

**Maquinas De Soporte Vectorial (svm) Para La Detección De Nódulos  
Pulmonares En Tomografía Axial Computarizada (TAC).**

**Willmar Ledezma Garrido**

**Pontificia Universidad Javeriana  
Facultad de Ingeniería  
Maestría en Ingeniería Electrónica  
Bogotá  
2012**

**Maquinas De Soporte Vectorial (svm) Para La Detección De Nódulos  
Pulmonares En Tomografía Axial Computarizada (TAC).**

**Willmar Ledezma Garrido**

**Trabajo de grado para optar al Título de Magister en Ingeniería Electrónica**

**Director  
Leonardo flórez Ph.D.  
Ingeniero de Sistemas**

**Pontificia Universidad Javeriana  
Facultad de Ingeniería  
Maestría en Ingeniería Electrónica  
Bogotá  
2012**



## Contenido

INTRODUCCIÓN.....	8
1. OBJETIVOS.....	10
1.1 GENERAL.....	10
1.2 ESPECÍFICOS.....	10
2. ESPECIFICACIONES.....	11
3. MARCO TEÓRICO.....	12
3.1 MARCO CLÍNICO .....	12
3.1.1 Pulmones, nódulos y cáncer de pulmonar.....	12
3.1.2 Tomografía Axial Computarizada (TAC).....	19
3.2 PROCESAMIENTO DE IMAGENES.....	20
3.2.1 Crecimiento de regiones.....	20
3.2.2 Maquinas de vectores soporte (svm).....	21
3.2.3 RBF gaussiana.....	26
3.2.4 Coeficiente kappa.....	27
4. PROPUESTA DE DETECCIÓN .....	29
4.1 REGION DE INTERÉS.....	29
4.2 CRECIMIENTO DE REGIONES.....	29
4.3 EXTRACCIÓN DE CARACTERÍSTICAS.....	30
4.3.1 Desviación estándar.....	31
4.3.2 Área.....	31
4.3.3 Perímetro.....	31
4.3.4 Promedio de la intensidad.....	31
4.3.5 Circularidad.....	31
4.3.6 Potencia.....	31
4.3.7 Distancia de la pared del pulmón.....	32
4.4 SVMs.....	32
4.5 INTEGRACION DEL ALGORITMO SVM CON EL VECTOR DE CARACTERÍSTICAS.....	33
4.6 DIAGRAMA DE BLOQUES.....	34
4.6.1 Esquema general del proyecto.....	34
4.6.2 Esquema de la segmentación.....	35
4.6.3 Esquema de la extracción de características.....	36
4.6.4 Esquema de la svm.....	37
5. RESULTADOS.....	38
5.1 DESCRIPCIÓN DEL PROTOCOLO DE EXPERIMENTACIÓN.....	38
5.2 PRESENTACIÓN DE RESULTADOS.....	39
5.2.1 Nódulos grandes.....	39

5.2.2 Nódulos pequeños.....	41
5.2.3 Nódulos totales.....	42
5.2.4 Integración mejores resultados.....	44
5.3 ANÁLISIS DE RESULTADOS.....	46
5.3.1 Nódulos grandes.....	46
5.3.2 Nódulos pequeños.....	46
5.3.1 Nódulos totales.....	46
6. CONCLUSIONES.....	48
6.1 VENTAJAS DE LA PROPUESTA DE DETECCIÓN CUANTIFICADAS EN TÉRMINOS DE LOS RESULTADOS.....	48
6.2 LIMITACIONES ENCONTRADAS A PARTIR DE LOS RESULTADOS.....	48
6.3 PERSPECTIVA.....	49
BIBLIOGRAFIA.....	50
ANEXOS.....	52

## Tablas

Tabla 1. tabla para el calculo del coeficiente kappa.....	27
Tabla 2. medición de coeficiente kappa y grado de acuerdo.....	28
Tabla 3. evaluación de nódulos resultados del svm para nódulos grandes.....	40
Tabla 4. evaluación de nódulos resultados del svm para nódulos pequeños.....	41
Tabla 5. evaluación de nódulos resultados del svm para nódulos totales.....	43
Tabla 6. mejores resultados de la evaluación kappa para nódulos grandes, pequeños y totales con referencia sigma.....	45

## Figuras

Figura 1. Top 10 de los cáncer mas comunes entre hombres y mujeres x cada 100mil habitantes en estados unidos en 2008.....	15
Figura 2. Top 10 de los cáncer mas comunes entre hombres x cada 100mil habitantes en estados unidos en 2008.....	16
Figura 3. Top 10 de los cáncer mas comunes entre mujeres x cada 100mil habitantes en estados unidos en 2008.....	16
Figura 4. Top 10 de la mortalidad del cáncer entre hombre y mujeres x cada 100mil habitantes en estados unidos en 2008.....	17
Figura 5. Imagen TAC con presencia de Nódulo pulmonar.....	18
Figura 6. conjunto de imágenes tomográficas en su corte axial.....	19
Figura 7. calculo de vectores soporte y maximización del margen.....	23
Figura 9. maximización del margen con soft margin.....	26
Figura 10. imagen de un nódulo enmarcado en una región de interés.....	29
Figura 11. nódulo segmentado por crecimiento de regiones.....	30
Figura 12. medición de la distancia del nódulo a la pared del pulmón.....	32
Figura 13. evaluación kappa resultados del svm para nódulos grandes.....	41
Figura 14. evaluación kappa resultados del svm para nódulos pequeños.....	42
Figura 15. evaluación kappa resultados del svm para nódulos totales.....	44
Figura 16. mejores resultados de la evaluación kappa para nódulos grandes, pequeños y totales con referencia sigma.....	45

## INTRODUCCIÓN

Un cáncer pulmonar en etapa inicial se puede diagnosticar precozmente al detectar un nódulo pulmonar en una Radiografía de Tórax o en una Tomografía Axial Computada, siendo esta última mucho más sensible y exacta. Los nódulos pulmonares pueden ser benignos (infecciones, cicatrices o tumores benignos) o representar un cáncer pulmonar en etapa inicial. La importancia de realizar un diagnóstico temprano, es que en etapas iniciales la cirugía es el único tratamiento que permite curar el cáncer pulmonar, este trabajo está orientado a la detección de nódulos mediante el uso de máquinas de vectores de soporte (svm); para el desarrollo del proyecto se cuenta con 50 estudios TAC de tórax con un espesor de 1,25mm en los cuales hay presencia de nódulos en los pulmones sobre los cuales se va a hacer el desarrollo, todos los nódulos detectados en los estudios tomográficos fueron previamente revisados por expertos en el diagnóstico de estos.

El desarrollo del proyecto se comienza con una selección de una región de interés por parte del experto en diagnóstico de cáncer pulmonar, que se define a partir de la imagen y donde se puedan tener sospechas de un nódulo pulmonar, a partir de esto se hace una extracción de la posición y el tamaño de la región, se introduce al algoritmo de segmentación para el cual se seleccionó el método de crecimiento de regiones para hacer la segmentación del posible nódulo, este proceso no lo hace exento de los posibles falsos nódulos, pero nos enmarca la región de análisis para las características a extraer del posible nódulo al cual se le extraen las características de las zonas previamente segmentadas y estas son: desviación estándar, área, perímetro, promedio, circularidad, potencia y distancia de la pared del pulmón con base en estas características se hace el entrenamiento de la svm creando una matriz que está compuesta con la suma de vectores resultado de la extracción de características y un vector o etiqueta, que define a que grupo pertenecen cada uno de los vectores; si es un posible nódulo o no, esto se define a partir de las anotaciones hechas por expertos sobre los estudios. Para probar el entrenamiento hecho se usa un vector o una matriz de la misma forma del entrenamiento pero sin etiqueta para que así, la svm realice una etiqueta sobre el vector o matriz introducido obteniendo como respuesta un vector con las etiquetas de identificación de verdadero o falso nódulo; junto con la matriz de pruebas a cada vector se le asignó una etiqueta con la observación del experto en el diagnóstico de los nódulos pulmonares; con las etiquetas resultantes de la svm se



prosigue con la evaluación, paralelo a esto se hizo la variación en la svm del parámetro sigma puesto que la svm funciona con kernel RBF Gaussiano.

Para la evaluación y repuesta de la svm se usó el coeficiente de kappa como medición de la concordancia entre el observador de los nódulos contra la respuesta del entrenamiento hecho a la svm, de este mismo proceso se derivan otras 3 mediciones que son la sensibilidad, especificidad y precisión teniendo así una visión clara del rendimiento de proyecto.

## **1. OBJETIVOS**

### **1.1 GENERAL**

- Detectar nódulos pulmonares por svm en imágenes TAC.

### **1.2 ESPECÍFICOS**

- Caracterizar clínicamente los nódulos pulmonares.
- Seleccionar los algoritmos para la segmentación de nódulos pulmonares.
- Seleccionar las características a extraer para svm.
- Extraer características para el vector de svm.
- Ejecutar el svm para detectar nódulos pulmonares.
- Evaluar el rendimiento y verificar resultados.

## 2. ESPECIFICACIONES

Para el desarrollo del proyecto se tienen 50 estudios TAC<sub>[13]</sub> con espesor de 1.25mm en los cuales hay presencia de nódulos pulmonares y están identificados y descritos por expertos en el diagnóstico de enfermedades en los pulmones. Teniendo esta base de estudios se hace la distinción de los nódulos grandes(>5mm) de los pequeños(<5mm) teniendo así las imágenes para la separación de región de interés.

Con una región de interés definida por el experto en diagnóstico, se introduce al algoritmo que hace la extracción de características, se generan los vectores para el entrenamiento de la svm que nos va a dar como resultado una detección del nódulo positivo, como el negativo, a partir de estos resultados se hace la evaluación por coeficientes de concordancia de svm contra observador.

### 3. MARCO TEÓRICO

#### 3.1 MARCO CLÍNICO

##### 3.1.1 Pulmones, nódulos y cáncer de pulmonar<sup>1</sup>

- **Los Pulmones**

Los pulmones son los órganos encargados de oxigenar la sangre y expulsar el dióxido de carbono, un producto de desecho producido por las células del cuerpo, Y Participa además en otras importantes funciones metabólicas y cardiovasculares. Los bronquios, en tanto transportan el aire inspirado hacia los pulmones. Desde la tráquea se van dividiendo en ramificaciones sucesivamente, dando origen a bronquios cada vez más pequeños hasta llegar a los alvéolos que son pequeños saquitos rodeados de vasos sanguíneos de pequeño calibre (capilares). Entre el aire contenido en los alvéolos y los capilares se produce el intercambio gaseoso.

- **Nódulos pulmonares**

Es una mancha (lesión) redonda u ovalada en los pulmones que se observa con una tomografía computarizada o una radiografía del tórax.

#### Causas

Más de la mitad de todos los nódulos pulmonares solitarios no son cancerosos (benignos). Los nódulos benignos tienen muchas causas, entre ellas, cicatrices viejas e infecciones.

Los granulomas infecciosos (reacciones a una infección pasada) causan la mayoría de las lesiones benignas. Las infecciones comunes que incrementan el riesgo de que se forme un nódulo pulmonar solitario abarcan:

- Tuberculosis o haber estado expuesto a ella
- Enfermedades pulmonares causadas por hongos, como: aspergilosis

---

<sup>1</sup>pulmones, nódulos y cáncer de pulmonar es un resumen de las referencias [1,2].

- coccidioidomicosis
- criptococosis
- histoplasmosis

El cáncer de pulmón es la causa más frecuente de nódulos pulmonares cancerosos (malignos).

### Síntomas

El nódulo en sí rara vez causa síntomas.

### Exámenes

Un nódulo pulmonar solitario casi siempre se detecta en una radiografía o en una tomografía computarizada del tórax, que generalmente se hace por otros síntomas o razones. El médico debe decidir si el nódulo en el pulmón probablemente es benigno (no es cáncer). Esto es más probable si:

- El nódulo es pequeño, tiene un borde liso y tiene una apariencia sólida y uniforme en una radiografía o tomografía computarizada.
- Usted es joven y no fuma.

El médico puede entonces optar simplemente por vigilar el nódulo en radiografías.

- Las radiografías o tomografías computarizadas repetitivas son la manera más común de hacerle seguimiento al nódulo. En raras ocasiones, se pueden hacer tomografía por emisión de positrones (TEP).
- Si las radiografías repetitivas muestran que el tamaño del nódulo no ha cambiado en dos años, muy probablemente sea benigno y no se necesite una biopsia .

El médico puede optar por hacer la biopsia del nódulo para descartar cáncer si:

- Usted es fumador.
- Usted tiene otros síntomas de cáncer pulmonar.
- El nódulo ha crecido en tamaño o ha cambiado comparado con radiografías previas.

Una biopsia pulmonar por punción se puede hacer colocando una aguja a través de la pared del tórax o durante una broncoscopia o mediastinoscopia.

También se pueden hacer exámenes de la piel para descartar tuberculosis y otras infecciones.

Los riesgos de una biopsia contra el monitoreo del tamaño del nódulo con radiografías regulares o tomografías computarizadas. El tratamiento se puede basar en los resultados de la biopsia u otros exámenes.

#### Pronostico

El pronóstico generalmente es bueno si el nódulo es benigno. Si el nódulo no aumenta de tamaño en un período de dos años, con frecuencia no es necesario hacer nada más.

- **Cáncer de pulmón**

El cáncer de pulmón es un tumor maligno que generalmente se origina en las células que recubren los bronquios (epitelio bronquial). Se produce principalmente por la irritación e inflamación crónica del epitelio bronquial por agentes externos (carcinógenos) donde destaca el humo del cigarrillo. Producto de esta irritación crónica y de factores genéticos se producen mutaciones que llevan al crecimiento rápido y descontrolado de algunas células, lo que se denomina transformación neoplásica, generándose así un cáncer. Finalmente al continuar creciendo el tumor, algunas de las células pueden desplazarse hacia otros órganos del cuerpo dando origen a metástasis.

#### Factores de riesgo

El tabaco es el principal factor de riesgo. También influyen otros agentes externos como la exposición a carcinógenos como asbestos otros minerales y gases radiactivos, el consumo de marihuana, la contaminación ambiental y la inflamación recurrente por tuberculosis y algunos tipos de neumonía, entre otros. Existe además una importante predisposición genética que hace que algunas personas

sean más susceptibles a desarrollar cáncer al exponerse a estos carcinógenos o aún sin exposición a ellos.

### Riesgo de desarrollar cáncer pulmonar al fumar

El 85% de los pacientes que tienen cáncer de pulmón son fumadores. El riesgo aumenta en pacientes mayores de 50 años y especialmente mayores de 60 años que fuman. Sin embargo, hay personas más jóvenes que desarrollan esta enfermedad. El peligro aumenta mientras mayor sea la cantidad de cigarrillos que se fuma y mayor sea el tiempo que se ha fumado. El riesgo es especialmente importante si se ha fumado más de 10 cigarrillos al día por 10 años. El cigarrillo es el tipo de tabaco con mayor riesgo especialmente los con mayor contenido de nicotina y alquitrán además de los sin filtro que incluyen un riesgo adicional aunque el tabaco de pipa y el de cigarro (habano) también son factores de riesgo para el desarrollo de cáncer pulmonar.

### Estadísticas

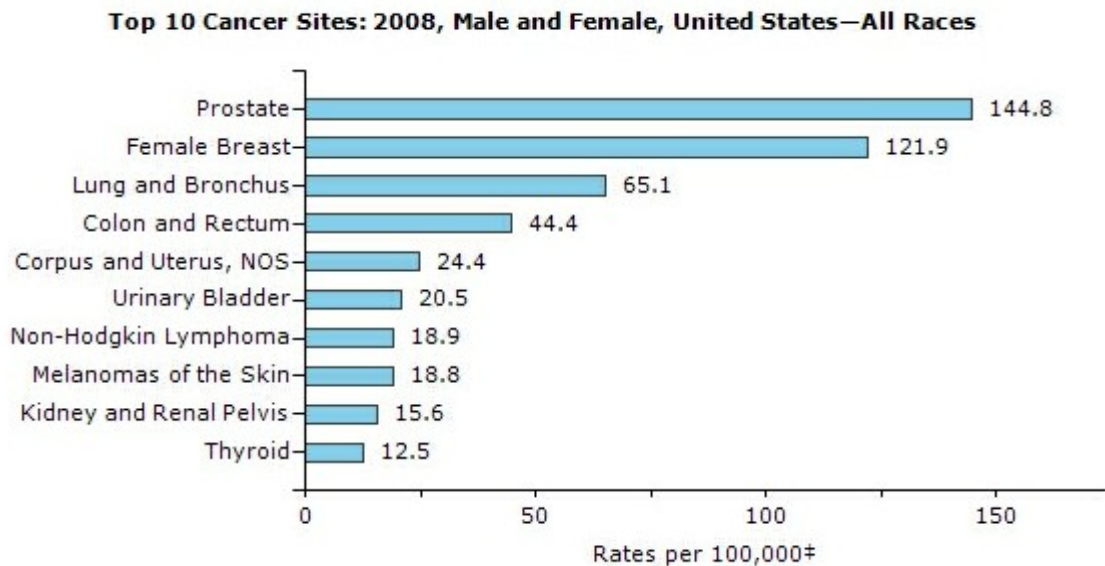


Figura 1. Top 10 de los cáncer mas comunes entre hombres y mujeres x cada 100mil habitantes en estados unidos en 2008 [14].

**Top 10 Cancer Sites: 2008, Male, United States—All Races**

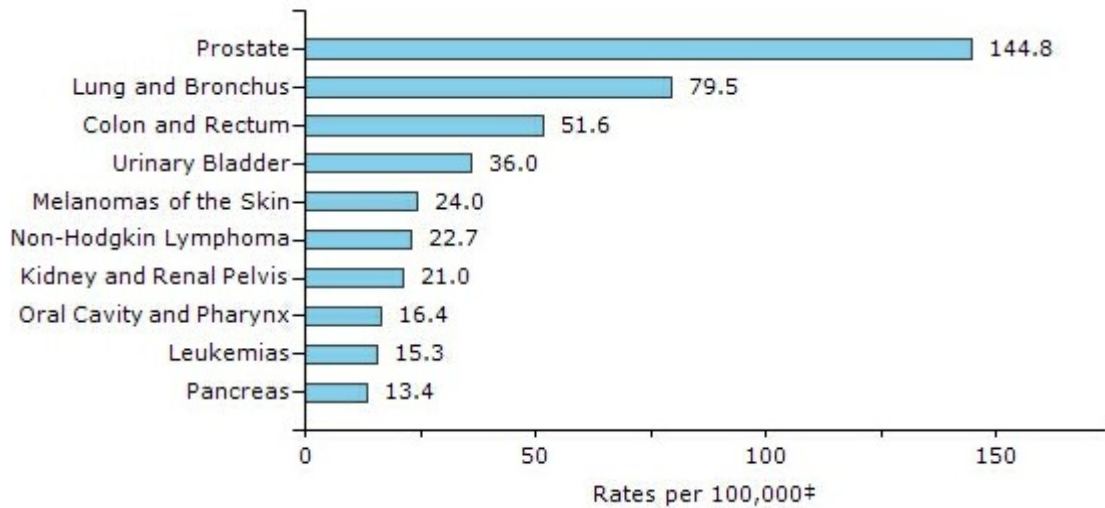


Figura 2. Top 10 de los cáncer mas comunes entre hombres x cada 100mil habitantes en estados unidos en 2008 [14].

**Top 10 Cancer Sites: 2008, Female, United States—All Races**

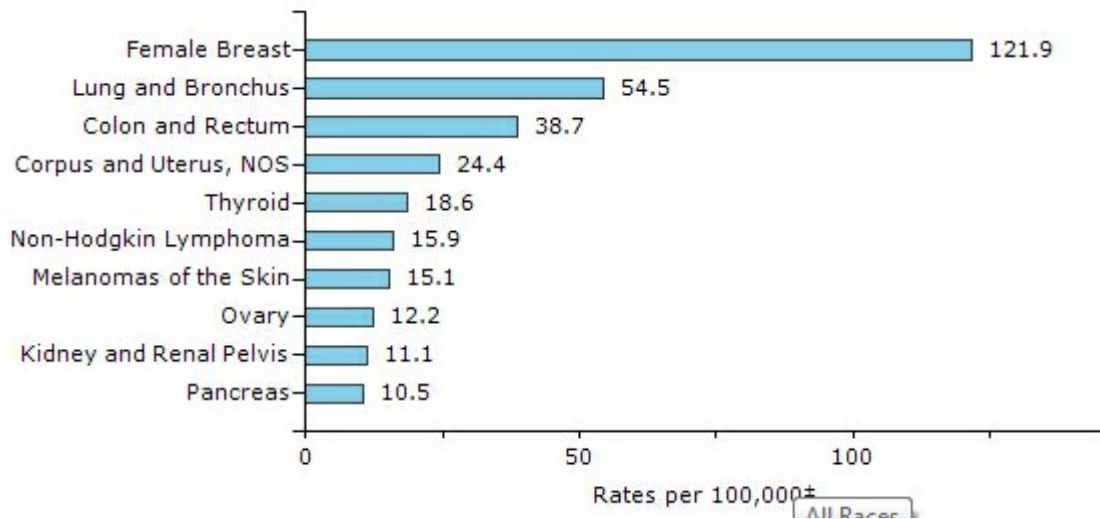


Figura 3. Top 10 de los cáncer mas comunes entre mujeres x cada 100mil habitantes en estados unidos en 2008 [14].



**Top 10 Cancer Sites: 2008, Male and Female, United States—All Races**

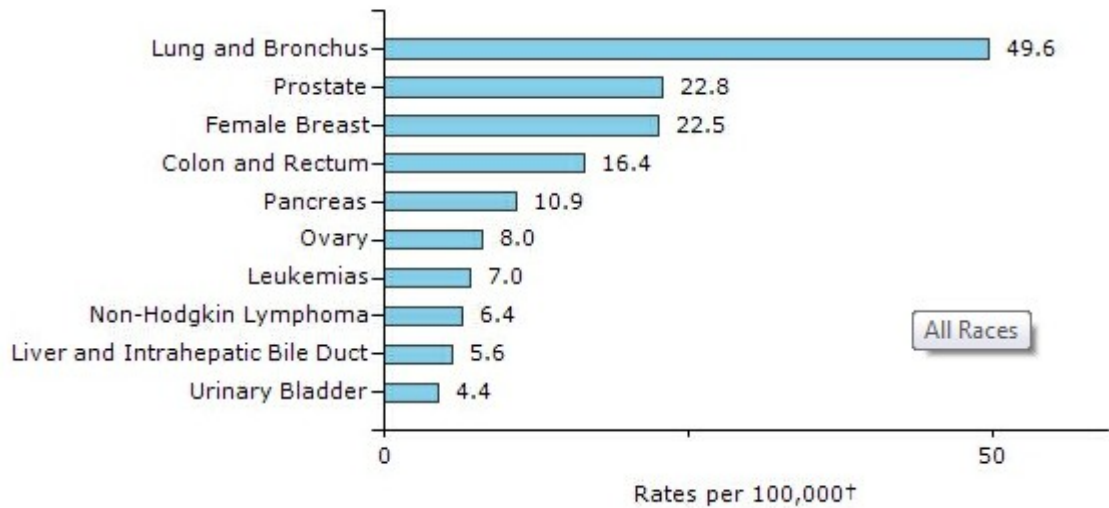


Figura 4. Top 10 de la mortalidad del cáncer entre hombre y mujeres x cada 100mil habitantes en estados unidos en 2008 [14].

## Síntomas

Los síntomas son muy variables y dependen de la localización del tumor y de la presencia de metástasis. En la mayoría de los casos, cuando el cáncer pulmonar da síntomas ya se encuentra en una etapa avanzada.

En general, los principales síntomas son:

- Tos, especialmente asociado a desgarro con sangre
- Disnea (sensación de ahogo y falta de aire)
- Neumonías a repetición
- Dolor torácico
- Baja de peso
- Dolores óseos

## Detección un cáncer pulmonar en etapas iniciales

Un cáncer pulmonar en etapa inicial se puede diagnosticar al detectar un nódulo pulmonar en una Radiografía de Tórax o en una Tomografía Axial Computada (ver Fig.5), siendo esta última mucho más sensible y exacta. Los nódulos pulmonares pueden ser benignos (infecciones, cicatrices o tumores benignos) o representar un cáncer pulmonar en etapa inicial. La importancia de realizar un diagnóstico temprano, es que etapas iniciales la cirugía es el único tratamiento que permite curar el cáncer pulmonar.

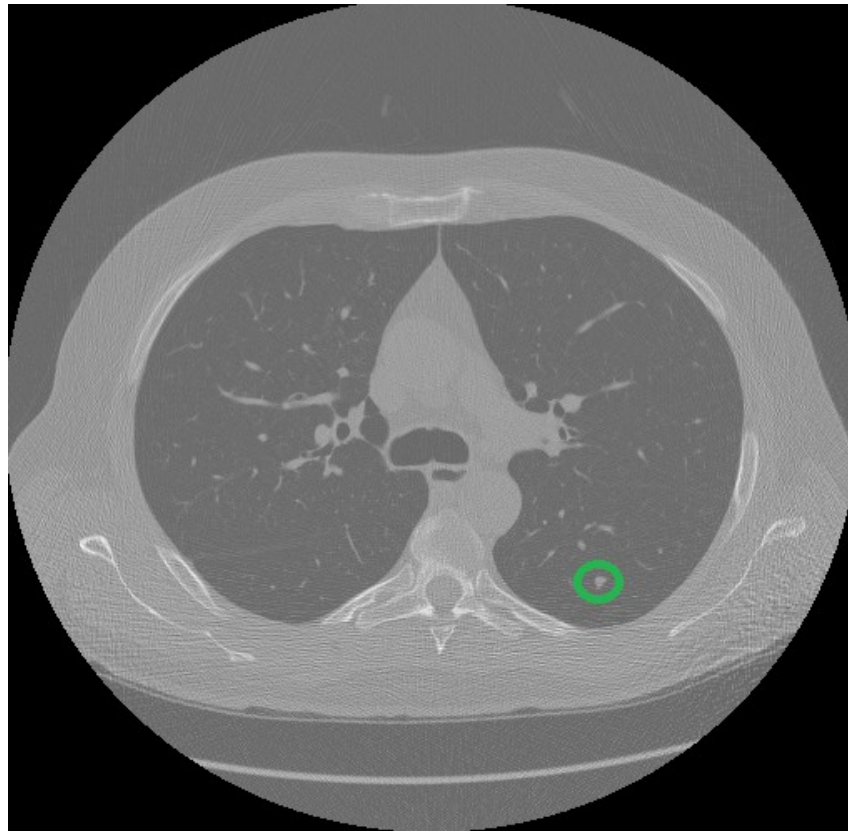


Figura 5. Imagen TAC con presencia de Nódulo pulmonar.

### 3.1.2 Tomografía Axial Computarizada (TAC)<sup>2</sup>

Creada por Godfrey Hounsfield en 1972, La imagen CT permite la adquisición de imágenes bidimensionales basadas en rayos X correspondientes a delgados cortes axiales a lo largo del cuerpo del paciente (ver Fig.6).

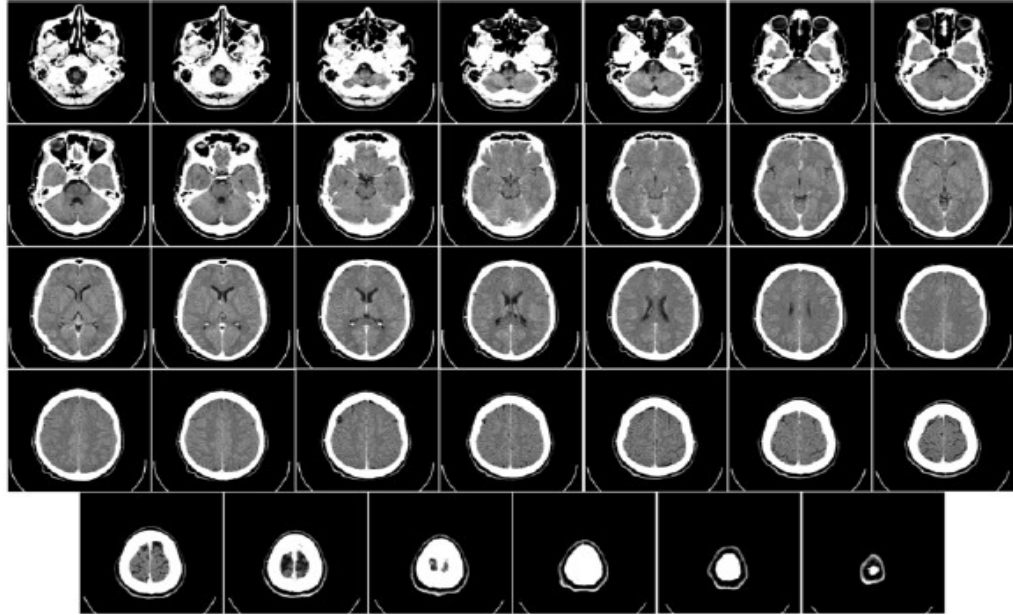


Figura 6. conjunto de imágenes tomográficas en su corte axial.

A partir de múltiples cortes axiales consecutivos se puede realizar una reconstrucción fiel del volumen correspondiente al área radiada. Las imágenes CT son capaces de mostrar de forma contrastada distintos tejidos blandos ya que, la radiación reflejada en la imagen correspondiente a la radiación incidente en un órgano determinado no solapa con la detectada correspondiente a otro órgano distinto. El principio que subyace en la adquisición de las imágenes CT es que la estructura bidimensional interna de un objeto puede ser reconstruida a partir de una serie de proyecciones unidimensionales del objeto adquiridas desde distintos ángulos. Los detectores, situados de frente al emisor, recogen el número total de rayos que han atravesado al paciente, produciendo una proyección unidimensional. Las intensidades de señal recogidas responden a la distribución de los coeficientes de atenuación a lo largo del corte axial correspondiente a la

<sup>2</sup>Tomografía Axial Computarizada (TAC) es un resumen de la referencia [3]

distribución bidimensional de tejidos. Entonces, el emisor y el receptor repetirán esta operación tras girar un cierto ángulo. Esta operación será repetida hasta que se haya conseguido un número suficiente de muestras que permita obtener una imagen de alta resolución espacial.

- **Adquisición de la imagen**

Diversos componentes de un sistema CT, tales como la fuente de rayos X, el colimador, y la rejilla antidispersión son similares a los utilizados en las radiografías proyectivas. A lo largo de los últimos treinta años, los sistemas CT han ido evolucionando, alterando el número de emisores y receptores y la configuración del sistemas.

## **3.2 PROCESAMIENTO DE IMAGENES**

### **3.2.1 Crecimiento de regiones<sup>3</sup>**

Como el nombre indica, el crecimiento de regiones es un procedimiento mediante el cual se agrupan píxeles o subregiones en regiones mayores. El procedimiento mas sencillo se denomina agregación de píxeles, que comienza a partir de un conjunto de píxeles semilla, de forma que a partir de cada semilla se crecen regiones añadiendo píxeles a dicha semilla entre aquellos píxeles vecinos que tienen propiedades similares. El resultado de la segmentación dará lugar como mucho a tantas regiones como semillas haya. Sin embargo, puede darse el caso de que dos de esas semillas correspondan a píxeles de la misma región, en este caso el crecimiento desde una de las semillas absorberá a la otra, que en este caso deberá ser descartada. Fijado un umbral T se va calculando la diferencia en valor absoluto del nivel de gris del píxel en cuestión; con respecto al nivel de gris de la semilla y si no se supera ese umbral T se añade a la región.

Dos problemas fundamentales en el crecimiento de regiones son: por un lado, la selección de las semillas o puntos de partida que representen adecuadamente a las regiones de interés; y por otro, la elección de las propiedades adecuadas que permitan ir agregando píxeles durante el proceso de crecimiento.

---

<sup>3</sup> Crecimiento de Regiones es un resumen de la referencia [7]

En cualquier caso siempre hay que tener en cuenta la conectividad durante el proceso de crecimiento para que el resultado tenga significado dentro de su contexto. Otro problema añadido en el crecimiento de regiones es la regla de parada. Básicamente, el crecimiento termina cuando no existen mas píxeles en el vecindario de la región ya crecida que cumplan el criterio de similitud.

### **3.2.2 Maquinas de vectores soporte (svm)<sup>4</sup>**

Las maquinas de vectores soporte fueron introducidas por Vladimir Vapnik y su grupo de trabajo. La primera mención fue en 1979, pero el principal documento fue publicado en 1995.

El apoyo de las maquinas de vectores soporte (SVM) son un método de aprendizaje relativamente nuevo utilizado para la clasificación binaria. La idea básica es encontrar un hiperplano que separa los datos de d-dimensionales perfectamente en sus dos clases. Sin embargo, ya que los datos de ejemplo a menudo no son linealmente separables, SVM introduce la noción de un kernel o núcleo que se introduce en un espacio de características que pone los datos a un nivel de dimensiones superiores donde los datos son separables. Normalmente, las dimensiones superiores podrían causar problemas de cómputo, y con el sobreajuste. La clave del éxito utilizada en SVM es que el espacio de dimensiones superiores no necesita tratarse directamente (como resulta, sólo la fórmula para el producto punto en ese espacio que se necesita), lo que elimina las preocupaciones mencionadas. Además, la CV-dimensión (una medida de la probabilidad de un sistema para un buen desempeño en los datos que no se ven) de svm se puede calcular de forma explícita, a diferencia de otros tipos de aprendizaje métodos como redes neuronales, para los que no hay ninguna medida. En general, las svm son intuitivas, teóricamente bien fundada. Las svm también se han extendido a resolver tareas de regresión.

La formulación matemática de las Máquinas de Vectores Soporte varía dependiendo de la naturaleza de los datos, es decir, existe una formulación para los casos linealmente separables y, por otro lado, una formulación para casos no linealmente separables. Las máquinas de vectores soporte siempre buscan encontrar un hiperplano óptimo que separe las clases.

---

<sup>4</sup> Maquinas de Vectores Soporte (svm) es un resumen de la referencias [4,5,6]

- **Casos linealmente separables**

En el caso de ser linealmente separable, las MVS conforman hiperplanos que separan los datos de entrada en dos subgrupos que poseen una etiqueta propia. En medio de todos los posibles planos de separación de las dos clases etiquetadas como  $\{-1, +1\}$ , existe sólo un hiperplano de separación óptimo, de forma que la distancia entre el hiperplano óptimo y el valor de entrada más cercano sea máxima (maximización del margen) (ver Fig.7) con la intención de forzar la generalización de la máquina que se esté construyendo como se muestra en la siguiente función.

Función de decisión para la svm

$$(w \cdot x^{+1} + b) = +1 \tag{1}$$

$$(w \cdot x^{-1} + b) = -1$$

$$y_i = \begin{cases} 1 & \text{si } x \in \text{Clase 1} \\ -1 & \text{si } x \in \text{Clase 2} \end{cases} \tag{2}$$

Resolviendo las ecuaciones

$$b = -\frac{1}{2}(w \cdot x^{+1} + w \cdot x^{-1}) \tag{3}$$

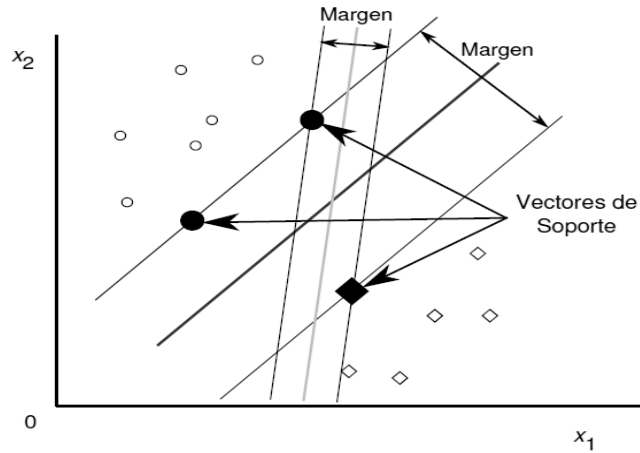


Figura 7. calculo de vectores soporte y maximización del margen

- **Casos linealmente no separables**

Para el caso no lineal existen dos casos que vale la pena mencionar:

El primero de estos se presenta cuando los datos pueden ser separables con margen máximo pero en un espacio de características (el cual es de una mayor dimensionalidad y se obtiene a través de una transformación a las variables del espacio de entrada) mediante el uso de una función kernel como se expresa en la siguientes funciones (ver Fig.8).

Función de transformación

$$x \rightarrow \phi(x) \tag{4}$$

Función de kernel

$$K(x_i, x_j) = \phi^T(x_i) \phi(x_j) \quad (5)$$

Aplicando la función kernel

$$F = (\alpha) \sum_{i=1}^M \alpha_i - \frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (6)$$

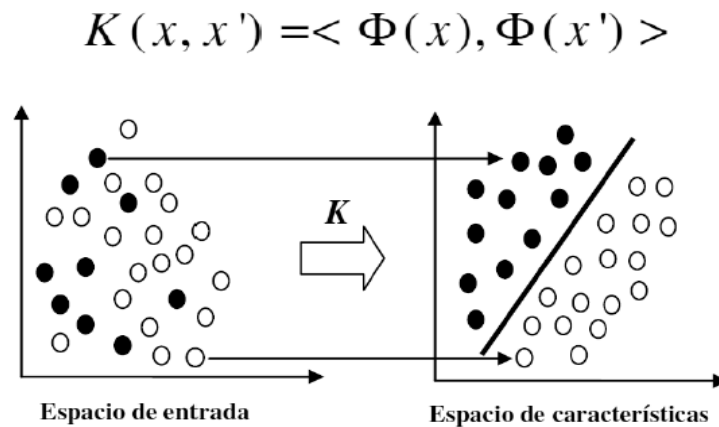


Figura 8. transformación del espacio de entrada a un espacio de características por medio de un kernel.



Los tipos de kernel mas usados son:

Lineales

$$K(x_i, x_j) = x_i^T x_j \quad (7)$$

Polinomiales

$$K(x_i, x_j) = (x_i^T x_j + 1)^d \quad (8)$$

Radial basis function (RBF)

$$K(x_i, x_j) = e^{-\gamma \|x_i - x_j\|^2} \quad (9)$$

El segundo caso especial de las svm denominado “Soft Margin” o margen blando, es utilizado cuando no es posible encontrar una transformación de los datos que permita separarlos linealmente, bien sea en el espacio de entrada o en el espacio de características como se expresa en la siguiente imagen(ver Fig.9). Esto incluye la constante C que determina la holgadura del margen y este debe ser estimado apriori, por último tener en cuenta que la elección de la C y el tipo de kernel influyen directamente en el desempeño de las svm (ver Fig.9).

No es difícil ampliar las ideas generales del caso separable al caso no separable introduciendo una variable  $\xi_j$  de holgura en las restricciones y plantear un nuevo conjunto de restricciones:

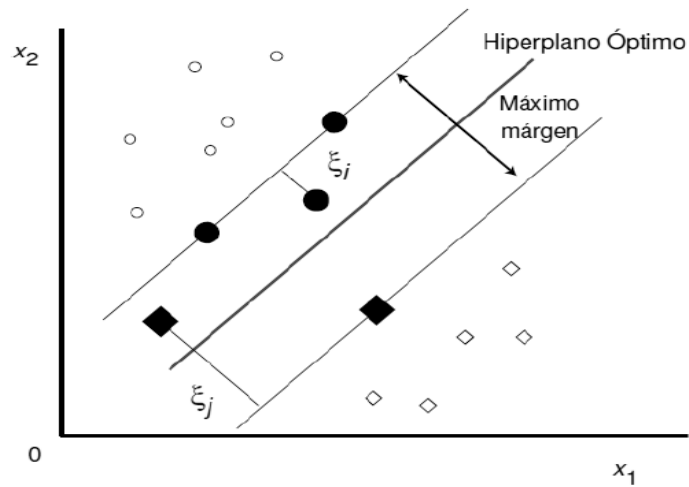


Figura 9. maximización del margen con soft margin

$$\begin{aligned}
 x_i \cdot w + b &\geq +1 - \xi_i \quad \text{para } y_i = +1 \\
 x_i \cdot w + b &\geq -1 + \xi_i \quad \text{para } y_i = -1 \\
 \xi_i &\geq 0 \quad \forall i = 1, \dots, n
 \end{aligned}
 \tag{10}$$

Ya que en el caso de margen suave, necesariamente se cometen errores, parece natural asignar a la función objetivo un coste extra que en cierto modo penalice los errores (función de pérdida). Por todo ello, una opción lógica sera plantear la función de minimización.

$$\frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i^k
 \tag{11}$$

### 3.2.3 RBF gaussiana<sup>5</sup>

Las funciones de base radial han recibido mucha atención, por lo general con un Gaussiana de la forma,

<sup>5</sup> RBF Gaussiana es un resumen de la referencia [8]

$$K(x, y) = \exp\left(-\frac{(x-y)^2}{2\sigma^2}\right) \quad (12)$$

Las técnicas clásicas que utilizan funciones de base radial emplean algún método de determinar un subconjunto de centros. Típicamente un método de agregación que se emplea antes de seleccionar un subconjunto de centros. Una característica atractiva de las svm es que esta selección es implícita, con cada uno de los vectores de soporte que contribuyen a un local de la función de Gauss, centrada en ese punto de datos. Por consideraciones adicionales que es posible seleccionar la base global ancho de la función,  $s(\sigma)$ , utilizando el principio svm.

### 3.2.4 Coeficiente kappa<sup>6</sup>

En una medidas de concordancia en la cual se puede medir entre dos observadores o mas, es muy usado en ensayos clínicos y se calcula de la siguiente manera:

		Observador B		
		Positivo	Negativo	
Observador A	Positivo	a	c	f1
	Negativo	b	d	f2
		c1	c2	n

Tabla 1. tabla para el calculo del coeficiente kappa.

$$pe = \frac{f1 * c1 + f2 * c2}{n^2} \quad po = \frac{a+b}{n} \quad (13)(14)(15)$$

$$kappa = \frac{po - pe}{1 - pe}$$

<sup>6</sup> Coeficiente kappa es un resumen de la referencia [10]

El grado de concordancia según el resultado de kappa se mide según la imagen (ver tabla 2) y se determina el grado de acuerdo.

<b>kappa</b>	<b>grado de acuerdo</b>
< 0	sin acuerdo
0 - 0,2	insignificante
0,2 - 0,4	bajo
0,4 - 0,6	moderado
0,6 - 0,8	bueno
0,8 - 1	muy bueno

Tabla 2. medición de coeficiente kappa y grado de acuerdo

## 4. PROPUESTA DE DETECCIÓN

Para el desarrollo del proyecto se cuenta con 50 estudios TAC<sub>[13]</sub> con espesor de 1.25mm, en los cuales hay presencia de nódulos pulmonares; teniendo esta base de imágenes se hace una separación de estudios para entrenamiento y otra cantidad de estudios para hacer pruebas o evaluaciones sobre el entrenamiento hecho a las svm y así probar el rendimiento.

El proyecto consta de varias etapas las cuales son: definición de la región de interés, segmentación de la región de interés, extracción de características, aplicación del svm, pruebas y evaluación; todas las etapas del proyecto fueron desarrolladas en el software matlab®.

### 4.1 REGION DE INTERÉS

La región de interés se define por un experto en el diagnóstico a partir de la imagen y donde se pueda tener sospechas de un nódulo pulmonar (ver Fig.10) a partir de esto, se hace una extracción de la posición y el tamaño de la región, y se introduce al algoritmo para el desarrollo de las siguientes fases.

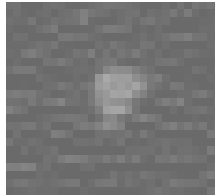


Figura 10. imagen de un nódulo enmarcado en una región de interés.

### 4.2 CRECIMIENTO DE REGIONES

Para hacer la segmentación del posible nódulo en la región de interés se usa el crecimiento de regiones como herramienta de segmentación del nódulo de la zona previamente seleccionada en las imágenes, este proceso no lo hace exento de los posibles falsos nódulos, pero nos enmarca la región de análisis para las

características a extraer del posible nódulo.

La segmentación de imágenes comienza con la selección de las semillas mas adecuada para el algoritmo de crecimiento de regiones, que en este caso se hace en el centro de la imagen porque antes se ha definido una región de interés y se presupone que el objetivo va a quedar en el centro o una porción de el, y así determinar un rango de crecimiento moderado para no provocar el desbordamiento de la región objetivo basados en la intensidad media de los cambios en borde de los nódulos en general.

Ya seleccionada las semillas se procede a implementar el algoritmo de crecimiento de regiones[9] con la función `regiongrow` de matlab®, en el cual se introducen una matriz que corresponde a la imagen a tratar, una matriz del mismo tamaño en la que se incluyen las semilla de crecimiento que tiene las semillas marcadas con 1 y lo que no pertenece a semilla con 0, por último un umbral de entrada para marcar lo limites de pertenencia de las semillas. La salida de algoritmo nos entrega 4 salidas que son la imagen con las regiones detectadas, numero de regiones detectadas, imagen con semillas utilizadas y por último una imagen que cumple con la condición de limite (ver Fig.11).

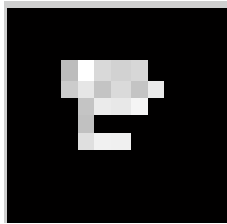


Figura 11. nódulo segmentado por crecimiento de regiones.

### 4.3 EXTRACCIÓN DE CARACTERÍSTICAS

Las características a extraer de las zonas previamente segmentadas son: desviación estándar, área, perímetro, promedio, circularidad, potencia y distancia de la pared del pulmón[12].

#### **4.3.1 Desviación estándar**

Para el desarrollo de la desviación estándar se hace una individualización de las zonas segmentadas y a cada una de ellas se le aplica la función `std2` de matlab® la cual calcula sobre la imagen y se obtienen el valor.

#### **4.3.2 Área**

Para el cálculo del área se hace la binarización de cada una de las regiones y se calcula el área con la función `bwarea` de matlab® que nos hace un conteo de los píxeles que pertenecen a la región binarizada en este caso el posible nódulo.

#### **4.3.3 Perímetro**

El calculo del perímetro se trabaja con las regiones binarizadas y la función `perim` de matlab® que nos da como resultado todo los píxeles que bordean la región segmentada y para obtener el resultado se hace la sumatoria de los píxeles que pertenecen al perímetro calculado.

#### **4.3.4 Promedio de la intensidad**

Con las regiones segmentada y usando la función `mean2` de matlab® para la media de las matrices se hace le promedio de intensidades de las zonas o regiones en las cuales se este trabajando.

#### **4.3.5 Circularidad**

Basados en los resultados de área y perímetro se aplica la característica de circularidad que se expresa con al formula  $(2 \cdot \text{área}) / \text{perímetro}$  y como resultado se obtiene circularidad.

#### **4.3.6 Potencia**

Para el calculo de la potencia sobre la imagen teniendo las imágenes segmentada e individualizadas para cada una de ella se le aplica la función `fft2` de matlab® que es la transformada de fourier para las matrices, a partir del este resultado se le

aplica la función `norm` de matlab® que nos arroja el resultado.

#### 4.3.7 Distancia de la pared del pulmón

Con la herramienta `imtool` de matlab® teniendo identificado el posible nódulo para introducirlo en la región de interés se calcula manualmente la distancia con la regla que ofrece esta aplicación, nos arroja la distancia en píxeles de cualquier medición sobre la imagen.(ver Fig.12).

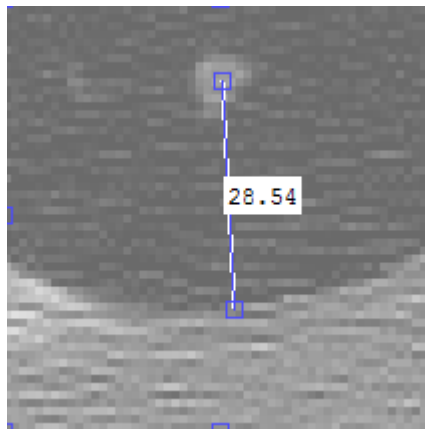


Figura 12. medición de la distancia del nódulo a la pared del pulmón.

#### 4.4 SVMs

Para este módulo se utiliza la función `svmtrain` de matlab® para desarrollar el entrenamiento de la svm y `svmclassify` de matlab® para hacer uso del entrenamiento hecho con `svmtrain`, para el proyecto se definió el kernel que es RBF gaussiano con variación del parámetro sigma entre 0,1 y 1 con variaciones de 0,1.

En fase de entrenamiento del svm se crea una matriz que esta compuesta con la suma vectores resultado de la extracción de características y un vector o etiqueta que define a que grupo pertenecen cada uno de los vectores, si es un posible nódulo o no. para la fase de prueba se usa un vector o una matriz de la misma



forma del entrenamiento pero sin etiqueta para que así la svm realice una etiqueta sobre le vector introducido.

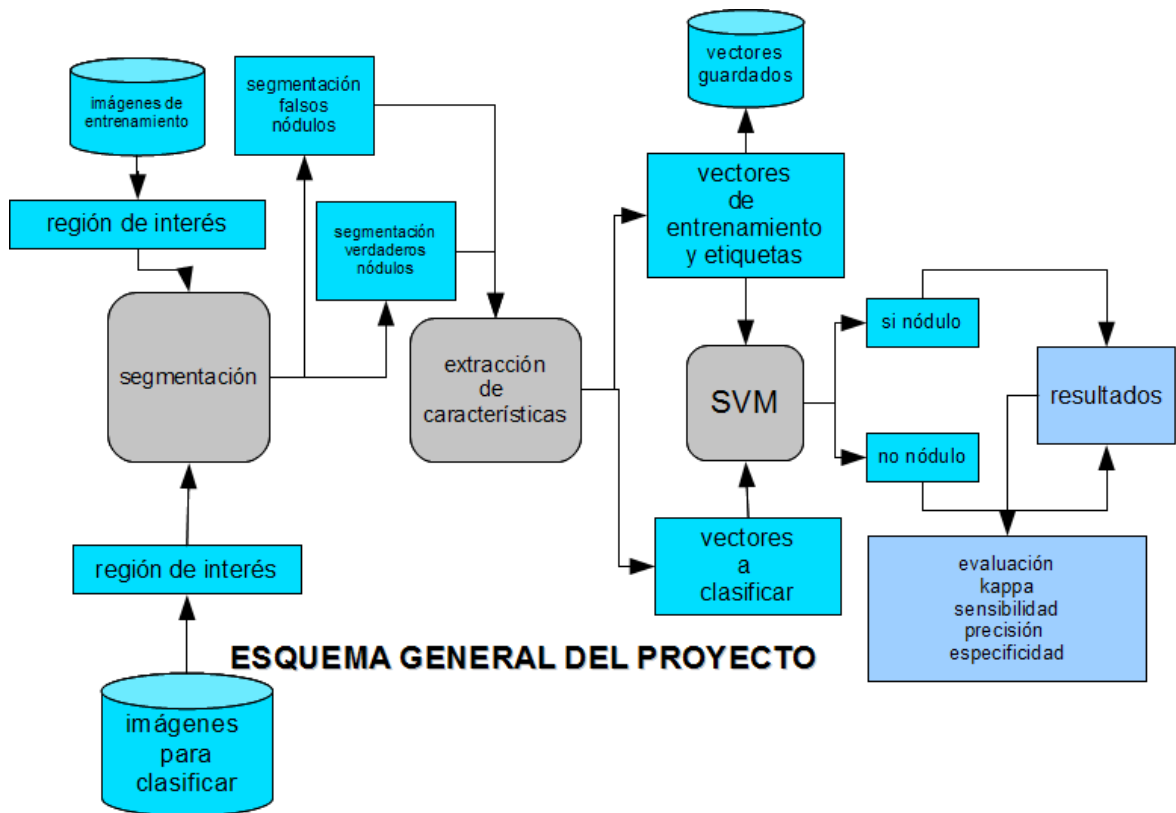
Para este proyecto en particular se hacen dos entrenamientos por separado uno para nódulos pequeños (<5mm), otro para nódulos grandes (>5mm) y por último para el conjunto de todos los nódulos con el objetivo de tener mejores resultados en las evaluaciones.

#### **4.5 INTEGRACION DEL ALGORITMO SVM CON EL VECTOR DE CARACTERÍSTICAS**

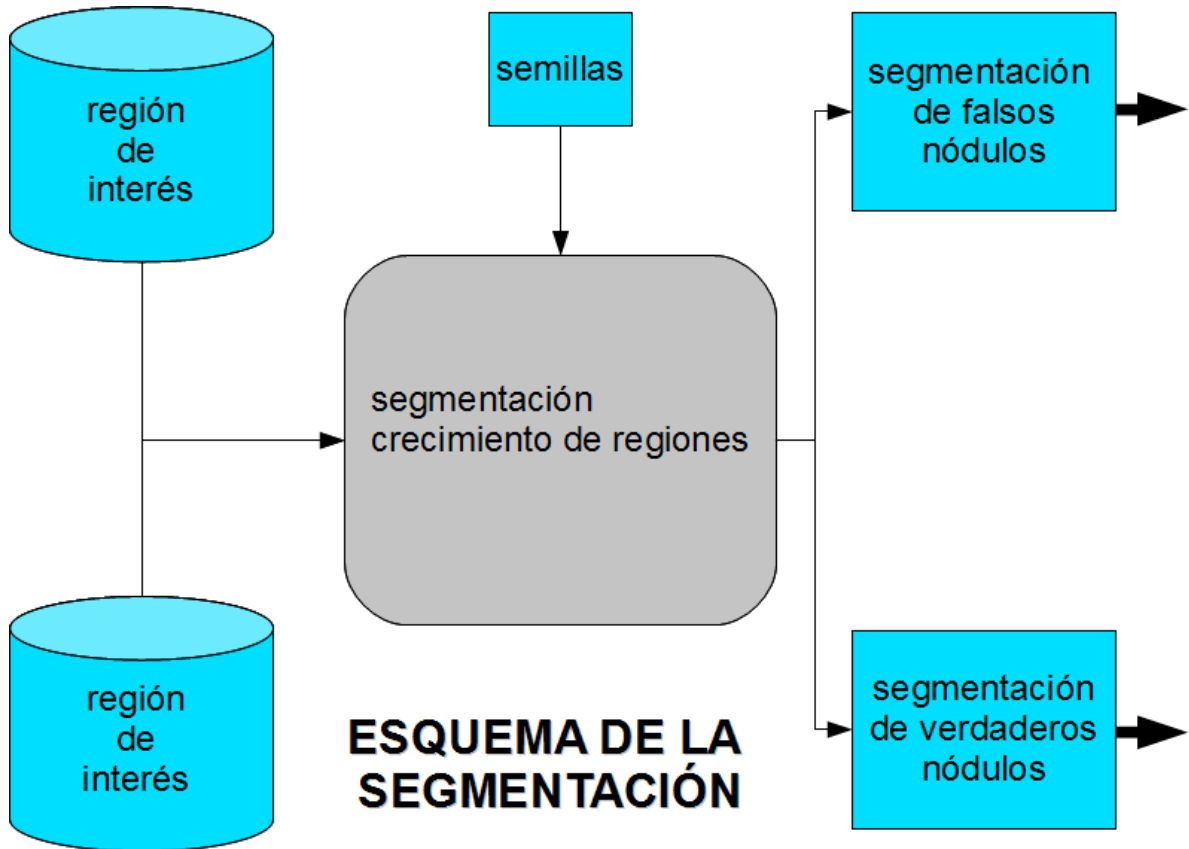
Cuando se está desarrollando la fase de extracción de características a cada una de ellas se les ha calculado un valor por medio de las funciones ya descritas, cada uno de esos valores pertenecen a una imagen y a cada imagen con el conjunto de características extraídas se le construye un vector que identifica esa imagen; con el conjunto de vectores de todas las imágenes tomadas tanto de nódulos verdaderos como de nódulos falsos se construyen matrices tanto de entrenamiento como de pruebas. Con la construcción de estas matrices se hace las aplicación del algoritmo de svm.

## 4.6 DIAGRAMA DE BLOQUES

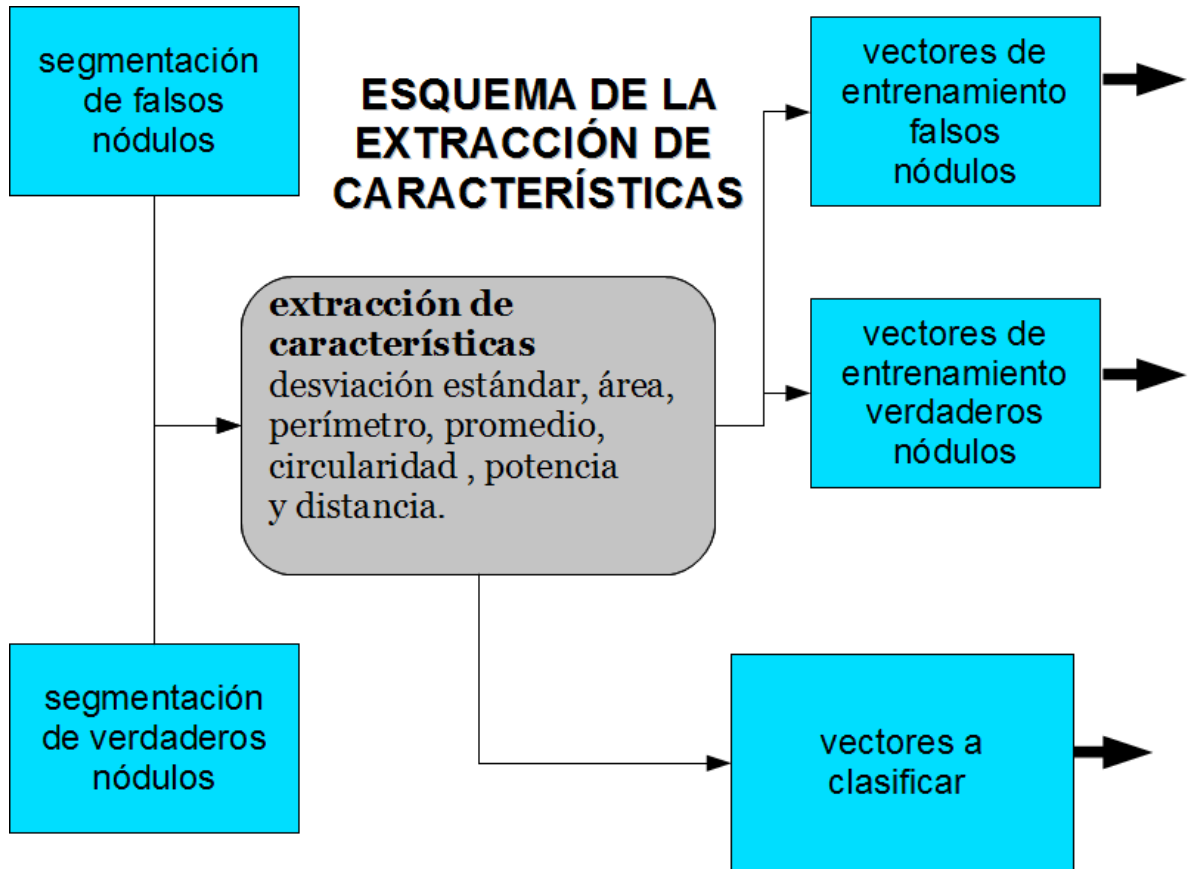
### 4.6.1 Esquema general del proyecto



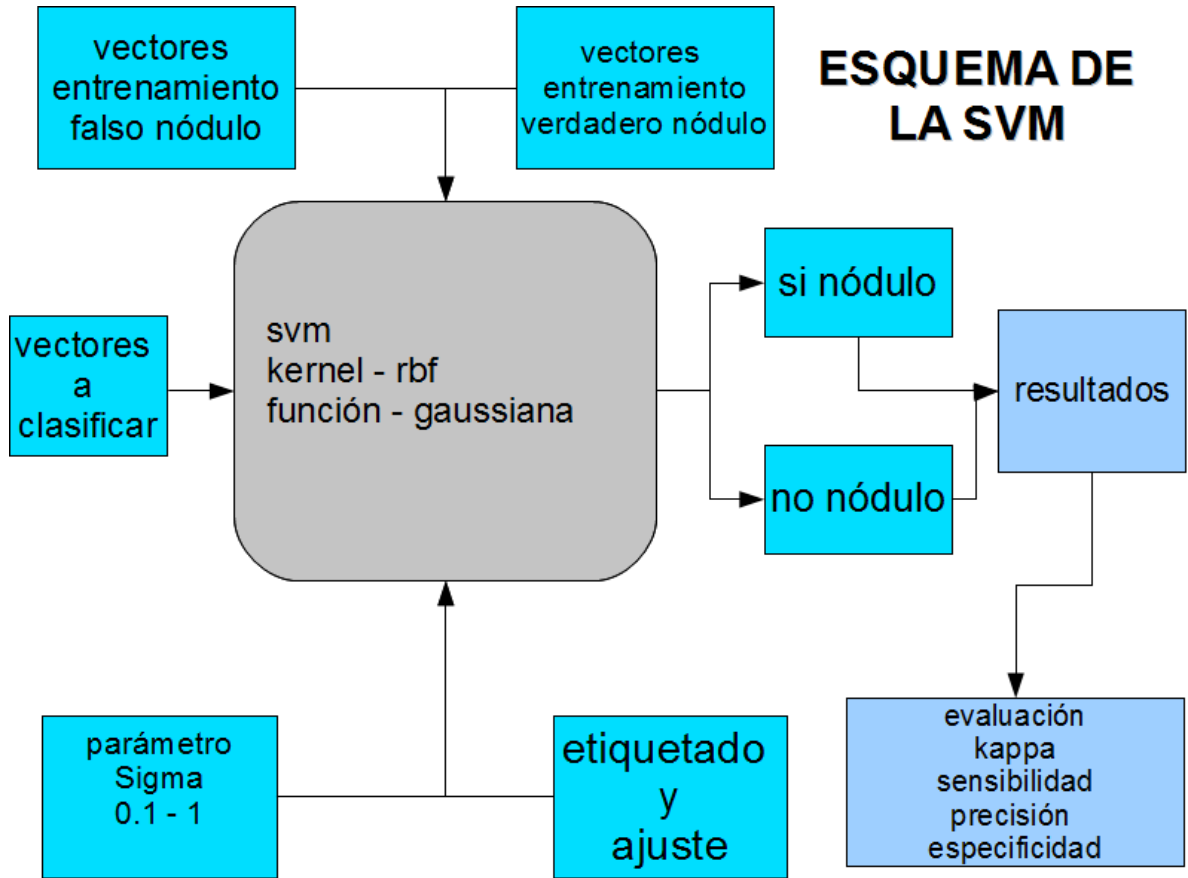
#### 4.6.2 Esquema de la segmentación



#### 4.6.3 Esquema de la extracción de características



#### 4.6.4 Esquema de la svm



## 5. RESULTADOS

### 5.1 DESCRIPCIÓN DEL PROTOCOLO DE EXPERIMENTACIÓN

Teniendo la imagen TAC se hace la identificación del nódulo basado en las anotaciones del experto en el diagnóstico, se hace luego recorte de la región de interés que cubre toda el área del nódulo identificado; se aplica el algoritmo de crecimiento de regiones que nos enmarca adecuadamente el nódulo; a partir de esto se puede hacer la extracción de características.

Para el entrenamiento de la svm se hace todo el experimento dividiendo en 3 grupos de análisis que son: para nódulos grandes (>5mm), para nódulos pequeños (<5mm) y para el conjunto de los dos anteriores.

Para cada conjunto se tiene una matriz previamente construida que tiene adjunta un vector de etiquetas el cual identifica cada vector que compone la matriz como un nódulo verdadero o uno falso con esta información se procede al entrenamiento de la svm.

Para el desarrollo del experimento se estructuraron matrices con vectores de características con verdaderos nódulos y falsos nódulos que se introducen a la svm obteniendo como respuesta un vector con las etiquetas de identificación de verdadero o falso nódulo basadas en el entrenamiento previo.

Junto con a la matriz de pruebas a cada vector se le asignó una etiqueta con la observación del experto en el diagnóstico de los nódulos pulmonares para tener dos puntos de observación, tanto de los expertos como la del algoritmo; con las etiquetas resultantes de la svm se prosigue con la evaluación, conjunto a esto se hizo la variación en la svm del parámetro sigma puesto que la svm funciona con kernel RBF Gaussiano

Para la medición rendimiento y repuesta de la svm se uso el coeficiente de kappa como medición de la concordancia entre el observador de los nódulos contra la respuesta del entrenamiento hecho a la svm y adicionalmente se tienen otras 3

mediciones que son la sensibilidad, especificidad y precisión<sup>[11]</sup> teniendo así una visión clara del rendimiento de proyecto.

Se hicieron 3 evaluaciones para nódulos grandes, nódulos pequeños y la mezcla o totalidad de los nódulos; para cada una de las evaluaciones tres pruebas del svm mezclando los nódulos extraídos para la prueba, como para entrenamiento aleatoriamente, para reentrenar la svm y realizar el promedio de los resultados.

$$sensibilidad = \frac{vp}{vp + fn}$$

$$especificidad = \frac{vp + vn}{vp + vn + fp + fn} \quad (16)(17)(18)$$

$$precisión = \frac{vn}{vn + fp}$$

## 5.2 PRESENTACIÓN DE RESULTADOS

### 5.2.1 Nódulos grandes

para esta evaluación se tiene una matriz de 68 a 72 nódulos entre positivos y falsos para la cual se midieron los falsos positivos (fp), falsos negativos (fn), verdaderos negativos (vn), verdaderos positivos (vp) (ver tabla 3).

Tabla3. Resultados pruebas aleatorias para nódulos grandes										
Sigma										
sigma	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
verdaderos negativos	18	18	16	11	11	12	13	13	13	13
verdaderos positivos	0	2	18	49	48	48	48	48	48	49
falsos negativos	0	0	2	7	7	6	5	5	5	5
falsos positivos	50	48	32	1	2	2	2	2	2	1
Sigma										
sigma	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
verdaderos negativos	1	3	8	12	12	12	13	14	15	15
verdaderos positivos	50	50	48	46	45	45	45	44	43	43
falsos negativos	21	19	14	10	10	10	9	8	7	7
falsos positivos	0	0	2	4	5	5	5	6	7	7
Sigma										
sigma	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
verdaderos negativos	21	20	16	17	17	17	17	18	18	18
verdaderos positivos	2	10	43	45	43	43	43	43	43	42
falsos negativos	0	1	5	4	4	4	4	3	3	3
falsos positivos	48	40	7	5	7	7	7	7	7	8

Tabla 3. evaluación de nódulos resultados del svm para nódulos grandes.

La evaluación del del coeficiente kappa de concordancia y la medición de la sensibilidad, especificidad y precisión (ver Fig.13).

### Evaluacion Kappa Nódulos Grandes

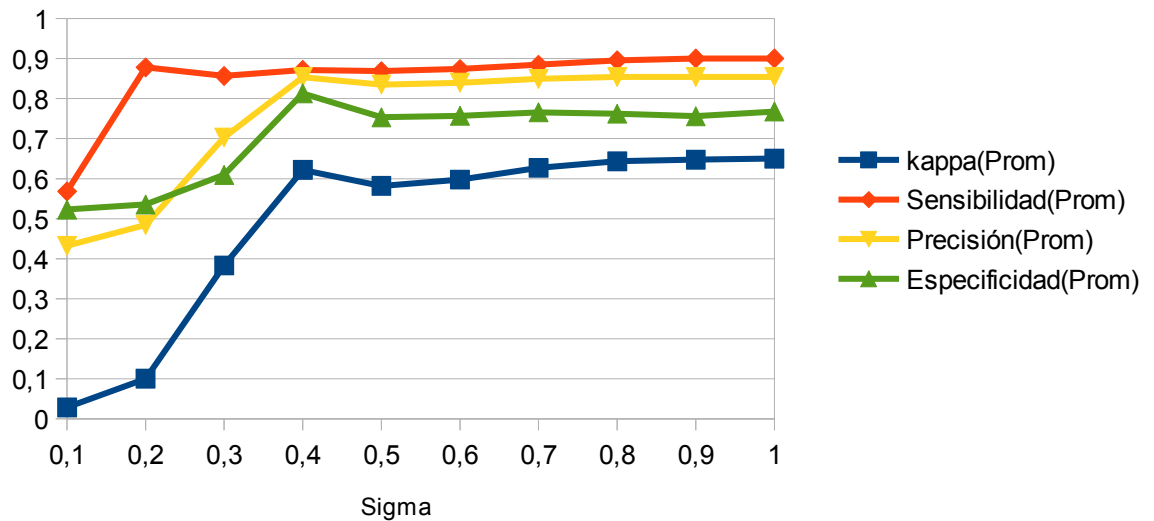




Figura 13. evaluación kappa resultados del svm para nódulos grandes.

### 5.2.2 Nódulos pequeños

para esta evaluación se tiene una matriz de 67 a 68 nódulos entre positivos y falsos para la cual se midieron los falsos positivos (fp), falsos negativos (fn), verdaderos negativos (vn), verdaderos positivos (vp) (ver tabla 4).

Tabla4. Resultados pruebas aleatorias para nódulos pequeños										
Sigma										
sigma	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
verdaderos negativos	17	17	7	12	12	13	13	14	14	15
verdaderos positivos	2	17	46	45	44	43	43	43	43	43
falsos negativos	0	0	10	5	5	4	4	3	3	2
falsos positivos	48	33	4	5	6	7	7	7	7	7
Sigma										
sigma	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
verdaderos negativos	4	11	13	13	13	13	13	14	14	15
verdaderos positivos	50	44	43	45	45	46	46	46	46	46
falsos negativos	14	7	5	5	5	5	5	4	4	3
falsos positivos	0	6	7	5	5	4	4	4	4	4
Sigma										
sigma	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
verdaderos negativos	18	7	11	15	14	14	15	15	15	15
verdaderos positivos	5	49	49	49	49	48	47	47	47	47
falsos negativos	0	11	7	3	4	4	3	3	3	3
falsos positivos	45	1	1	1	1	2	3	3	3	3

Tabla 4. evaluación de nódulos resultados del svm para nódulos pequeños.

La evaluación del coeficiente kappa de concordancia y la medición de la sensibilidad, especificidad y precisión (ver Fig.14).

### Evaluacion Kappa Nódulos Pequeños

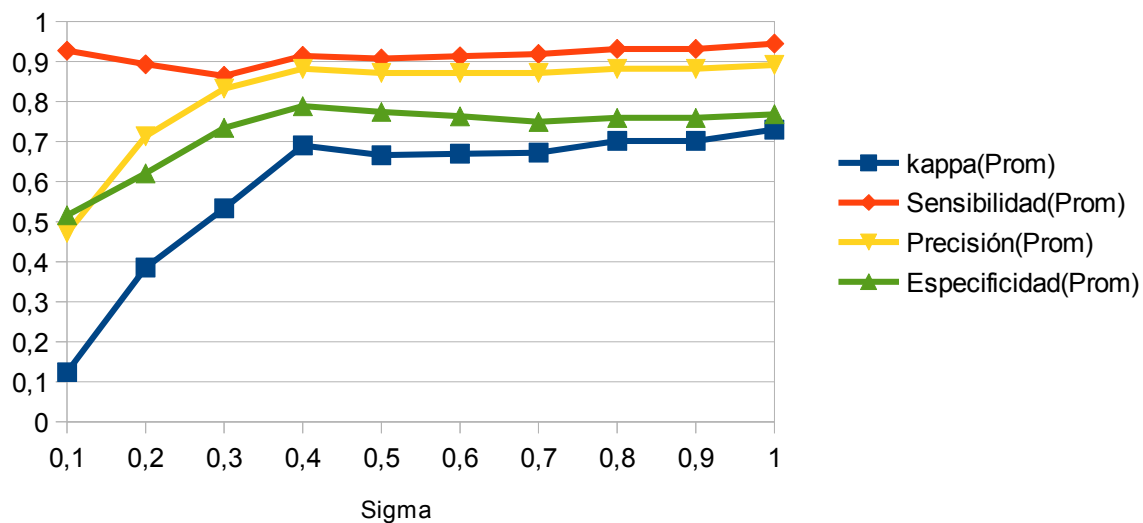


Figura 14. evaluación kappa resultados del svm para nódulos pequeños.

### 5.2.3 Nódulos totales

para esta evaluación se tiene una matriz de 135 a 139 nódulos entre positivos y falsos para la cual se midieron los falsos positivos (fp), falsos negativos (fn), verdaderos negativos (vn), verdaderos positivos (vp) (ver tabla 5).

Tabla5. Resultados pruebas aleatorias para nódulos totales										
Sigma										
sigma	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
verdaderos negativos	35	35	21	24	24	27	25	27	28	28
verdaderos positivos	6	33	95	93	93	93	94	94	93	92
falsos negativos	0	0	14	11	11	8	10	8	7	7
falsos positivos	94	67	5	7	7	7	6	6	7	8
Sigma										
sigma	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
verdaderos negativos	36	35	23	23	25	25	26	27	27	27
verdaderos positivos	7	41	88	90	90	91	90	89	90	91
falsos negativos	0	1	13	13	11	11	10	9	9	9
falsos positivos	93	59	12	10	10	9	10	11	10	9
Sigma										
sigma	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
verdaderos negativos	39	35	31	32	32	33	32	32	32	32
verdaderos positivos	11	47	93	90	88	88	88	87	87	87
falsos negativos	0	4	8	7	7	6	7	7	7	7
falsos positivos	89	53	7	10	12	12	12	13	13	13

Tabla 5. evaluación de nódulos resultados del svm para nódulos totales.

La evaluación del coeficiente kappa de concordancia y la medición de la sensibilidad, especificidad y precisión (ver Fig.15).

### Evaluación Kappa Nódulos Totales

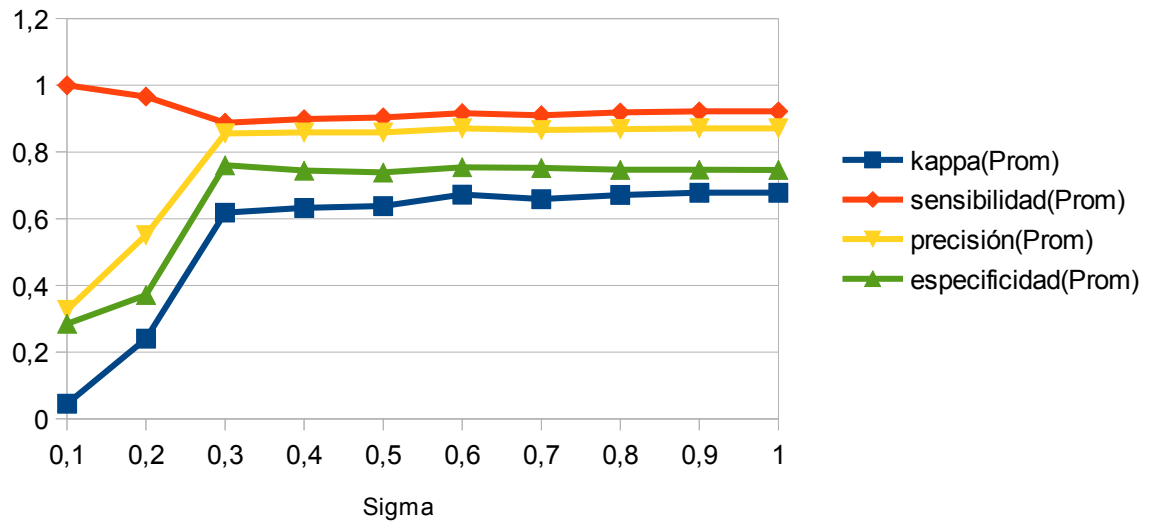


Figura 15. evaluación kappa resultados del svm para nódulos totales.

#### 5.2.4 Integración mejores resultados

Los mejores resultados de la evaluación kappa con respecto a la variable sigma están consignados en la tabla y gráfica (ver tabla 6 y Fig.16) y se toman como resultados del mejor rendimiento del proyecto con respecto al coeficiente kappa.

### Mejores resultados evaluacion kappa promedio

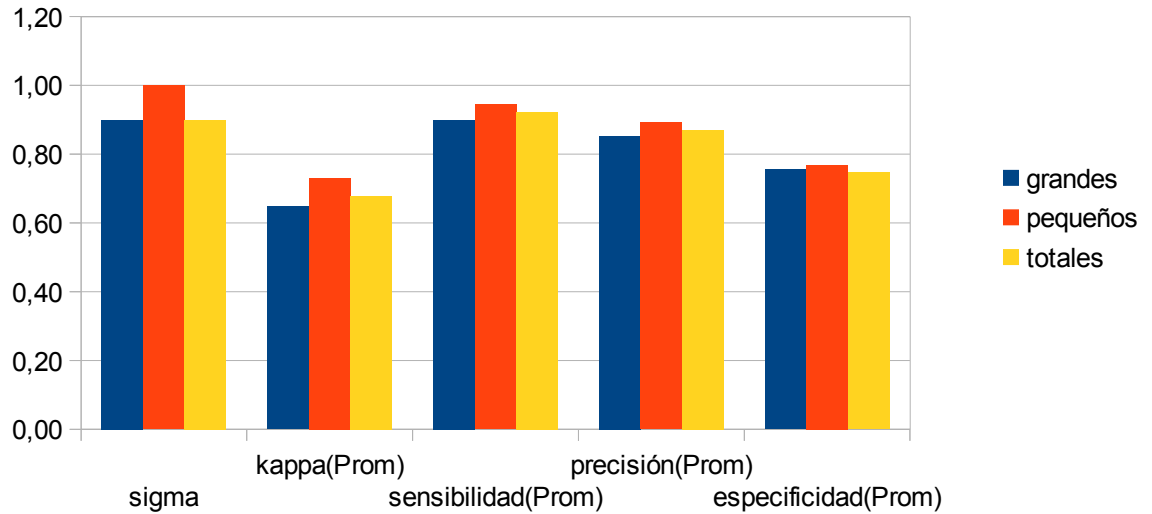


Figura 16. mejores resultados de la evaluación kappa para nódulos grandes, pequeños y totales con referencia sigma.

nódulos	sigma	kappa(Prom)	sensibilidad(Prom)	precisión(Prom)	especificidad(Prom)
grandes	0,90	0,65	0,90	0,85	0,76
pequeños	1,00	0,73	0,94	0,89	0,77
totales	0,90	0,68	0,92	0,87	0,75

Tabla 6. mejores resultados de la evaluación kappa para nódulos grandes, pequeños y totales con referencia sigma.

## **5.3 ANÁLISIS DE RESULTADOS**

en los 3 conjuntos de análisis que se tomaron se obtuvieron resultados muy buenos basados en la escala del coeficiente kappa.

### **5.3.1 Nódulos grandes**

El mejor rendimiento o mejor respuesta para el entrenamiento hecho con nódulos de tamaño grande (>5mm) se obtuvo un valor de kappa de 0,65 para un sigma de 0,9, que nos ubica en el valor de concordancia bueno basado en la escala. Los valores de sensibilidad precisión y especificidad con valores respectivamente de 0.90 , 0.85, 0.76 nos dan a entender que en conjunto la presencia de falsos negativos como de falsos positivos es aceptable para este conjunto de análisis.

### **5.3.2 Nódulos pequeños**

El mejor rendimiento o mejor respuesta para el entrenamiento hecho con nódulos de tamaño pequeño (<5mm) se obtuvo como mayor valor el 0.73 para un sigma de 1 el cual nos ubica en el valor de concordancia bueno basado en la escala. El valor sensibilidad que es 0.94 nos refleja que se tiene una presencia de falsos negativos menor al 6% en cambio del valor de precisión que resulta un valor de 0.89 que nos arroja un 11% de falsos positivos.

### **5.3.1 Nódulos totales**

El mejor rendimiento o mejor respuesta para el entrenamiento hecho con nódulos en conjunto se obtuvo como mayor valor el 0,68 para un sigma 0.9 el cual nos ubica en el valor de concordancia bueno basado en la escala. Los valores de sensibilidad precisión con valores respectivamente de 0.92 , 0.87 nos dan a entender que en conjunto la presencia de falsos negativos como de falsos

positivos es aceptable para este conjunto de análisis y es menor al 13 % aunque la especificidad nos arroje un valor de tan solo 0.75.

## **6. CONCLUSIONES**

### **6.1 VENTAJAS DE LA PROPUESTA DE DETECCIÓN CUANTIFICADAS EN TÉRMINOS DE LOS RESULTADOS**

En este proyecto se propone el uso de las svm para la detección de nódulos pulmonares, del desarrollo del proyecto se destaca el tener un muy buen rendimiento en la detección con un listado de características de solo 7 comparado con otros estudios similares<sup>[12]</sup> que usan un listado grande de características que según los resultados no aportan mejoras visibles a la temática o al problema propuesto.

El procedimiento de extracción de características se hace de una manera automática a partir de las región de interés, a excepción de la característica de distancia de la pared del pulmón al nódulo; se puede afirmar que esta etapa es uno de los fuertes para la buena respuesta que tuvo a la evaluación por lo tanto en los resultados.

Se destaca el buen resultado del coeficiente kappa que fue de 0.73 en el grupo de los nódulos pequeños, que fue el mas alto dentro de la evaluación promediada y nos pone en una buena posición con respecto a la clasificación kappa .

Basados en los resultados obtenidos en la evaluación del proyecto se observa que se tiene una buena confiabilidad y se podría usar como ayuda diagnostica en la detección de nódulos pulmonares en un futuro.

### **6.2 LIMITACIONES ENCONTRADAS A PARTIR DE LOS RESULTADOS**

El rendimiento del conjunto de los nódulos grandes tuvo un bajo rendimiento con respecto al grupo de los nódulos pequeños teniendo en cuenta que tuvo una respuesta buena en el coeficiente kappa.



### **6.3 PERSPECTIVA**

En vista de los buenos resultados que se obtuvieron en el desarrollo del proyecto se propone mejorar en proceso de extracción de características y hacerlo de una manera completamente automática y evitar los errores humanos en la obtención de estas al usar la ayuda diagnóstica.

## BIBLIOGRAFIA

1. MedLine plus  
<http://www.nlm.nih.gov/medlineplus/spanish/ency/article/000071.htm>
2. Pontificia Universidad Catolica de Chile  
[http://www.toraxuc.cl/cancer\\_pulmonar.html](http://www.toraxuc.cl/cancer_pulmonar.html)
3. Trabajos Academicos de la Universidad de Sevilla  
<http://bibing.us.es/proyectos/abreproy/11854> Aportaciones a la Segmentación y Caracterización de Imágenes Médicas 3D, Ignacio García Fenoll.
4. California Institute of Technology  
<http://www.work.caltech.edu/~boswell/IntroToSVM.pdf> Introduction to Support Vector Machines Dustin Boswell August 6, 2002
5. Asociación Internacional de Economía aplicada  
<http://www.asepelt.org/ficheros/File/Anales/2003%20-%20Almeria/asepeltPDF/55.pdf> Modelos de Clasificación basados en Maquinas de Vectores Soporte L. Gonzalez Abril luisgon@us.es -Universidad de Sevilla
6. Inteligencia Artificial Máquinas de Vectores de Soporte Capítulo IV Gerardo Colmenares
7. Laboratory of mathematics in Imaging – harvard medical school  
<http://lmi.bwh.harvard.edu/papers/pdfs/2003/martin-fernandezCOURSE03b.pdf> Técnicas Clásicas de Segmentación de Imagen Marcos Martin 4 de mayo de 2004.
8. <http://www.svms.org/tutorials/Gunn1998.pdf> ISIS Technical Report Support Vector Machines for Classification and Regression Steve Gunn 14 May 1998
9. R. C. Gonzalez, R. E. Woods, & S. L. Eddins Digital Image Processing Using matlab®, Prentice-Hall, 2004
10. Asociación de la Sociedad Española de Hipertensión <http://www.seh-lelha.org/concor2.htm> Luis M. Molinero
11. Yang Liu; Jinzhu Yang; Dazhe Zhao; Jiren Liu; , "Computer aided detection of lung nodules based on voxel analysis utilizing support vector machines," *BioMedical Information Engineering*, 2009. *FBIE 2009. International Conference on Future* , vol., no., pp.90-93, 13-14 Dec. 2009 doi:10.1109/FBIE.2009.5405784

URL:<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5405784&isnumber=5405767>

12. Xujiang Ye; Xinyu Lin; Dehmeshki, J.; Slabaugh, G.; Beddoe, G.; , "Shape-Based Computer-Aided Detection of Lung Nodules in Thoracic CT Images," *Biomedical Engineering, IEEE Transactions on* , vol.56, no.7, pp.1810-1820, July 2009  
doi:10.1109/TBME.2009.2017027 URL:  
<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5073252&isnumber=5073244>
13. ELCAP Public Lung Image Database  
<http://www.via.cornell.edu/databases/lungdb.html>
14. Center for Disease Control and Prevention - National Program of Cancer Registries (NPCR) <http://apps.nccd.cdc.gov/uscs/toptencancers.aspx>

## ANEXOS

**Anexo A.** CD con los algoritmos del desarrollo del proyecto, matrices de prueba y entrenamiento, evaluación del proyecto y 50 estudios TAC.