomienda utilizar el clasificador kNN con la distancia Manhattan en la herramienta de clasificación desarrollada.

### Comparación con otros enfoques de la literatura estudiada

La tabla 10 muestra la comparación de la evaluación estadística obtenida durante la investigación con los resultados obtenidos en otras investigaciones analizadas. Dichas investigaciones utilizan al igual que la presente investigación la base de datos Herlev, aunque no especifican la forma en que se realizó la partición de los conjuntos de rasgos para las pruebas y el entrenamiento.

**Tabla 10.** Comparación de la clasificación en términos de la media armónica  $H$  ( $H$ -mean)

Clasificadores	# de Rasgos utilizados	% Media Armónica ( $H$ -mean)
Fuzzy C-means <sup>3</sup>	7	90.58
Spectral Clustering <sup>3</sup>	9	88.77
SVM (RBF) <sup>4</sup>	9	96.89
$kNN$ (Manhattan)	5	91.35

Se establece una comparación en función de la media armónica debido a que es la medida de evaluación utilizada por dichas investigaciones. El resultado obtenido por la presente investigación de acuerdo a la media armónica muestra que se logra superar los resultados obtenidos en,<sup>3</sup> mientras que son ligeramente inferiores a<sup>14</sup> aunque se debe tener en cuenta que este utiliza un mayor número de rasgos.

## CONCLUSIONES

Según,<sup>2</sup> en el año 2011 se reportaron en Cuba 1334 casos nuevos de cáncer cervicouterino, y unas 455 mujeres fallecieron por esta causa. Atendiendo a tal situación, se plantea que los programas de pesquisa y diagnóstico precoz del cáncer cervicouterino y sus lesiones precursoras cobran cada vez mayor relevancia.

La presente investigación se basa fundamentalmente en la realización de una herramienta de clasificación automática que permita la reducción de los falsos negativos resultantes de la clasificación de las imágenes en la prueba de Papanicolaou. A medida que avanzó la investigación y en la realización de las pruebas se evidenció que las células sanas pertenecientes a la clasificación Normal\_Columnar de la base de datos Herlev, son células propensas a falsos positivos debido a que sus características son similares a las pertenecientes a las células enfermas.

A continuación se enuncian algunos de los resultados arrojados:

- El empleo de técnicas de clasificación en la realización de pruebas a células cervicales permite obtener una clasificación con baja tasa de falsos negativos.
- A través de una comparación fueron obtenidos los parámetros propios de funcionamiento (kernel, distancias, números de clusters) con los cuales los algoritmos muestran los mejores resultados en la clasificación de este tipo de células. El algoritmo SVM mostró ser más apropiado cuando hace uso del kernel RBF, debido a su comportamiento con respecto al kernel lineal, y a su bajo costo computacional. En el caso de las redes neuronales comparadas se evidencia que la RBF Network es más eficaz en el momento de la clasificación por los resultados obtenidos, así como en el caso del algoritmo  $kNN$ , la distancia Manhattan evidenció obtener mejores resultados además de un bajo costo computacional. Dichos algoritmos arrojaron resultados que demuestran su eficacia en la clasificación de las células cervicales.

- Se demostró que a partir de un conjunto simplificado de datos geométricos básicos, es posible realizar de forma efectiva la clasificación binaria de imágenes en la prueba de Papanicolaou.

- Los procedimientos estadísticos llevados a cabo por cada índice de efectividad evaluado, determinaron que con respecto a la medida AUC y a la medida F los clasificadores RNA y kNN poseían diferencias significativas con respecto a SVM, no siendo así entre kNN y RNA para ambas medidas. Luego del análisis estadístico kNN mostró un alto desempeño manteniendo los valores de AUC superiores al 91% y llegando hasta un 97.1%.

La herramienta de clasificación desarrollada como resultado de la presente investigación, si bien no es una herramienta de diagnóstico, sirve como ayuda en el proceso. La misma es capaz de clasificar entre células anómalas y normales obteniendo resultados alentadores. Las iteraciones de prueba realizadas y la comparación de los resultados obtenidos en otras investigaciones relacionadas con el dominio abordado demostraron la reducción de falsos negativos que es uno de los propósitos principales de la presente investigación.

## REFERENCIAS BIBLIOGRÁFICAS

1. Lorenzo J.V Rodríguez I. Aplicación de técnicas de visión computacional en la prueba de Papanicolaou. *Medicentro Electrónica*. 2012. 16(3): p. 196-198.
2. Acosta L.F. El diagnóstico temprano es garantía de vida. In *Periódico Granma*. 2012: Ciudad de la Habana.
3. Plissiti M, Nikou C. Cervical Cell Classification Based Exclusively on Nucleus Features. In *Image Analysis and Recognition*. Campilho A, Kamel M, Editors. 2012, Springer Berlin Heidelberg. p. 483-490.
4. Plissiti M.E, Nikou C. On the importance of nucleus features in the classification of cervical cells in Pap smear images. University of Ioannina. 2012.
5. Riana D, Murni A. Performance evaluation of Pap smear cell image classification using quantitative and qualitative features based on multiple classifiers. In *International Conference on Advanced Computer Science and Information Systems, ACSIS*. 2009.
6. Mat-Isa N.A, Mashor M.Y, Othman N.H. An automated cervical pre-cancerous diagnostic system. *Artificial Intelligence in Medicine*. 2008. 42(1): p. 1-11.
7. Huang P.-C, et al. Quantitative Assessment of Pap Smear Cells by PC-Based Cytopathologic Image Analysis System and Support Vector Machine, In *Medical Biometrics*. D. Zhang, Editor. 2007, Springer Berlin Heidelberg. p. 192-199.
8. Marinakis Y, Dounias G, Jantzen J. Pap smear diagnosis using a hybrid intelligent scheme focusing on genetic algorithm based feature selection and nearest neighbor classification. *Computers in Biology and Medicine*. 2009. 39(1): p. 69-78.
9. Plissiti M.E, Nikou C, Charchanti A. Combining shape, texture and intensity features for cell nuclei extraction in Pap smear images. *Pattern Recognition Letters*. 2011. 32(6): p. 838-853.



10. Plissiti M.E, et al. Automated Detection of Cell Nuclei in Pap Smear Images Using Morphological Reconstruction and Clustering. IEEE Transactions on information technology in biomedicine. 2011. 15(2): p. 233-241.
11. Lorenzo J.V, et al. Cervical Cell Classification Using Features Related to Morphometry and Texture of Nuclei, in Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications. 2013, Springer. p. 222-229.
12. Velezmore G.A.B, Villafuerte D.F. Factores de riesgo que pronóstican el hallazgo de citologías cervicales anormales en dos poblaciones: mujeres de obreros de construcción civil vs. mujeres control en la posta médica "Construcción Civil" ESSALUD, de junio a septiembre del 2000, in Facultad de Medicina Humana. 2001, Universidad Nacional Mayor de San Marcos: Lima, Perú. p. 67.
13. Jantzen J, et al. Pap-smear Benchmark Data For Pattern Classification. In Proc. NiSIS 2005. 2005, Nature inspired Smart Information Systems (NiSIS): Albufeira, Portugal. p. 1-9.
14. Bamber D. The area above the ordinal dominance graph and the area below the receiver operating characteristic graph. Journal of Mathematical Psychology. 1975. 12(4): p. 387-415.
15. Hanley J.A, McNei IB.J. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 1982. 143(1): p. 29-36.
16. Demšar J. Statistical Comparisons of Classifiers over Multiple Data Sets. The Journal of Machine Learning Research. 2006. 7: p. 1-30.
17. García S, Herrera F. An Extension on "Statistical Comparisons of Classifiers over Multiple Data Sets" for all Pairwise Comparisons. Journal of Machine Learning Research. 2008. 9: p. 2677-2694.

Recibido: 22 de marzo de 2016.

Aprobado: 12 de mayo de 2016.