

Herramientas para elaborar métodos de CAD del cáncer de mama

Una colaboración entre el CIEMAT y las Facultades de Ingeniería y Medicina de la Universidad de Porto

Tools for building Breast Cancer CAD methods

A collaboration between CIEMAT and the Faculties of Engineering and Medicine of the University of Porto

Guillermo DÍAZ-HERRERO; Jose Miguel FRANCO-VALIENTE; César SUÁREZ-ORTEGA; Manuel RUBIO DEL SOLAR y Raúl RAMOS-POLLÁN CETA-CIEMAT / CETA-CIEMAT

Miguel Ángel GUEVARA-LÓPEZ; Naimy GONZÁLEZ DE POSADA INEGI-FEUP Instituto de Ingeniería Mecánica y Gestión Industrial, Facultad de Ingeniería, Universidad de Porto / INEGI-FEUP Institute of Mechanical Engineering and Industrial Management, Faculty of Engineering, University of Porto

Isabel RAMOS y Joana LOUREIRO HSJ-FMUP Hospital de São João, Facultad de Medicina, Universidad de Porto / HSJ-FMUP Sao Paulo, Faculty of Medicine, University of Porto

En este artículo, se describen los principales resultados de la colaboración que existe actualmente entre CETA-CIEMAT y la Facultad de Ingeniería de la Universidad de Porto. Se presenta el “Terminal de Imágenes de Mamografía para el Análisis y Diagnóstico” (MIWAD), así como los procesos de análisis de datos mediante los cuales se utilizan infraestructuras de computación Grid para desarrollar métodos de Diagnóstico Asistido por Computador (CAD) del cáncer de mama. MIWAD es un software integral y novedoso que incorpora una interfaz gráfica de usuario especializada donde se combinan el procesamiento de imágenes digitales, el reconocimiento de pautas y técnicas de inteligencia artificial. En este trabajo, se describe su aplicación con éxito, en una experiencia piloto, al primer repositorio digital de cáncer de mama portugués.

Introducción

En este trabajo, describimos el proceso mediante el cual (1) se utilizó el MIWAD para crear conjuntos de datos (*datasets*) extraídos de mamografías anotadas por radiólogos especializados, y (2) se usaron recursos de computación Grid para elaborar clasificadores de aprendizaje automático para el CAD del cáncer de mama. En conjunto, esto constituye una experiencia piloto de éxito con MIWAD llevado a cabo con casos reales de cáncer de mama, donde hemos obtenido clasificadores automáticos produciendo valores del área bajo la curva ROC (*Receiver Operating Characteristic*) de entre 0,661 y 0,996 para los distintos conjuntos de datos elaborados.

El repositorio digital de cáncer de mama

Construimos el “Repositorio Digital de Cáncer de Mama” (BCDR) con casos anónimos procedentes de los archivos médicos históricos (en cumplimiento con la actual reglamentación de privacidad, ya que se utilizan también para la enseñanza de estudiantes médicos regulares o de post-grado) facilitados por la Facultad de Medicina – Hospital São João, Universidad de Porto (FMUP), Portugal. El BCDR está soportado y alojado en la Infraestructura de Repositorios Digitales (DRI) desarrollada por los autores en CETA-CIEMAT [1], y su modelo de datos es un subconjunto del formato de fichero médico DICOM [2] personalizada por radiólogos de la FMUP. En el momento de la redacción de este artículo, el BCDR está compuesto

por 950 casos, y las lesiones en mamografías son segmentadas (delimitadas) por especialistas y posteriormente clasificadas según la familia de clases del Sistema de Informes y Registro de Datos de Imagen de Mama (BI-RADS) [3]. Luego se elaboran los *datasets* para el entrenamiento de clasificadores de aprendizaje automático, que serán posteriormente integrados en MIWAD para su clasificación automática y su uso por especialistas como segunda opinión para respaldar su diagnóstico y sus decisiones relativas a la gestión del paciente.

Terminal de Imágenes de Mamografía para el Análisis y Diagnóstico (MIWAD)

MIWAD es una aplicación Java que permite el almacenamiento, recuperación y gestión de información sobre pacientes y casos en el BCDR y que constituye un terminal gráfico (ver Figura 1) para el procesamiento, análisis y diagnóstico de imágenes de mamografía. MIWAD incorpora distintos grupos de algoritmos de procesamiento y/o análisis de imágenes, tales como: filtros clásicos de pre-procesamiento, operadores morfológicos matemáticos, filtros de realce de imagen, etc. MIWAD se diferencia de los productos CAD comerciales porque está dedicado al diagnóstico asistido por computador (CADx) de posibles lesiones patológicas sobre regiones definidas por el usuario, y no a la detección asistida por computador (CADe). El proceso mediante el cual se diagnostica una imagen de mamografía con la ayuda de MIWAD sigue los pasos mostrados

Computación GRID • GRID Computing

en la Figura 1: (1) Selección de la Región de Interés (ROI), (2) Pre-procesamiento de las imágenes, (3) Segmentación, (4) Extracción de características, (5) Clasificación.

En el paso 4 (extracción de características), se extraen de la región segmentada medidas (características) de distinta naturaleza para producir un vector de características que es representativo de la región segmentada (perímetro, superficie, elongación, aspereza, contraste, etc.). En el paso 5 (clasificación), se asigna el vector de características extraído a una determinada clase, correspondiente a un tipo de lesión y/o a un estado de benignidad/malignidad.

Preparación de conjuntos de datos

A partir del BCDR, unos radiólogos especializados de la FMUP evaluaron, utilizando MIWAD, y clasificaron, utilizando BI-RADS, 286 casos. Solo se seleccionaron los casos en los que había imágenes de mamografía tanto *craneocaudal* (CC) como *oblicua medio-lateral* (OML) de las mamas izquierda y derecha, incluyendo información crítica asociada como el tipo de lesión, resultados de biopsias, etc. Unos pocos casos eran normales (sin lesión), mientras que el resto presentaba una de las siguientes lesiones: microcalcificaciones, calcificaciones, masas, distorsiones de la arquitectura o asimetrías. Los radiólogos trataron especialmente de localizar y segmentar posibles ROI para la misma lesión en las imágenes de mamografía asociadas tanto a la CC como a la OML en cada caso. Se realizó esta doble segmentación con éxito en 126 casos, produciendo cada uno dos vectores de rasgos (uno para cada imagen CC y OML), mientras que en los restantes 160 casos, solo se segmentó una ROI en la imagen CC o en la OML, produciendo un solo vector de características. Esto es atribuible a varias causas, entre ellas problemas técnicos, dificultades en la identificación de ROI's tanto en las imágenes CC como las OML o contingencias fortuitas. Para cada ROI segmentada (en imágenes OML y/o CC),

This paper describes the main results of the collaboration currently ongoing between CETA-CIEMAT and the Faculty of Engineering at the University of Porto. It introduces the “Mammography Image Workstation for Analysis and Diagnosis” (MIWAD) and the data analysis processes through which Grid infrastructures are used to develop Breast Cancer Computer-Aided Diagnosis (CAD) methods. MIWAD is a novel integrated software framework that integrates a specialized graphical user interface combining digital image processing, pattern recognition and artificial intelligence techniques. This paper describes its successful application in a pilot experience, on the first Portuguese Breast Cancer Digital Repository.

Introduction

In this work we describe the process by which (1) MIWAD was used to create datasets extracted from annotated mammograms by specialized radiologists and (2) Grid computing resources were used to build machine learning classifiers for breast cancer CAD. Overall, this constitutes a successful MIWAD pilot experience with real breast cancer cases, where we obtained automatic classifiers producing areas under the receiver operating characteristic (ROC) curve values between 0.661 and 0.996 for the different datasets built.

The Breast Cancer Digital Repository

We built the “Breast Cancer Digital Repository” (BCDR) with anonymous cases from medical historical archives (complying with current privacy regulations as they are also used to teach regular and postgraduate medical students) supplied by the Faculty of Medicine - Hospital São João, University of Porto (FMUP), Portugal. BCDR is supported and hosted on the Digital Repositories Infrastructure (DRI) platform developed by the authors at CETA-CIEMAT [1] and its data model is a subset of the DICOM medical file format [2] customized by radiologists at the FMUP. At the time of writing, BCDR is composed of 950 cases and lesions in mammograms are to be segmented (marked out) by specialists and then classified according to the Breast Image Reporting and Data System (BI-RADS) class family [3]. Datasets are then built to train machine learning classifiers to be later integrated in MIWAD for automatic classification to be used for second opinion by specialists as second opinion to support their diagnosis and patient management decisions.

Mammography Image Workstation for Analysis and Diagnosis (MIWAD)

MIWAD is a Java application that allows storage, retrieval and management of patient and case information in the BCDR and provides a graphical workstation (see figure 1) for processing, analyzing and diagnosing mammography

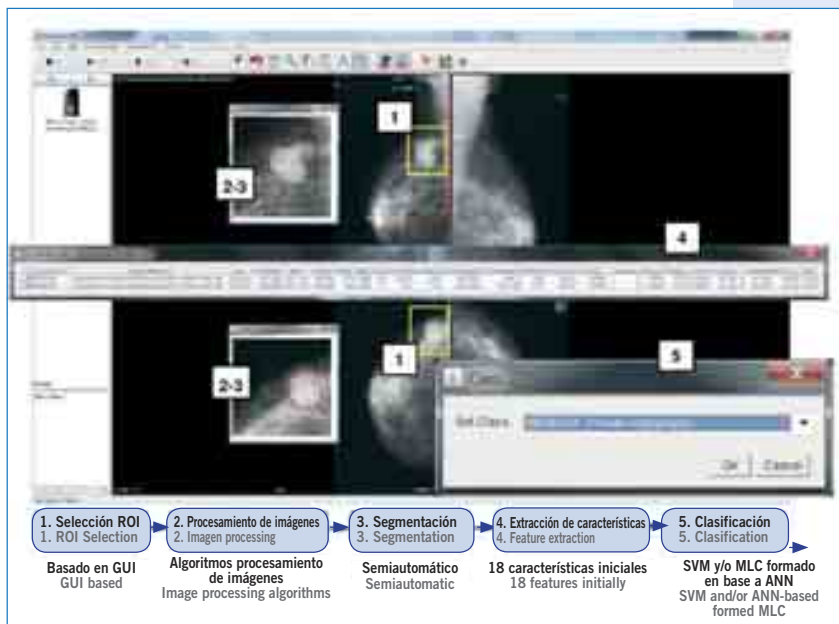


Figura 1. Mamografía CAD.
Figure 1. Mammography CAD.

se extrajo un vector de características y se construyeron varios datasets combinando distintos conjuntos de características extraídas y operaciones de normalización. La clasificación BIRADS efectuada por los especialistas dio lugar a un 62% de los casos benignos y un 38% malignos.

Exploraciones del espacio de búsqueda de clasificadores de aprendizaje automático basados en SVM y ANN

Usamos recursos de computación Grid para realizar exploraciones masivas de clasificadores basados en redes neurales artificiales (ANN) y máquinas de vectores de soporte (SVM), con el fin de buscar clasificadores de alto rendimiento para nuestros conjuntos de datos. Entrenamos miles de configuraciones, consumiendo unos 200 días de tiempo CPU. La validación fue realizada con *leave-one-out cross-validation* [4] y un variante del método bootstrap [5, 6] donde, antes del entrenamiento de cada configuración de una exploración, se marcó de forma aleatoria el 40% del conjunto de datos para el test y el resto para el entrenamiento. Después del entrenamiento y la validación, utilizado dos medidas para evaluar el rendimiento del clasificador en la parte del test del conjunto de datos: TESTPCT, el porcentaje de elementos clasificados correctamente (precisión) y TESTAZ, el área bajo la curva ROC (AUC). Solo los clasificadores de mejor desempeño seleccionados que resultasen de la exploración de todos los datasets se integrarán en MIWAD para el diagnóstico asistido del cáncer de mama.

Resultados y conclusiones

En la Tabla 1, se resumen los resultados de clasificación para los *datasets* en los que se seleccionó solo una ROI en la imagen OML o en la CC. Para cada tipo de SVM y ANN, se muestran los valores medios y máximos para TESTAZ y TESTPCT para todas las configuraciones entrenadas. En la Tabla 2, se muestran los mejores resultados de clasificación para los conjuntos de datos donde se podían segmentar las lesiones en ambas imágenes, junto con las configuraciones obtenidas de clasificadores específicos.

Los resultados de nuestra investigación muestran como se puede utilizar MIWAD para identificar qué combinaciones de procesamiento de imágenes digitales, reconocimiento de patrones y técnicas de inteligencia artificial son las más apropiadas para el diagnóstico del cáncer de mama, y que la identificación de la misma lesión tanto en las imágenes de mamografía CC como las OML aumenta de forma significativa el desempeño de

images. MIWAD implements distinctive groups of image processing and/or analysis algorithms, such as: classical pre-processing filters, mathematical morphological operators, image enhancing filters, etc. MIWAD differs from commercial CAD products because it is devoted to computer assisted diagnosis (CADx) for possible pathological lesions on user defined regions, rather than to computer assisted detection (CADe). The process by which a mammography image is diagnosed with the help of MIWAD goes through the steps shown in figure 1: (1) Region of Interest (ROI) selection, (2) Image Preprocessing, (3) Segmentation, (4) Features Extraction, (5) Classification.

In step 4 (features extraction) quantitative measures (features) of different nature are extracted out from the segmented region to produce a features vector which is representative of the segmented region (perimeter, area, elongation, roughness, contrast, etc.). In step 5 (classification), the extracted features vector is assigned to a certain class, corresponding to a lesion type and/or a benignancy/malignancy status.

Dataset Preparation

From BCDR, specialized radiologists at the FMUP using MIWAD evaluated and BI-RADS classified 286 cases. Only cases having both craniocaudal (CC) and mediolateral oblique (MLO) mammography images of left and right breasts were selected, including associated critical information such as lesion type, biopsies results, etc. A few cases were normal ones (no lesion) and the rest showed one of the following lesions: microcalcifications, calcifications, masses, architectural distortions and asymmetries. Radiologists made a special effort trying to locate and segment possible ROIs for the same lesion in both CC and MLO associated mammography images for each case. This double-segmentation was successfully performed in 126 cases producing each one two features vectors (one for each CC and MLO image), whereas in the remaining 160 cases only one ROI was segmented, either in the CC or in the MLO image, producing one single features vector. This was attributable to various reasons, including technical issues, difficulties in ROI identification in both CC and MLO images or casual contingencies. For each segmented ROI (in MLO and/or CC images) a features vector was extracted and several datasets were constructed combining different sets of extracted features and normalization operations. BIRADS classification by specialists produced 62% of the cases benign and 38% malignant.

Explorations of SVM and ANN based Machine Learning Classifiers Search Space

We used Grid computing resources to make massive explorations of artificial neural networks (ANN) and support vector machines (SVM) classifiers to find well performing classifiers for our datasets. We trained thousands of configurations taking around 200 days of CPU time. We used leave-one-out cross validation [4] and a variant of the bootstrap validation method [5, 6] where before training each configuration of an exploration, 40% of the dataset was labeled randomly for testing and the rest for

| MOTOR/ ENGINE | MAX TESTPCT | AVG TESTPCT | STDDEV TESTPCT | MAX TESTAZ | AVG TESTAZ | STDDEV TESTAZ |
|------------------|----------------|----------------|-------------------|---------------|---------------|------------------|
| libsvm.csvc | 0,750 | 0,654 | 0,031 | 0,778 | 0,660 | 0,053 |
| encog.ffsa | 0,709 | 0,656 | 0,025 | 0,788 | 0,704 | 0,044 |
| encog.ffbpiroc | 0,721 | 0,640 | 0,054 | 0,758 | 0,679 | 0,032 |
| encog.ffaaroc | 0,726 | 0,650 | 0,038 | 0,752 | 0,683 | 0,044 |
| encog.ffbp | 0,709 | 0,635 | 0,039 | 0,744 | 0,662 | 0,045 |
| encog.ffga | 0,709 | 0,618 | 0,052 | 0,737 | 0,622 | 0,060 |

Tabla 1.
Table 1.

Computación GRID • GRID Computing

| MOTOR/ ENGINE | PARÁMETROS DE CONFIGURACIÓN / CONFIGURATION PARAMETERS | VALORES DE CONFIGURACIÓN / CONFIGURATION VALUES | TESTPCT / TESTPCT | TESTAZ / TESTAZ |
|------------------|---|--|----------------------|--------------------|
| libsvm.csvc | kernel degree gamma shrink coef0 cost weight probestimates | pol 2 0.0048 true 1.0 64.0 0.5 true | 0,937 | 0,984 |
| libsvm.csvc | kernel degree gamma shrink coef0 cost weight probestimates | pol 2 0.0010 true 0.1 1.0 0.1 true | 0,937 | 0,982 |
| libsvm.csvc | kernel degree gamma shrink coef0 cost weight probestimates | sigm 2 0.0010 true 0.1 1.0 0.1 true | 0,921 | 0,981 |
| libsvm.csvc | kernel degree gamma shrink coef0 cost weight probestimates | rbf 2 0.01 true 0.6 1.0 1.0 true | 0,913 | 0,981 |
| encog.ffbp | layers and neurons learnrate momentum epochs | [36:54:27:14:2] 0.1 0.2 500 | 0,941 | 0,996 |
| encog.ffsa | layers and neurons starttempendtemp cycles epochs | [36:54:27:14:2] 100.0 2.0 100 200 | 0,940 | 0,983 |
| encog.ffbpiroc | ayers and neurons learnrate momentum epochs | [36:54:27:14:2] 0.1 0.2 500 | 0,902 | 0,964 |
| encog.ffsaroc | layers and neurons starttempendtemp cycles epochs | [36:54:27:14:2] 100.0 2.0 100 200 | 0,902 | 0,960 |

Tabla 2.
Table 2.

los clasificadores, lo que confirma los hallazgos descritos en la literatura. En términos generales, MIWAD ha resultado ser una aplicación informática apropiada para soportar al ciclo de vida completo del desarrollo y validación de métodos CAD del cáncer de mama.

Agradecimientos

Este trabajo forma parte del proyecto de colaboración en la investigación GRIDMED entre INEGI (Portugal) y CETA-CIEMAT (España). El CETA-CIEMAT agradece la aportación del Fondo Europeo de Desarrollo Regional. El Prof. Guevara agradece el POPH - QREN-Tipología 4.2 – Promoción de empleo científico, financiado por el ESF y MCTES, Portugal.

Referencias / References

- [1] R. Ramos-Pollan, et al., "Exploiting elnfrastructures for medical image storage and analysis: A Grid application for mammography CAD," in The Seventh IASTED International Conference on Biomedical Engineering, Innsbruck, Austria, 2010.
- [2] NEMA. (2010). Digital Imaging and Communications in Medicine. Available: <http://dicom.nema.org/>
- [3] C. J. D'Orsi, et al., Breast Imaging Reporting and Data System: ACR BI-RADS-Mammography, 4th Edition ed.: American College of Radiology, 2003.
- [4] B. Efron, "Estimating the Error Rate of a Prediction Rule: Improvement on Cross-Validation," Journal of the American Statistical Association, vol. 78, pp. 316-331, 1983.
- [5] J.-H. Kim, "Estimating classification error rate: Repeated cross-validation, repeated hold-out and bootstrap," Computational Statistics & Data Analysis, vol. 53, pp. 3735-3745, 2009.
- [6] B. Efron and G. Gong, "A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation," The American Statistician, vol. 37, pp. 36-48, 1983

training. After training and validation, we used two measures to evaluate classifier performance on the test part of the dataset: TESTPCT, the percentage of elements correctly classified (accuracy) and TESTAZ, the area under the ROC curve (AUC). Only selected well performing classifiers resulting from exploring all datasets would be integrated back into MIWAD for assisted breast cancer diagnosis.

Results and conclusions

Table 1 summarizes the classification results for datasets where only one ROI was selected either on the MLO or in the CC images. For each SVM and ANN engine it shows averages and maximum values for TESTAZ and TESTPCT for all configurations trained. Table 2 shows the best classification results for datasets where the lesions could be segmented on both images, along with the specific classifier configurations obtained.

The results of our investigation show how MIWAD can be used to identify what combination of digital image processing, pattern recognition and artificial intelligence techniques are most appropriate to diagnose breast cancer; and that identifying the same lesion in both CC and MLO mammography images increased meaningfully classifiers performance, which confirms literature reported findings. Overall, MIWAD proved to be an appropriate software framework supporting the full lifecycle to develop and validate breast cancer CAD methods.

Acknowledgements

This work is part of the GRIDMED research collaboration project between INEGI (Portugal) and CETA-CIEMAT (Spain). CETA-CIEMAT acknowledges the support of the European Regional Development Fund. Prof. Guevara acknowledges POPH-QREN-Tipología 4.2 – Promotion of scientific employment funded by the ESF and MCTES, Portugal.