

ARTÍCULO BREVE

imageROC: software para la evaluación diagnóstica de la calidad de imágenes médicas

David Adame Brooks*, Andrés Ramírez Aguilera y Rafael A. Miller

Centro de Biofísica Médica, Universidad de Oriente, Santiago de Cuba, Cuba

Recibido el 21 de septiembre de 2011; aceptado el 13 de enero de 2012

PALABRAS CLAVE

Curvas ROC;
Sensibilidad;
Especificidad;
Calidad de imagen

Resumen

El análisis de la curva característica de operación del receptor (ROC) es una de las herramientas analíticas más importantes en el proceso de detección de la presencia o ausencia de una anomalía en una imagen. El análisis de las curvas ROC aporta, de manera sencilla, un magnífico indicador de la precisión de una prueba diagnóstica. En este trabajo, se propone la herramienta imageROC desarrollado para la evaluación diagnóstica de la calidad de imágenes. También se presentan algunos aspectos de esta metodología y su utilidad. Se concluyó que el programa propuesto proporciona una forma eficaz y rápida para evaluar completamente el desempeño conjunto del observador y del sistema de imágenes en una tarea de diagnóstico.

© 2011 ACTEDI. Publicado por Elsevier España, S.L. Todos los derechos reservados.

KEYWORDS

ROC curves;
Sensitivity;
Specificity;
Image quality

ROC image: software for the diagnostic assessment of medical imaging quality

Abstract

The analysis of the receiver operating characteristic (ROC) curve is one of the most important analytical tools in the process of detecting the presence or absence of an abnormality in an image. ROC curve analysis provides a simple and excellent indicator of the accuracy of a diagnostic test.

In this work we propose the imageROC tool developed for the quality of images for diagnostic assessment. Some aspects of this methodology and its usefulness are also presented. We concluded that the proposed program, along with the input by the observer, provides an effective and rapid means of assessing a diagnosis in an imaging system.

© 2011 ACTEDI. Published by Elsevier España, S.L. All rights reserved.

*Autor para correspondencia.

Correo electrónico: david.adame@cbiomed.cu (D. Adame Brooks).

Introducción

El análisis de la curva característica de operación del receptor (ROC) es una de las herramientas analíticas más importantes para caracterizar el desempeño humano en una tarea de análisis de imágenes, y es un área activa de investigación en imágenes médicas¹.

La efectividad clínica de un sistema de imágenes médicas depende usualmente no sólo de las características físicas del sistema y de la presencia o no de anomalías en el paciente, sino también de las características perceptuales del observador humano y del nivel crítico de confianza que el observador utiliza para determinar si una imagen es positiva o negativa¹.

Al cambiar el nivel crítico de confianza, cambian la sensibilidad y la especificidad. Entonces no se puede evaluar de forma completa el desempeño conjunto del observador y del sistema de imágenes en una tarea de clasificación, utilizando solamente los índices de especificidad y sensibilidad. Es necesario estimar y reportar todas las combinaciones entre especificidad y sensibilidad que pueden lograr, en conjunto, un cierto observador y un cierto sistema de imágenes. La curva ROC es uno de los métodos modernos que superan muchas de las limitaciones de los índices tradicionales².

Las curvas ROC proporcionan un buen índice de la capacidad de una prueba diagnóstica para diferenciar entre estados alternativos de salud, cuando los resultados se miden en escala ordinal, por intervalo o continua³. También son útiles para comparar distintos métodos diagnósticos y seleccionar umbrales de decisión⁴.

En este trabajo se propone un acercamiento a esta metodología y se presenta una herramienta que permite la construcción de la curva y el cálculo del área bajo la curva.

Base teórica

Curvas ROC

La metodología ROC se desarrolló en el contexto de la detección de señales electrónicas en los inicios de la década de 1950. A mediados de la década de 1960 se habían usado las curvas ROC en psicología y psicofísica experimental⁵.

El radiólogo Leo Lusted fue el primero en usarlas en 1967 en el proceso de toma de decisiones médicas, y en 1969 comenzó a aplicarlas en estudios con imágenes^{6,7}. El análisis ROC (que algunos autores han traducido como curvas de rendimiento diagnóstico) es una técnica matemática para describir y estudiar las propiedades de las pruebas diagnósticas⁸. Mediante ella, es posible hacer una estimación de la exactitud diagnóstica de un método determinado, modalidad de imagen o evaluador, mediante la medición de la capacidad de éste de clasificar correctamente los casos en categorías⁹.

Construcción de las curvas ROC

Clásicamente, la exactitud de una prueba diagnóstica se ha evaluado en función de dos características: la sensibilidad y

la especificidad⁷. Sin embargo, éstas varían en función del criterio elegido como punto de corte entre la población sana y la enferma.

El método tradicional para determinar experimentalmente la curva ROC de un observador humano es el siguiente:

- Se adquiere un conjunto representativo de imágenes que incluyan casos positivos y negativos.
- Se le pide al observador que califique cada imagen en una escala de varios niveles, que pueden ir desde “definitivamente anormal” hasta “definitivamente normal”.
- Entonces se pueden tabular diferentes valores de sensibilidad y especificidad para el conjunto de imágenes utilizando cada nivel de certeza como un umbral de decisión.

La curva ROC es un gráfico en el que se observan todos los pares de sensibilidad/especificidad resultantes de la variación continua de los puntos de corte en todo el rango de resultados observados¹⁰.

En el eje “Y” de coordenadas se sitúa la sensibilidad o la fracción de verdaderos positivos, definida como la probabilidad de clasificar correctamente a un individuo cuyo estado real sea el definido como positivo respecto a la condición que estudia la prueba, razón por la que también se denomina fracción de verdaderos positivos¹⁰.

En el eje “X” está la fracción de falsos positivos o 1-especificidad, definida como la probabilidad de clasificar correctamente a un individuo cuyo estado real sea el definido como negativo. Es igual al resultado de restar a uno la fracción de falsos positivos¹⁰.

Una prueba con diferenciación perfecta tiene una curva ROC que pasa por la esquina superior izquierda, donde S y E toman valores máximos (S y E = 1). Una prueba sin diferenciación, con igual distribución de resultados en los dos subgrupos, da lugar a una línea diagonal de 45°, desde la esquina inferior izquierda hasta la superior derecha.

Área bajo la curva ROC

El área bajo la curva ROC (ABC_{ROC}) es una medida global de la exactitud de una prueba diagnóstica. En el caso de las imágenes médicas, ésta mide la probabilidad de que, en forma aleatoria, pares de imágenes normales y anormales sean correctamente clasificadas debido a las anomalías percibidas en ellas⁹.

Por convenio, el ABC_{ROC} es siempre mayor o igual que 0,5. Toma valores comprendidos entre 0,5 (si no hay diferencias en la distribución de resultados de la prueba entre los subgrupos enfermo y sano) y 1 (cuando hay separación perfecta entre las dos distribuciones). Swets¹¹ interpreta el ABC_{ROC} de la manera siguiente: valores entre 0,5 y 0,7 indican baja exactitud; entre 0,7 y 0,9 pueden ser útiles para algunos propósitos, y un valor mayor de 0,9 indica exactitud alta.

En función del tipo de curva, el cálculo del ABC_{ROC} , así como de su error estándar, puede abordarse mediante un enfoque paramétrico o no paramétrico. Si se utiliza el

criterio no paramétrico, el ABC_{ROC} puede calcularse mediante un método trapezoidal o estimarse como el estadístico W de Wilcoxon⁹. Ambas metodologías pueden expresarse a través de las ecuaciones 1 y 2, respectivamente.

Método trapezoidal

$$\int_0^1 f(x)dx = \lim_{n \rightarrow +\infty} \sum_{K=1}^n f(x_K) \Delta_K x \quad (1)$$

Estadístico W de Wilcoxon

$$W = \frac{1}{n_A \cdot n_N} \sum_{j=1}^{n_A} \sum_{k=1}^{n_N} S(x_j, x_k) \quad (2)$$

donde $f(x)$ representa la ecuación que describe la curva; n_A y n_N representan el número de imágenes anormales y normales, y $S(x_j, x_k)$ es el puntaje dado al par compuesto por el j th paciente con la enfermedad y el k th paciente sin la enfermedad. $S(x_j, x_k)$ es 1 si $x_j > x_k$, $1/2$ si $x_j = x_k$ y 0 si $x_j < x_k$.

Bamber¹² demostró que el ABC_{ROC} coincide con la suma de rangos W obtenida mediante la prueba no paramétrica de comparación de medias de Wilcoxon. Hanley y McNeil⁹ exponen esta demostración matemática en el contexto médico y el método directo para el cálculo del error estándar de W , que permite obtener su intervalo de confianza. La metodología propuesta por Hanley y McNeil se expresa en la ecuación 3:

$$SE(W) \sqrt{\frac{W(1-W) + (n_A - 1)(Q_1 - W^2) + (n_N - 1)(Q_2 - W^2)}{n_A \cdot n_N}} \quad (3)$$

$$\text{Donde } Q_1 = \frac{W}{(2-W)} \text{ y } Q_2 = \frac{2W^2}{(1+W)}$$

La herramienta imageROC se implementó usando Matlab 2009a con el sistema operativo Windows XP. Su uso es muy sencillo e incluye tres módulos, como se muestra en la figura 1.

Introducción de los datos

Para evaluar las imágenes, primero se pasa a introducir información referente a las imágenes a través de este procedimiento de recolección de información. Se pueden introducir datos relacionados con la técnica, la institución, el equipo y las imágenes. Al introducir las imágenes, éstas se alistarán con sus nombres. Primero se introducen las imágenes que presentan alguna anomalía y después las imágenes de pacientes normales. Los tipos de imágenes que pueden introducirse son: Dicom (dcm), Bitmap (bmp) y jpg.

Niveles de decisión

La elección del número y niveles de decisión puede efectuarse de acuerdo a criterios ya establecidos por trabajos

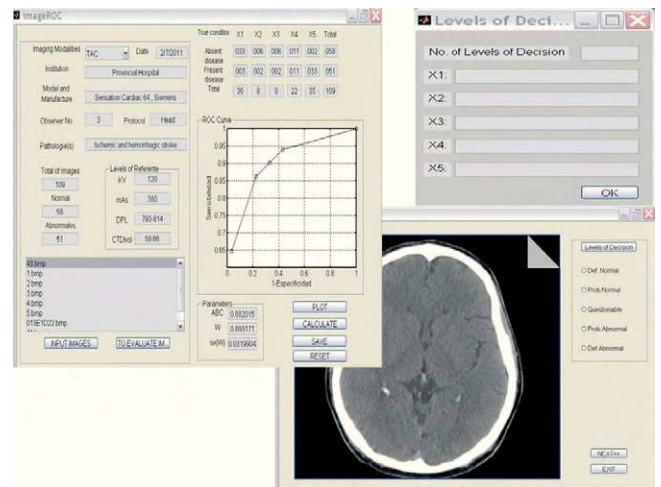


Figura 1 Visión general de la interfaz gráfica de imageROC.

anteriores, por razones teóricas basadas en la información clínica o fisiológica, pero otras veces es el propio investigador quien tiene que decidir los niveles de decisión que va a establecer. En esta herramienta tiene niveles ya preestablecidos que van desde seguramente normal, hasta seguramente anormal. El usuario tiene la opción de definir el número y los niveles que va a utilizar para evaluar su sistema de imágenes.

Cálculo del ABC_{ROC}

Después de evaluar las imágenes, aparecerán los datos relacionados a la evaluación en la tabla de contingencia. Posteriormente, el usuario puede visualizar la curva ROC obtenida y obtener el ABC a través de la regla trapezoidal (ABC_T), por el estadístico de Wilcoxon (ABC_W) y el error estándar (SE_W).

Almacenamiento de datos

Toda la información relacionada con la evaluación de un sistema de imágenes puede almacenarse para su utilización posterior en otros estudios y análisis. En la tabla 1 se observan los datos que se conservan con la herramienta imageROC.

En la figura 2 se observa la curva ROC obtenida a través de la herramienta imageROC, que se conserva junto con el resto de los datos.

Aplicación del software

En este trabajo se presenta un ejemplo que ilustra la aplicación de la herramienta imageROC en el análisis del efecto de la reducción de dosis en la detectabilidad de lesiones. Para esto, se tomaron un conjunto de imágenes de tomografía computarizada (TC) a diferentes índices de dosis, con el fin de determinar la capacidad de detección de

Tabla 1 Conjunto de datos salvados en un archivo txt por imageROC

2/7/2011
 Modalidad. imagenología: TC
 Institución: Hospital Provincial
 Modelo y fabricante: Sensation Cardiac 64, Siemens
 Observador: 3
 Protocolo: cabeza
 Enfermedades: ACV isquémico y hemorrágico
 Total de imágenes: 109
 Normales: 58
 Anormales: 51
 Niveles de referencia
 KV: 120-mAs: 380-DPL: 793-814-CTDI_{VOL}: 59,66
 Imágenes: '1.bmp' '2.bmp' '3.bmp' '4.bmp' '5.bmp' '41.bmp' '42.bmp' '43.bmp' '44.bmp' '45.bmp' '1030.bmp' '6030.bmp' 'A11.bmp' 'A12.bmp' 'CTDI1.bmp' 'CTDI2.bmp' 'E.bmp'
 ABC_r: 0,882015
 W: 0,893171
 SE(W): 0,0319904

ABC_r: área bajo la curva por el método trapezoidal;
 ACV: accidente cerebrovascular; CTDI_{VOL}: índice de dosis volumétrico; DPL: producto dosis longitud;
 Kv: kilovoltaje; mAs: miliamperaje/segundo;
 SE (W): error estándar de W; W: estadístico de Wilcoxon.

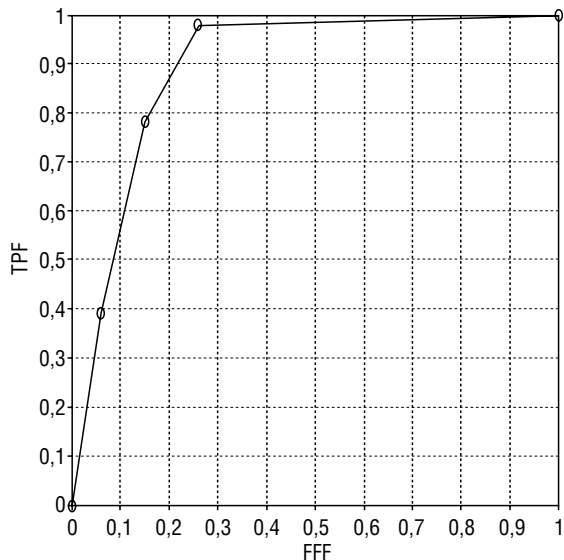


Figura 2 Curva ROC obtenida a través de la herramienta imageROC.

un observador con este sistema de imágenes. La tabla 2 muestra los valores de ABC_{ROC} y SE obtenidos a través de la herramienta imageROC. Con la disminución del ABC_{ROC}, disminuye la capacidad de clasificar correctamente las lesiones, hecho que se relaciona con la pérdida de calidad en la imagen, teniendo en cuenta el compromiso existente

Tabla 2 Valores del área bajo la curva ROC obtenidas mediante la herramienta imageROC

| Rangos de CTDI _{VOL} (mGy) | ABC _{ROC} ± SE |
|-------------------------------------|-------------------------|
| CTDI _{VOL} : 50-60 | 0,92 ± 0,069 |
| CTDI _{VOL} : 40-50 | 0,89 ± 0,073 |
| CTDI _{VOL} : 30-40 | 0,80 ± 0,103 |
| CTDI _{VOL} : 20-30 | 0,75 ± 0,120 |

ABC_{ROC}: área bajo la curva ROC; CTDI_{VOL}: índice de dosis volumétrico; mGy: miliGray; SE: error estándar.

entre la calidad de la imagen y la dosis de radiación en la TC. Como puede observarse, esta herramienta podría ser muy útil para evaluar la detectabilidad de lesiones en imágenes que han sido sometidas a algún tipo de procesamiento.

Conclusión

imageROC proporciona un medio eficaz y rápido por métodos no paramétricos para realizar la evaluación diagnóstica de la calidad de imágenes médicas. Por otra parte, permite almacenar información para la utilización posterior en otros estudios y análisis.

Con todo este procedimiento en pocos minutos se puede conocer la calidad diagnóstica del sistema de imágenes, partiendo de la percepción de los observadores, así como de datos referentes a los parámetros bajo los cuales se obtuvieron estas imágenes.

Esta herramienta no sólo es útil como apoyo a investigaciones, sino que también podría ser de gran utilidad para los profesionales vinculados con el análisis del desempeño conjunto observador-sistema de imágenes en una tarea de diagnóstico.

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

Bibliografía

1. Arámbula Cosío F. Análisis manual de imágenes médicas: ROC (Receiver Operating Characteristic). En: Curso de Introducción al Análisis de Imágenes Médicas y Cirugía Asistida por Computadora. México: UNAM; Curso semestre; 2011.
2. Deeks JJ. Systematic reviews of evaluations of diagnostic and screening test. BMJ. 2001;323:157-62.
3. Zweig MH, Campbell G. Receiver Operating Characteristic (ROC) Plots: A fundamental evaluation tool in Clinical Medicine. Clin Chem. 1993;39:561-77.
4. Piatt JH. Receiver operating characteristic curves. J Neurosurg. 2001;95:918-9.
5. Swets JA. The relative operating characteristic in psychology. Science. 1973;182:990-1000.

6. Lusted LB. Decision making studies in patient management. *NEJM*. 1971;284:416-24.
7. Lusted LB. Signal detectability and medical decisionmaking. *Science* 1971;171:1217-9.
8. Suetens P. *Fundamentals of Medical Imaging*. Leuven: Cambridge University Press; 2002. p. 213-20.
9. Hanley JA, McNeil BJ. The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve. *Radiology*. 1982;143:29-36.
10. Burgueño MJ, García Bastos JL, González Buitrago JM. Las curvas ROC en la evaluación de las pruebas diagnósticas. *Med Clin (Barc)*. 1995;104:661-70.
11. Swets JA. Measuring the accuracy of diagnostic systems. *Science*. 1988;240:1285-93.
12. Bamber D. The area above the ordinal dominance graph and the area below the receiver operating graph. *J Math Psych*. 1975;12:387-415.