



Kongens Lyngby 2012
IMM-M.Sc.-2012-91

Easing scoring in ER and Ki-67 breast cancer histopathological images

Resumen PFC

Gonzalo R. Ríos Muñoz

Institución: Danmarks Tekniske Universitet

Supervisores: Anders Lindbjerg Dahl y Rasmus Larsen

Coordinador Erasmus: Francisco Valera Pintor

Cotutor: Francisco Valera Pintor

Área de conocimiento: Ing. Telemática

Fecha de lectura: 25 de enero de 2012

Información del proyecto

El proyecto se realizó en el departamento de Informática y Modelado Matemático de la **Universidad Técnica de Dinamarca (DTU)** durante el primer cuatrimestre del curso 2011-2012 bajo el programa de intercambio Erasmus. En la universidad de destino se cursó bajo el nombre de Master Thesis Project y estuvo supervisado por los profesores. La carga total de la tesis fue de 30 ECTS, siendo defendida el día 25 de enero de 2012 en Dinamarca ante un tribunal formado por varios profesores, el jefe del departamento, los supervisores y varios estudiantes que fueron invitados, así como un evaluador externo designado por la universidad.

En la **Universidad Carlos III Madrid** mi supervisor fue **Francisco Valera Pintor**, quien también fue mi coordinador académico Erasmus. El proyecto se incluyó dentro del departamento de Telemática de la universidad.

Introducción

La tesis centra su estudio en el análisis de imágenes realizadas a nivel microscópico sobre muestras de biopsias de tejido del pecho y de cómo estas se procesan para evaluar el desarrollo del cáncer para poder agilizar y obtener una puntuación para las mismas.

Para ello es necesaria una segmentación de los núcleos de las células que conforman las muestras, las cuales han sido previamente tintadas gracias a dos biomarcadores, Ki-67 y ER. Esto es simplemente poder diferenciar mediante software los tipos de elementos que componen la imagen para poder asignarles una clase. El principal objetivo de la segmentación es poder diferenciar y aislar las células en estado de proliferación de las normales. Después de eso, un sencillo y rápido algoritmo de evaluación es aplicado basado en el tamaño de los núcleos, en base a observaciones estadísticas de los mismos.

El estudio intenta facilitar y reducir la carga de trabajo que los patólogos han de procesar para facilitar los resultados del diagnóstico con la máxima rapidez disponible, incrementando las posibilidades de recuperación satisfactoria en el caso de que el cáncer esté presente.

La tesis emplea tratamiento y análisis de imágenes de tejido procedente del pecho de las pacientes, detectando si hay células sospechosas de contribuir al desarrollo de tumores para así implementar una manera de puntuar y evaluar el desarrollo de la enfermedad en los pacientes. Es por esto que un análisis a posteriori de los resultados obtenidos es adecuado para poder cuantificar y expresar con números la eficiencia de esta aproximación, que a fin y al cabo, termina siendo una estimación en base al comportamiento estadístico del tamaño de las células presentes en las biopsias. El algoritmo de segmentación empleado para cumplir la tarea que permite distinguir las células positivas de las negativas se describe más adelante. En caso de interés el texto detallado se puede consultar en [1] y [3].

El proceso para obtener las muestras de tejido sigue el procedimiento médico pertinente y puede variar de un laboratorio a otro. En primer lugar, una muestra del tejido de la paciente es tomada, biopsia (normalmente suelen ser cilíndrica). Seguidamente la muestra es cortada en finas capas a las que se aplica un tinte biológico. De esta forma los biomarcadores tratan de unirse a las células con las que tienen mayor afinidad, es este proceso bioquímico el que produce el cambio de tonalidad necesario para poder diferenciar las distintas células visualmente.

Una vez que este proceso es completado se toma una imagen en alta resolución, siendo esta la imagen la que se estudiará y analizará. Con estos archivos, el procesado de la imagen se realiza, obteniendo una nueva imagen con información acerca de la probabilidad de que las células del tejido se estén multiplicando (comportamiento característico del cáncer) o no, haciendo posible la puntuación de la enfermedad en función del número de células positivas frente a negativas. Los registros médicos y software para la adquisición de la imagen fueron proporcionados por la empresa danesa colaboradora del proyecto **Visiopharm**.

Problemas que solucionar

Ahorrar tiempo de procesado y reducir la carga de trabajo son los problemas que motivaron el proyecto y que se trataron de solventar para así contribuir a que los médicos y especialistas del sector confíen en el análisis de imágenes asistido por ordenador.

Muchas de las veces los propios expertos tienen que contar de manera manual las células presentes en las biopsias, una a una. Este método asegura la precisión y los buenos resultados en el diagnóstico final, pero el tiempo hasta que el paciente es notificado del grado de desarrollo del cáncer puede variar de semanas a meses dependiendo de los recursos y prioridades de los laboratorios. La velocidad de desarrollo del cáncer es también un factor limitante, así que cuanto más pronto se descubre la enfermedad mejor y más rápido se puede empezar con el preoperatorio, tratamiento o intervención necesaria, incrementando significativamente las posibilidades de afrontar con éxito la enfermedad que cada año afecta a más y más mujeres en el mundo.

Esta tesis trata de proponer un nuevo método para puntuar el desarrollo del cáncer a los ya existentes. Su objetivo no es sustituir los demás, sino probar que otros acercamientos son posibles e igualmente válidos independientemente de la segmentación o la técnica empleada a la hora de contar los núcleos. Son las personas las que deben beneficiarse de estos avances y descubrimientos.

Extracción de las imágenes

El estudio llevado a cabo se basa en la extracción de imágenes de tejido procedentes de biopsias de pecho. Las biopsias son cortadas en finas capas (de aproximadamente una célula de espesor y de 0.6 a 2 mm de diámetro) que a su vez son distribuidas en planchas de parafina siguiendo un patrón de coordenadas. Esta técnica es la conocida como TMA (Tissue Microarray). Una vez las muestras están ordenadas se procede a aplicar el proceso de revelado haciendo cambiar de color a las células. De esta forma es posible discernir y clasificar las células.

Proceso de tintado

El proceso de tintado inmunohistoquímico consiste en la reacción bioquímica que produce un cambio de color en las células bajo la presencia de biomarcadores. Estos biomarcadores se unen al núcleo celular normalmente durante el estado de división de las células cambiando la coloración de las mismas como consecuencia de las reacciones químicas. Es el cambio de color en concentraciones grandes de células lo que permite a los doctores situar el tumor.

Captura de las imágenes

Las imágenes fueron capturadas empleando microscopios Hamamatsu, utilizando dos niveles de aumento a la hora de realizar la adquisición. Los archivos obtenidos de los TMAs se almacenaron en el formato especial de Hamamatsu, *ndpi*. Cada uno de los *cores*, secciones circulares de la biopsia, tenían un tamaño de unos 5500x5500 pixels para el mayor nivel de magnificación. Los archivos fueron convertidos a su vez a formato *tif* gracias al software proporcionado por Visiopharm para, en la medida de lo posible, acceder a los mismos de una manera mucho más cómoda empleando programas de imagen más comunes así como hacerlo compatible con *MatLab*.

Base de datos

La base de datos consistía de una serie de imágenes de los TMAs obtenidas de los microscopios. Estos archivos ocupaban alrededor de 700 MB cada uno, con lo que procesar cada TMA por separado era algo engorroso y que implicaba analizar todas y cada una de las secciones a la vez. Es por ello que se decidió emplear el software '**VIS**' (Visiopharm Integrator System) y extraer uno a unos los *cores* que fueron almacenados independientemente bajo la extensión '*.tif*'.

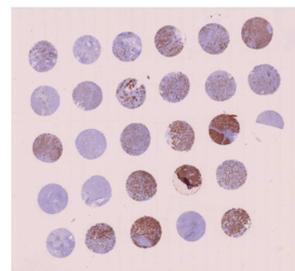
Independientemente, me fueron proporcionados 10 TMA's, 5 correspondientes al caso ER y otros 5 para el Ki-67. Cada TMA a su vez consistía de unas 27-28 secciones circulares. Aislado estos últimos, cada sección rondaba los 60-70MB, dando un número total de píxeles cercano a los 30 millones por core. El hecho de emplear un nivel de aumento tan alto, con la respectiva alta resolución en las imágenes, ayuda a obtener mejores resultados al poder distinguir mucho mejor las células, pero tienen un problema añadido, el tiempo de cómputo se incrementa, así como el almacenamiento necesario requerido para la base de datos. Para cada uno de los núcleos de la biopsia se contaba con una puntuación realizada a mano por expertos patólogos, para poder comparar y arrojar datos acerca del rendimiento final del algoritmo propuesto.

<i>TMA</i>	<i>Number of cores</i>
ER_KK1	28
ER_KK2	27
ER_KK3	27
ER_KK4	28
ER_KK5	8
Total	118

(a) ER TMA database.

<i>TMA</i>	<i>Number of cores</i>
KI67_KK1	28
KI67_KK2	27
KI67_KK3	27
KI67_KK4	28
KI67_KK5	8
Total	118

(b) Ki-67 TMA database.



Los *cores*, finas secciones circulares obtenidas de las biopsias, fueron tintados bioquímicamente empleando las técnicas habituales que vienen utilizando los laboratorios. La principal característica de este tipo de proceso es el cambio en la coloración de los elementos que componen las imágenes, importante en el diagnóstico.

Los principales elementos que constituyen una sección de la biopsia son:

- **Núcleos azules:** Células sanas que no reaccionan ante la presencia del tinte de los biomarcadores. Células normales que no contribuyen directamente a la presencia del cáncer.

- **Núcleos marrones:** Éstos son los “malos”, células en su estado de proliferación. No por ello han de ser confundidas como células cancerígenas. Las células en su ciclo de división son más sensibles a los biomarcadores. Lo que verdaderamente caracteriza el cáncer son las grandes concentraciones de este tipo de células en un área reducida de tejido, por lo que contando la presencia de este tipo de elementos es lo que permite a los doctores obtener un porcentaje para así evaluar el grado de desarrollo de la enfermedad.
- **Fondo:** Dejando a un lado los dos tipos de núcleos, el resto del tejido es considerado como elementos que componen el fondo de la imagen y que en este estudio no ofrecen información relevante.

Criterios de puntuación

Dependiendo del tipo de tinte aplicado se empleará un criterio diferente para la evaluación:

- **Puntuación ER = positiva o negativa**

Siendo **positiva** cuando más del 1% de las células presentes en la región son marrones y **negativa** en caso contrario.

- **Puntuación Ki-67 = % de núcleos positivos**

En el que el porcentaje se obtiene directamente de dividir los casos positivos entre el número total de núcleos en la región (positivos y negativos, o lo que es lo mismo marrones y azules).

Algoritmo de segmentación

El concepto de segmentación reside en diferenciar los píxeles que pertenecen a núcleos azules y marrones, así como distinguir el fondo de la imagen de los mismos. A cada pixel según su proximidad con las clases deseadas se le asignó una etiqueta para permitir al software una gestión de los mismos más limpia y simplificada. Un ejemplo de cómo quedaría una segmentación ideal lo podemos encontrar en la siguiente imagen.

Previo a la clasificación de los píxeles de la imagen, el algoritmo fue entrenado, necesitando dos diccionarios, uno de intensidad (color) y otro correspondiente a la máscara de clase para cada elemento. Para ello varios parámetros de entrada son necesarios:

- **Imagen de entrenamiento:** una porción representativa de textura/tejido de la cual se extrajeron los diccionarios.
- **Imagen de Máscara de entrenamiento:** matriz multidimensional representando todos los tipos de texturas que contribuyen a la formación de la imagen de entrenamiento. Una dimensión por clase, 3 en este caso.
- **Imagen de Test:** imagen a estudiar sobre la que se aplicó el segmentado empleando los diccionarios. Es la salida del algoritmo, la imagen debidamente segmentada.

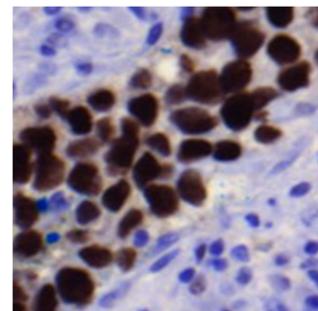
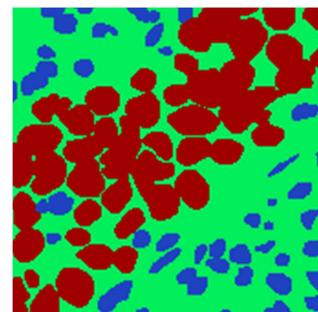
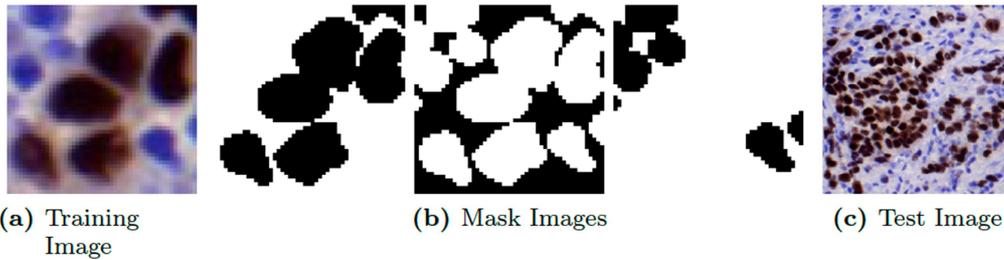


Imagen inicial



Segmentación deseada

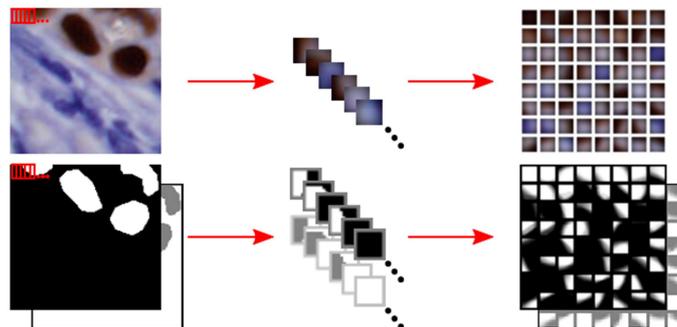


Para la obtención de las máscaras de entrenamiento, en la histopatología de cáncer de mama, me centré en el estudio de los núcleos azules y marrones, diferenciándolos entre ellos y a su vez de la membrana citoplasmática y demás regiones que contribuyen a formar el fondo de la imagen. Cada nuevo elemento que se desea clasificar no hace sino añadir una nueva capa a las ya existentes, una nueva categoría.

Luego este proceso se realiza a mano, empleando programas de tratamiento de imágenes como Windows Paint o Photoshop. Sobre la imagen de entrenamiento se toman 3 capas, si el pixel indicado pertenece a la categoría de la capa en cuestión este se pinta de negro, en caso contrario de blanco. De esta forma la posición de un pixel solo podrá pertenecer a una única clase, siendo por tanto negro en una única capa y en el resto ha de permanecer blanco, sinónimo de no pertenecer a la misma.

Una vez que todas las imágenes necesarias se han obtenido, el siguiente paso que ha de realizarse es la obtención de los diccionarios. Sobre la imagen de entrenamiento se eligen de manera aleatoria unos parches que representen la imagen y que se incluyen para formar parte del diccionario de intensidad. Este diccionario tiene información acerca del color de los píxeles del parche. El mismo parche correspondiente a la misma posición es tomado de la matriz con las capas de las máscaras y es a su vez incluido en el diccionario de etiquetas.

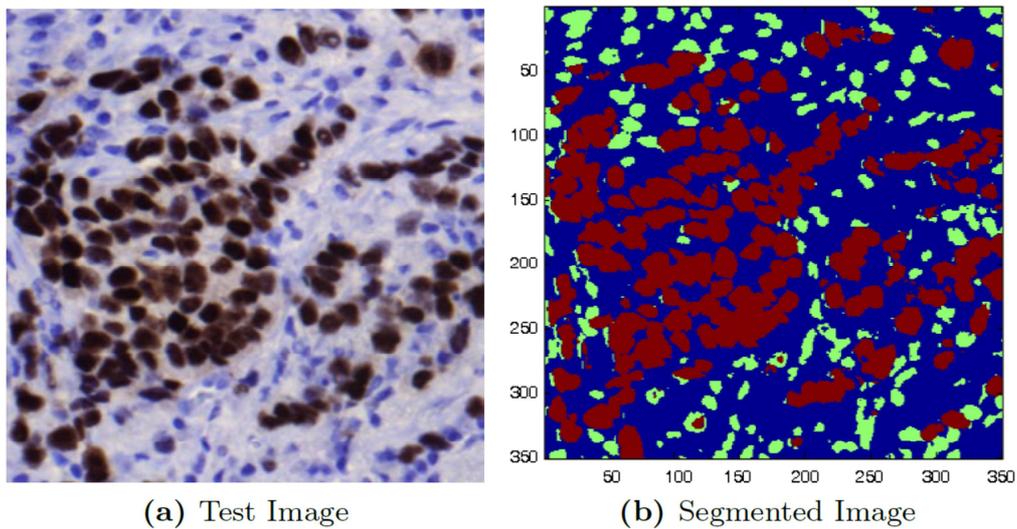
Para mejorar la selección de los parches y tomar los más representativos y cargados de información algunos parches se añaden o quitan de los diccionarios para mejorarlo y optimizarlo. Este proceso se realiza varias veces. Un esquema con la descripción del proceso se puede encontrar a continuación:



El siguiente paso es la segmentación propia de los núcleos. La imagen de test es analizada en esta ocasión. A mayor tamaño mayor tiempo de procesado. A la hora de procesar la imagen se van seleccionando parches continuos y se comparan con los de los diccionarios. Los parches se van solapando ofreciendo una mejor y suavizada segmentación a la vez que se obtiene una información más veraz acerca de la imagen. Es entonces cuando por el método del vecino más próximo se asocia el parche de intensidad de la imagen de test con uno perteneciente al

diccionario de intensidad y se le asocia la correspondiente máscara. Es así como la imagen de test queda segmentada.

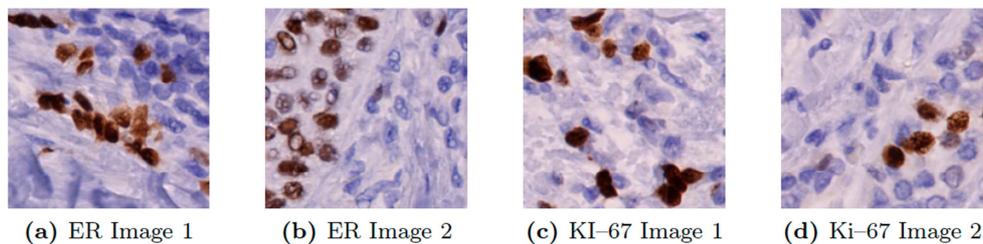
Este algoritmo ofrece una eficiencia muy alta, con un error a la hora de clasificar los píxeles entorno al 7%, un porcentaje muy esperanzador. Un ejemplo de segmentación puede apreciarse en la siguiente imagen:



Estimador

El estimador trata de describir mediante una distribución normal el tamaño en píxeles de los núcleos azules y marrones, para luego aplicar el caso general a imágenes independientes de tejido y obtener la deseada puntuación.

Para el estudio extraje algunas imágenes de las muestras. De esta manera pude contar todos y cada uno de los núcleos marrones y azules con los que obtener una base de datos fiel y verídica con la que contrastar la eficiencia del estimador. El tamaño de las imágenes elegido fue de 200x200 píxeles, ni muy grandes ni muy pequeñas. De esta manera la media y varianzas de los valores del tamaño de cada núcleo mejoraban. Para tener una buena base de datos realicé manualmente la cuenta sobre un total de 20 imágenes para cada tipo de tinto, contando más de 2000 núcleos. Un ejemplo de las imágenes de empleé:



<i>Image</i>	<i>brown nuclei</i>	<i>blue nuclei</i>	<i>Scoring Ratio(R)</i>
ER Image 1	21	33	0.3889
ER Image 2	24	25	0.4898
Ki-67 Image 1	17	28	0.1522
Ki-67 Image 2	7	39	0.3778

Para el completo estudio del algoritmo de segmentación, se emplearon varios valores para los parámetros dando un total de 12 combinaciones finales sobre las que se analizó el estimador. Éstos fueron los parámetros elegidos:

- **Tamaño de los parches (*atomSize*):** 3, 4, 5 & 6.
- **Número de parches por diccionario:** 500, 1000 & 1500.

Criterio de puntuación

Como así fue introducido anteriormente, el objetivo de este proyecto es la estimación de la relación de núcleos marrones frente a azules en base a los píxeles que contienen las imágenes empleando una estimación probabilística basada en una aproximación gaussiana. La fórmula que se aplicó fue:

$$R = \frac{\#brown\ nuclei}{\#brown\ nuclei + \#blue\ nuclei}$$

En lo respectivo a la segmentación y al método para realizar la cuenta de núcleos, varios parámetros nuevos fueron introducidos. Para facilitar su lectura, se empleó una notación al respecto:

- **Blue pixels (p_{bl}):** número de píxeles pertenecientes más probablemente a la segmentación de célula azul.
- **Brown pixels (p_{br}):** número de píxeles pertenecientes más probablemente a la segmentación de célula marrón.
- **Blue nuclei (n_{bl}):** número de núcleos azules contenidos en una imagen de la base de datos. Valor obtenido a mano.
- **Brown nuclei (n_{br}):** número de núcleos marrones contenidos en una imagen de la base de datos. Valor obtenido a mano.

Aplicando estas nuevas abreviaturas la ecuación anterior queda simplificada a:

$$R = \frac{n_{br}}{n_{br} + n_{bl}}$$

Fórmula corregida

En un primer acercamiento traté de cuantificar el número de células presentes en la imagen simplemente con la relación entre el número de píxeles pertenecientes a cada clase de segmentación:

$$\hat{R} = \frac{\#brown\ pixels}{\#brown\ pixels + \#blue\ pixels}$$

Pero como pude comprobar por experiencia los tamaños de los núcleos azules y marrones no eran parecidos, luego no se puede estimar con esta aproximación.

Luego para conseguir un nuevo modelo para evaluar las imágenes de tejido, propuse una nueva ecuación basada en la anterior con algunas variaciones que incluían la definición de un par de nuevas constantes. Con esta nueva aproximación la limitación del tamaño es tomada en consideración. El otro problema que se solucionó fue el relacionado con los distintos niveles de zoom que tenían las imágenes, que fue compensado gracias al uso de las nuevas variables.

Continuando con la notación anterior, una nueva variable era necesaria para relacionar el número de píxeles (p) y el número de núcleos (n) para cada caso (marrón y azul). La nueva variable fue definida bajo el nombre de w medida en píxeles por célula o núcleo. Consecuentemente la relación entre estas variables puede adivinarse fácilmente:

$$p = w \cdot n$$

Como en el caso anterior, habrá una variable w_{br} para el caso marrón y w_{bl} para el azul. Para la nueva estimación de \hat{R} la formula obtenida fue:

$$\begin{aligned} \hat{R} &= \frac{\hat{n}_{br}}{\hat{n}_{br} + \hat{n}_{bl}} = \frac{\frac{p_{br}}{w_{br}}}{\frac{p_{br}}{w_{br}} + \frac{p_{bl}}{w_{bl}}} = \frac{\frac{1}{w_{br}}}{\frac{1}{w_{br}} + \frac{w_{br}}{w_{bl}} \cdot \frac{p_{bl}}{p_{br}}} \cdot \frac{p_{br}}{p_{br} + \frac{w_{br}}{w_{bl}} \cdot p_{bl}} \\ &= \frac{p_{br}}{p_{br} + \frac{w_{br}}{w_{bl}} \cdot p_{bl}} = \frac{p_{br}}{p_{br}} \cdot \frac{1}{1 + \frac{w_{br}}{w_{bl}} \cdot \frac{p_{bl}}{p_{br}}} = \frac{1}{1 + \frac{w_{br}}{w_{bl}} \cdot \frac{p_{bl}}{p_{br}}} \\ &= \frac{1}{1 + \hat{C} \cdot P}, \end{aligned}$$

en el que el estimador se modeló como una función del número de células estimado en la imagen. Las células fueron expresadas en términos de relaciones existentes entre el número de píxeles por células y los píxeles obtenidos directamente después de aplicar el algoritmo de segmentación.

Resultados de las pruebas

Doce diferentes tests se hicieron para cada caso de estudio, combinando los parámetros de segmentación y discutiendo su eficiencia y grado de error. Para la obtención del error se empleó la siguiente función de error:

$$|e| = |R - \hat{R}|$$

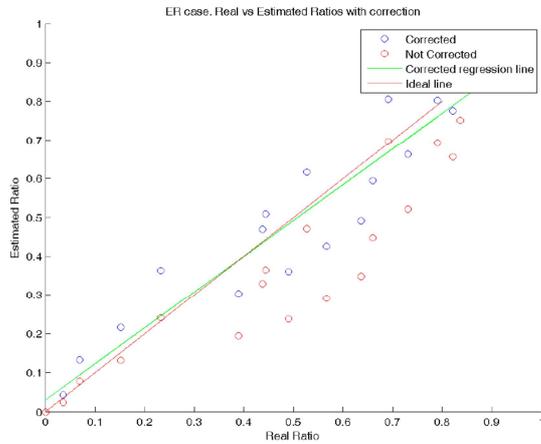
Pasos de prueba

Primero se obtuvo la base de datos con las imágenes de 200x200 píxeles, un total de 20 por caso. Se contaron y anotaron los núcleos azules y marrones encontrados.

Seguidamente se realizaron 12 segmentaciones diferentes sobre cada una de las imágenes que componían la base de datos empleando diferentes imágenes de entrenamiento. Con la segmentación se pudo obtener información crucial acerca de los píxeles, con lo que el estimador de \hat{C} pudo ser calculado como una nueva distribución obtenida de la división de las variables \hat{w}_{br} y \hat{w}_{bl} .

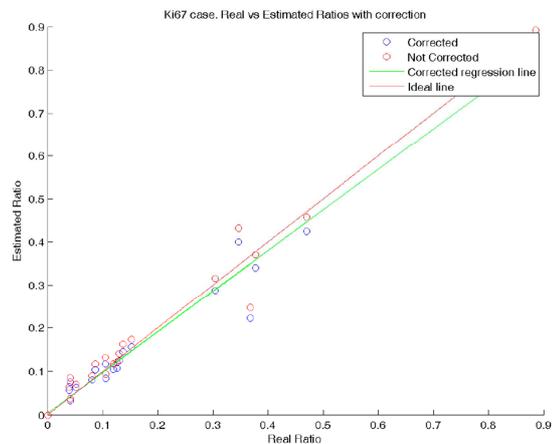
Para cada caso los resultados fueron almacenados y el error cometido calculado basado en las anotaciones de las cuentas realizadas a mano.

Ejemplos de los resultados obtenidos



Caso ER.

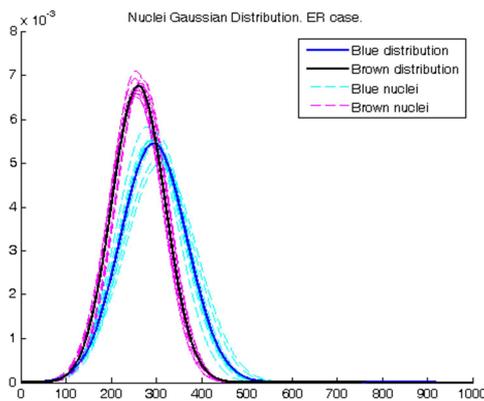
Real vs Estimated ratios. Los puntos corregidos y no corregidos se muestran a la vez junto con la regresión lineal en verde y la ideal en rojo.



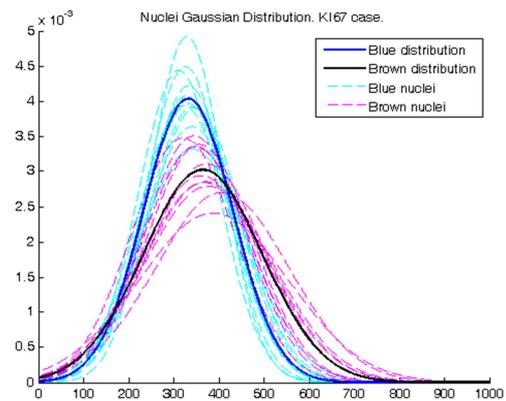
Caso Ki-67.

Distribuciones de núcleos

Las distribuciones de los núcleos se pudieron aproximar a sendas normales. Obteniendo para cada caso el aspecto que a continuación se muestran:



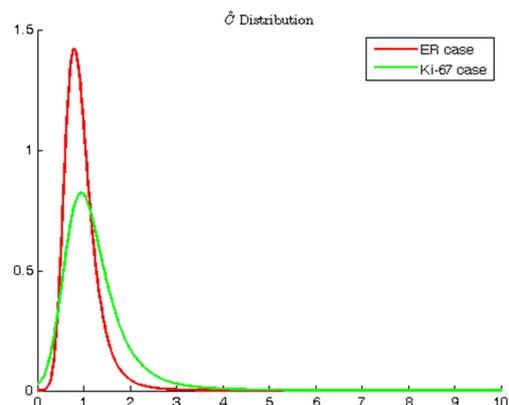
Caso ER. \hat{w}_{br} y \hat{w}_{bl} .



Caso Ki-67. \hat{w}_{br} y \hat{w}_{bl} .

Como \hat{C} es la relación entre ambos estimadores, y ya que la división de dos gaussianas no da otra distribución gaussiana, tuve que consultar varias publicaciones en estadística para finalmente obtener la distribución de \hat{C} que tenía este aspecto:

En el caso de Ki-67, las distribuciones tenían diferentes desviaciones típicas, lo que hacía que las curvas obtenidas fuesen más planas. En el



caso de ER, los resultados fueron mucho mejores, debido a lo similares que eran las medias obtenidas en cada caso. Mejor este último caso en lo que a distribuciones se refiere. Como se puede comprobar \hat{C} ronda el valor 1, con lo que el tamaño del núcleo azul y marrón son muy cercanos.

Discusión del error

El propósito de evaluar el acercamiento es decidir si los resultados obtenidos con el algoritmo basado en la estimación rápida de núcleos es válido y suficientemente bueno para ser empleado en la práctica. Por esa razón fue necesario el análisis de las 12 segmentaciones.

Como se puede apreciar en las tablas:

Test	atomSize	nPatches	$E[e]$	$E[e ^2]$
Test01	3	500	0.0479	0.0036
Test02	3	1000	0.0527	0.0045
Test03	3	1250	0.0539	0.0047
Test04	4	500	0.0498	0.0038
Test05	4	1000	0.0497	0.0039
Test06	4	1250	0.0549	0.0049
Test07	5	500	0.0504	0.0041
Test08	5	1000	0.0507	0.0039
Test09	5	1250	0.0500	0.0040
Test10	6	500	0.0507	0.0040
Test11	6	1000	0.0516	0.0041
Test12	6	1250	0.0518	0.0043

Errores en el caso ER

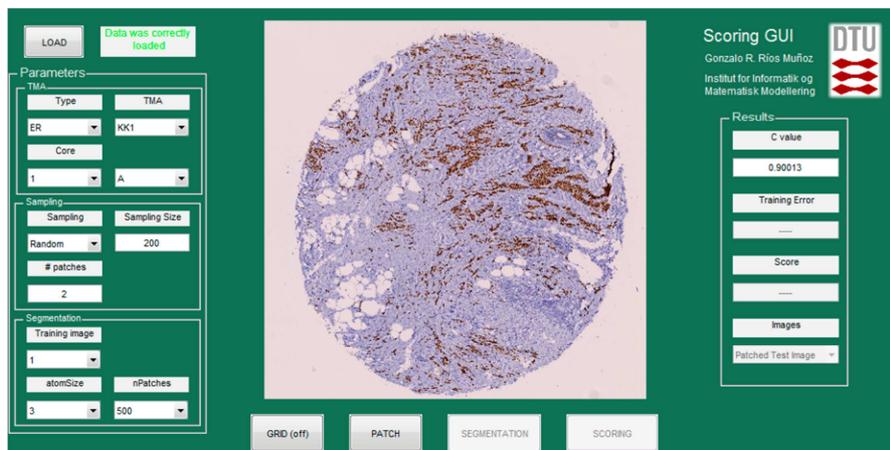
Test	atomSize	nPatches	$E[e]$	$E[e ^2]$
Test01	3	500	0.0263	0.0015
Test02	3	1000	0.0264	0.0019
Test03	3	1250	0.0249	0.0016
Test04	4	500	0.0276	0.0019
Test05	4	1000	0.0314	0.0024
Test06	4	1250	0.0260	0.0016
Test07	5	500	0.0267	0.0018
Test08	5	1000	0.0239	0.0015
Test09	5	1250	0.0238	0.0015
Test10	6	500	0.0302	0.0019
Test11	6	1000	0.0253	0.0016
Test12	6	1250	0.0230	0.0016

Errores en el caso Ki-67

El error absoluto obtenido en el caso ER es de alrededor del **2.5%** y con el Ki-67 es de **5%**. Resultados altamente dependientes de la imagen de entrenamiento que se empleó. Una mala elección en este proceso puede llevar a la obtención de resultados nefastos para nuestros intereses. Sin embargo, la segmentación en el caso Ki-67 fue mejor que para el ER.

GUI

Se realizó paralelamente una interfaz gráfica de usuario para facilitar el uso y obtención de resultados y así colaborar y ayudar a posibles estudiantes que quieran continuar el trabajo aquí explicado.



Comentarios

El principal objetivo de la tesis fue el desarrollo de un nuevo método de evaluación de las imágenes de tejido de pecho empleando los biomarcadores ER y Ki-67. Esta tarea fue satisfactoriamente realizada con la introducción de un nuevo estimador basado en el tamaño del núcleo de la célula. Los tests llevados a cabo sobre imágenes de pequeño tamaño dieron evidencias claras de que el método propuesto era adecuado siempre y cuando el algoritmo sea entrenado con una imagen adecuada de referencia. Para el salto a imágenes mayores, debido al descomunal tamaño de las imágenes, llevó mucho tiempo analizar una simple sección. Solamente dos de ellas fueron analizadas, obteniendo no obstante sendos buenos resultados (**Ki-67**: Real 10%. Estimado 15%. **ER**: Real: Positivo. Estimado: Positivo).

Otros problemas que afectaron negativamente al correcto desarrollo de la tesis fueron la carencia de conocimiento acerca de la evaluación histopatológica así como la no familiarización con los métodos que en ella se emplean. Como no soy ningún doctor ni experto en el tema en cuestión, llegó a ser en ocasiones frustrante el hecho de poseer el software necesario pero carecer de información sobre cómo o dónde aplicarlo. A veces el área del tumor era muy evidente, pero muchas imágenes escondían áreas afectadas por la enfermedad que solo son detectables por un ojo experto.

La herramienta aquí expuesta así como la interfaz gráfica está totalmente concebida para profesionales que sepan donde hay que aplicarla. La desarrollé lo más intuitiva posible a la vez que la hacía útil e interesante.

Hay todavía mucho que hacer en lo referente a la detección de tumores, afortunadamente estamos en el buen camino. Durante estos últimos 20 años ha habido un intenso avance y desarrollo tanto en aplicaciones como en estudios basados en técnicas asistidas por ordenador. Estoy muy satisfecho con mi contribución así como con los resultados obtenidos, espero que todo mi esfuerzo pueda ser continuado algún día y extendido con el único fin de ayudas a las personas que así lo necesiten.

Bibliografía

- [1] A.L. Dahl A. Karsnas and R. Larsen. Learning histopathological patterns.2011.
- [2] M. Dowsett A. Urruticoechea, I.E. Smith. Proliferation marker ki-67 in early breast cancer. *Journal of Clinical Oncology*, 23(28):7212–7220, October 2005.
- [3] R. Larsen A.L. Dahl. Learning dictionaries of discriminative image patches. In *British Machine Vision Conference*, 2011.
- [4] P. Stephan at About.com Guide. Her2/neu and diagnosis. http://breastcancer.about.com/od/diagnosis/p/her2_diagnosis.htm/. [Online; accessed 11-October-2011].
- [5] Medical Dictionary at thefreedictionary.com. Immunohistochemistry definition. <http://medical-dictionary.thefreedictionary.com/immunohistochemistry/>. [Online; accessed 11-October-2011].
- [6] M.C. Habib-H. Vacheret L. Xerri B. Devictor M.N. Lavaut M.Toga C. Charpin, L. Andrac. Immunodetection in fine-needle aspirates and multiparametric (samba) image analysis. receptors (monoclonal antiestrogen and antiprogesterone) and growth fraction (monoclonal ki67) evaluation in breast carcinomas. *Cancer*, 63(5):863–72, 1989.
- [7] C. Kim C. Jung. Segmenting clustered nuclei using h-minima transformbased marker extraction and contour parameterization. Number 57(10):2600-2604 in 57(10):2600-2604. *IEEE Trans. on Biomed. Eng.*, 2010. 57(10):2600-2604.
- [8] M.E. Burnett D.L. Commins, R.D. Atkinson. Review of meningioma histopathology. *Neurosurgical Focus*, 23(4):3–800, 2007.
- [9] A. Fischer. How breast cancer is diagnosed. <http://mammary.nih.gov/reviews/tumorigenesis/Fischer001/>. [Online; accessed 4-October-2011].
- [10] J E Gardi, J R Nyengaard, and H J G Gundersen. The proportionator: unbiased stereological estimation using biased automatic image analysis and non-uniform probability proportional to size sampling. *Computers in Biology and Medicine*, 38(3):313–328, 2008.
- [11] David V. Hinkley. On the ratio of two correlated normal random variables. *Biometrika*, 56(3):635–639, December 1969.
- [12] H. Lemke-H. Stein J. Gerdes, U. Schwab. The production of a mouse monoclonal antibody reactive with a human nuclear antigen associated with cell proliferation. *Int Journal Cancer*, 1983. 31:13-20.
- [13] Kononen J. Kallioniemi O.-P. Nocito, A. and G Sauter. Tissue microarrays (tmas) for high-throughput molecular pathology research. *International Journal of Cancer*, 94:1–5, October 2001.
- [14] U.S. National Library of Medicine Medical Subject Headings. Ki-67 antigen description. http://www.nlm.nih.gov/cgi/mesh/2011/MB_cgi?mode=&term=Ki-67+Antigen/. [Online; accessed 11-October-2011].

[15] F.A. Tavassoli P. Devilee. World Health Organization: Tumours of the Breast and Female Genital Organs. Oxford University Press.

[16] S. Ferretti R. Rinaldi E. Magri M. Indelli I. Nenci P. Querzoli, G. Albonico. Mib- 1 proliferative activity in invasive breast cancer measured by image analysis. Journal of Clinical Pathology, 1996. 49:926–930. doi: 10.1136/jcp.49.11.926.

[17] Keerthana Prasad, Avani Tiwari, Sandhya Ilanthodi, GopalakrishnaPrabhu, and Muktha Pai.

[18] R. Mathur L. Wise L.B. Kahn R. Mir, H. Johnson.

[19] Inc. The McGraw-Hill Companies. Animation: How the cell cycle works. http://highered.mcgraw-hill.com/sites/0072495855/student_view0/chapter2/animation__how_the_cell_cycle_works.html. [Online;accessed 13-October-2011].

[20] National Cancer Institute (USA). Dictionary of cancer terms: estrogen receptor. <http://www.cancer.gov/dictionary?cdrid=46409>. [Online; accessed 25-October-2011].

[21] National Cancer Institute (USA). Understanding cancer series. estrogen receptors/serms. <http://www.cancer.gov/cancertopics/understandingcancer/estrogenreceptors>. [Online; accessed 25-October-2011].

[22] V.R. Korde, H. Bartels, et al. Automatic segmentation of cell nuclei in bladder and skin tissue for karyometric analysis. In Biophotonics, Proceedings of the SPIE, volume 6633, 2007.