



This item was submitted to Loughborough's Institutional Repository (<https://dspace.lboro.ac.uk/>) by the author and is made available under the following Creative Commons Licence conditions.

 **creative commons**
C O M M O N S D E E D

Attribution-NonCommercial-NoDerivs 2.5

You are free:

- to copy, distribute, display, and perform the work

Under the following conditions:

 **Attribution.** You must attribute the work in the manner specified by the author or licensor.

 **Noncommercial.** You may not use this work for commercial purposes.

 **No Derivative Works.** You may not alter, transform, or build upon this work.

- For any reuse or distribution, you must make clear to others the license terms of this work.
- Any of these conditions can be waived if you get permission from the copyright holder.

Your fair use and other rights are in no way affected by the above.

This is a human-readable summary of the [Legal Code \(the full license\)](#).

[Disclaimer](#) 

For the full text of this licence, please go to:
<http://creativecommons.org/licenses/by-nc-nd/2.5/>

Individualised Grid-enabled Mammographic Training System

M.H. Yap, A.G. Gale
AVRC, Loughborough University, LE11 3TU

ABSTRACT

The PERFORMS self-assessment scheme measures individuals skills in identifying key mammographic features on sets of known cases. One aspect of this is that it allows radiologists' skills to be trained, based on their data from this scheme. Consequently, a new strategy is introduced to provide revision training based on mammographic features that the radiologist has had difficulty with in these sets. To do this requires a lot of random cases to provide dynamic, unique, and up-to-date training modules for each individual. We propose GIMI (Generic Infrastructure in Medical Informatics) middleware as the solution to harvest cases from distributed grid servers. The GIMI middleware enables existing and legacy data to support healthcare delivery, research, and training. It is technology-agnostic, data-agnostic, and has a security policy. The trainee examines each case, indicating the location of regions of interest, and completes an evaluation form, to determine mammographic feature labelling, diagnosis, and decisions. For feedback, the trainee can choose to have immediate feedback after examining each case or batch feedback after examining a number of cases. All the trainees' result are recorded in a database which also contains their trainee profile. A full report can be prepared for the trainee after they have completed their training. This project demonstrates the practicality of a grid-based individualised training strategy and the efficacy in generating dynamic training modules within the coverage/outreach of the GIMI middleware. The advantages and limitations of the approach are discussed together with future plans.

Keywords: mammography, training, grid computing, DICOM, GIMI

1. BACKGROUND

The national UK Breast Screening programme is now 20 years old [1] and was established using analogue X-ray film to screen women. In recent years, trials of digital FFDM (full field digital mammography) have been undertaken in the UK and it is planned that by 2012 there will be a significant roll out of digital breast imaging nationally. At the same time the current age range of women being screened, aged 50 – 70 years, will be extended to encompass women aged 47 to 73 years. This will significantly increase the number of women screened annually from the current figures of almost 2.5 million women who are invited for screening [2]. Additionally, the use of digital imaging should be particularly useful for the screening of the younger aged women in this age range who tend to have mammographically denser breasts.

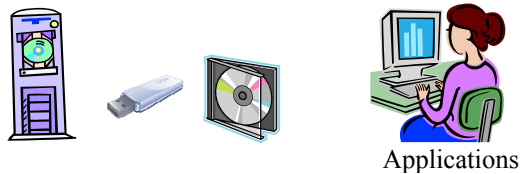
Currently training in interpreting screening cases in the UK is carried out in several ways (e.g. dedicated training courses, multi-disciplinary meetings), however because of the need to use mammographic film images it is always inherently somewhat limited in its user interaction ability as well as location of training; essentially being tied to where multi-viewers are located. Digital imaging offers the potential to use images more interactively as well as potentially in a wider range of different places. Additionally, other techniques such as CAD can be utilised in both digital screening and in training support. Training programmes can be developed based on certain general criteria and educational aims, however the ideal is for a training programme to be specifically targeted to the known needs of the individual. One way of doing this is to first measure the skills of individuals and ascertain how they differ from their peers and then offer training to improve any individual deficiencies.

In the UK all screeners (over 650 individuals) are invited to participate in a self assessment scheme (PERFORMS) twice a year from which they receive feedback on how well they have fared in interpreting sets of difficult recent screening cases as compared to the rest of the UK screening population [3]. From these data it is possible for an individual to realise that they have some difficulties/challenges in detecting, recognising and identifying certain key mammographic features as compared to their peers [4]. They can then seek out additional information which will improve their abilities in these key areas.

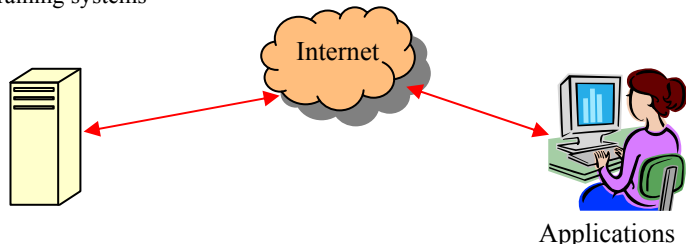
2. INTRODUCTION

Basic computer aided training can be conceptualised as starting with a very simple model, as shown in Figure 1 with software sitting on local hard disk, CD/DVD, or flash disks. The major limitation of this approach is the lack of information sharing and difficulty in trying to update the training system. Web-based systems facilitate information sharing and can provide a dynamic training system. However, security and speed of transmission are issues; the internet cannot deliver very large high quality images alone. This is where grid based systems can help. A grid-enabled training system can provide secure access to a very large dataset of large high resolution images for training purposes as well as facilitating a high speed network. This current development is part of the GIMI project [5] where the middleware provides a service oriented interoperability framework (sif) to enable existing and legacy data to support healthcare delivery, research and training [6].

(a) Computer-based training systems



(b) Web-based training systems



(c) Individualised grid-enabled training systems

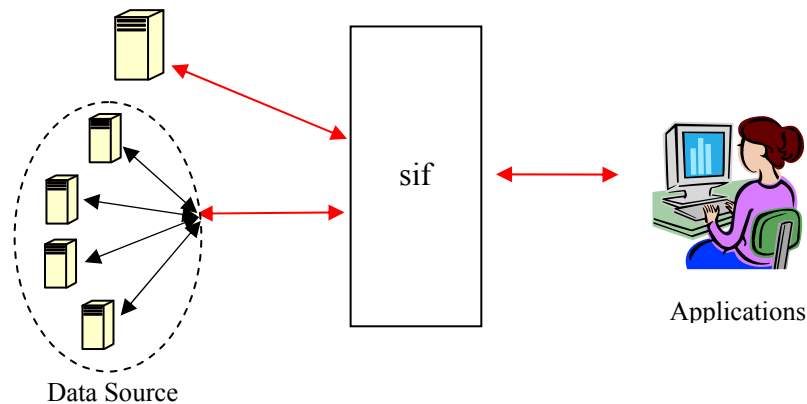


Figure 1 Illustration of training systems in different architectures: (a) computer-based training system, (b) web-based training system, (c) Individualised grid-enabled training system

3. METHODOLOGY

A grid enabled training system has been developed based on using middleware developed within the GIMI project which facilitates the deployment of medical images and other data between grid nodes situated at different hospital or university sites. This development is based on a large database of some 800 digitised screening cases which have been previously annotated, together with other patient-related information (e.g. pathology). From these exemplars a case can be selected based upon various criteria. The cases in the database are first audited to ensure that each case is useful for training purposes and that it contains the required information and new cases can be added to expand the dataset [7].

Case abnormal appearances in this dataset have been characterised using the terms shown in table 1. In the PERFORMS scheme participants identify mammographic features as being present in a case using the terminology which is also shown in table 1. Because of the good correspondence between the two sets of characterisations then it is feasible to utilise an individual's PERFORMS data which includes information concerning features that they have had difficulty either in identifying or in recognising [8] as a source so as to generate an individualised training programme. Essentially knowledge about what difficulties an individual has in identifying certain key mammographic features as elaborated from the PERFORMS scheme is then used to identify specific cases from the GIMI database which contain exemplars of those features for subsequent inspection and reporting. Therefore training can be tailored to the needs of the individual, appropriate feedback can be provided and their progress can be suitably measured.

Table 1. Correspondence mammographic feature labelling in PERFORMS and GIMI description.

| Feature labelling in PERFORMS | GIMI mammographic database feature descriptions |
|---|--|
| Mass | |
| <i>Well Defined Mass</i> | Circumscribed (Well-Defined or Sharply-Defined) Margins |
| <i>Ill Defined Mass</i> | Indistinct (Ill-Defined) Margins Microlobulated Margins |
| <i>Spiculate Mass</i> | Spiculated Margins |
| Architectural Distortion | Architectural Distortion |
| Asymmetry | Parenchymal Asymmetry |
| Suspicious Calcification | |
| <i>Diffuse Suspicious Calcification</i> | Diffuse/Scattered |
| <i>Benign Calcification</i> | Group or Clustered Linear |
| <i>Diffuse Benign Calcification</i> | Segmental |
| Other | Special Cases |

This training system has been developed by using the Eclipse Rich Client Platform (RCP) [9] which comprises a set of plug-ins to build a rich client application in Eclipse. Firstly, the individual uses their PERFORMS data to generate their own trainee profile as an input to the system and then their anonymous profile is stored in the training database (developed in Apache Derby). In the example trainee profile (figure 2), the two mammographic features that this particular trainee scored below average with in a recent PERFORMS assessment are shown, together with other pertinent data (e.g. mammographic case density, abnormality locations, case difficulty) that can also be used to identify cases for training purposes.

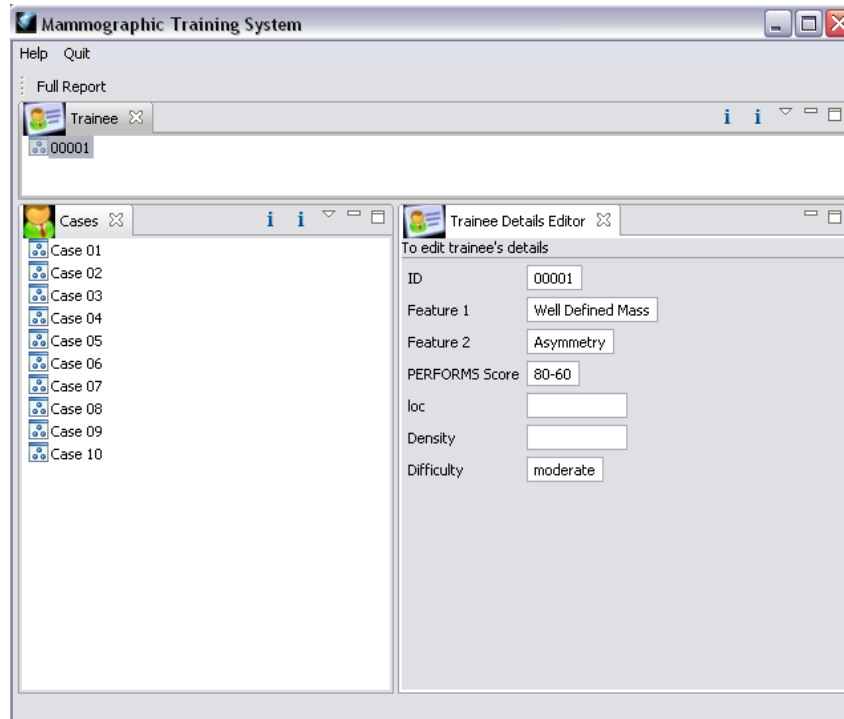


Figure 2. Example trainee details editor and training cases list.

When a trainee then logs in to the GIMI system, it selects potential training cases from the GIMI distributed servers based upon, for instance, the features indicated in the profile as being where the trainee has a need for improvement. This case selection process is currently random although it could be constrained by various criteria. The trainee then chooses a particular case to start with and the mammographic images of that particular case are then loaded in real-time from the servers and displayed using the jViewbox DICOM image viewer [10].

We have added some features to the image viewer, for instance to allow the user to mark up regions of interest in the images (figure 3) and also as feedback to show the actual location of regions of interest (here shown as highlighted circles around an image area) as previously defined by an experienced radiologist (figure 4). The trainee examines each case and then makes various decisions, indicating locations on the images where any mammographic feature/abnormality is considered to be located and completes an online evaluation form (figure 3), to indicate the feature labelling and case decisions. For feedback, the trainee can choose to have immediate feedback after examining each case or batch feedback after examining a number of cases. Feedback comprises comparing the trainee's responses to those of the experienced radiologist who initially described all the cases in the database. Current work is examining different types of feedback which is why the detailed location co-ordinates of the abnormality are given in figure 4. All the trainee's results are recorded in the database and a more detailed report can be generated after they have completed their training session. Currently (example given in figure 5) this indicates the performance of the individual by indicating the number of cases answered examined, accuracy in abnormality location detection, recall rates, positive predictive value (PPV), and negative predictive value (NPV).

4. DISCUSSION

The prototype system demonstrates the concept of using the grid to develop an individualised mammographic training system using GIMI middleware. From the prototype the image download speed remains an issue for such a real time distributed system. For instance, to download the four DICOM images of a case, at 33 Megabytes per image, it currently takes 30-60 seconds. Approaches to shortening this image fetch time are to pre-load the selected cases before the participant views them. Figure 6 shows the system running on a workstation with the images shown on one monitor and the text on the other. Although a workstation has been used in the development other client computers could be used such as a PC or laptop. Any type of client could be used as long as it has been granted the certificate to access the servers. However, the display resolution of such devices would engender different user interaction techniques.

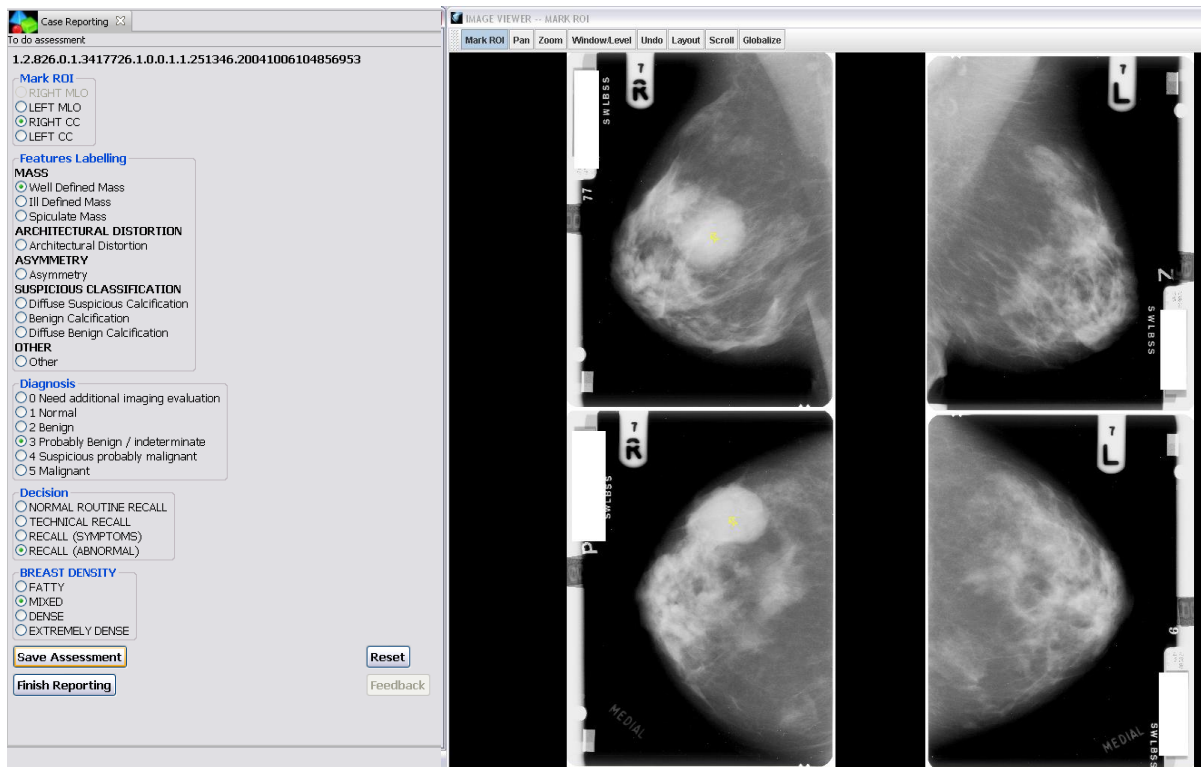


Figure 3. Reporting module.

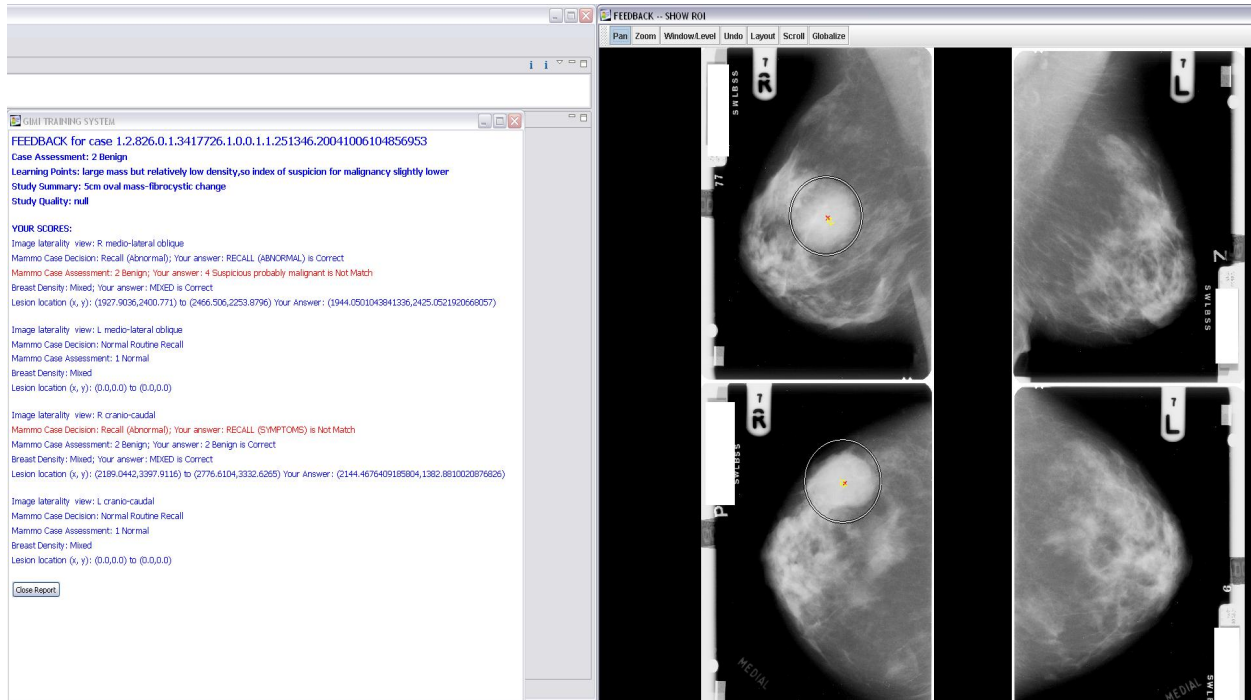


Figure 4. Feedback Module

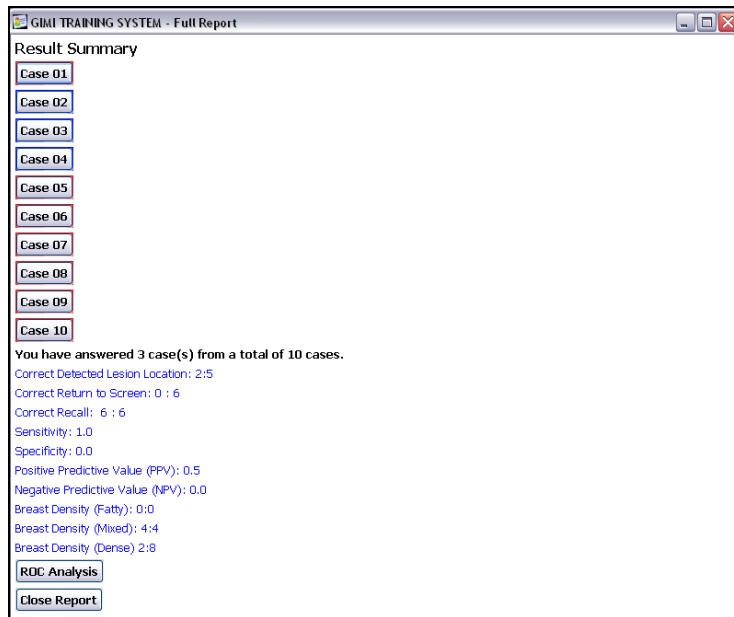


Figure 5. Example developmental report



Figure 6. System development workstation

A potential issue with an approach of basing training on data from the PERFORMS scheme is that in the scheme participants are presented with only a small number of difficult cases and their skills measured. The question can then be raised as to whether data from such carefully selected exemplar cases relate to an individual's real life data in screening. We have studied this separately on different occasions and our available evidence does point to this being the case which then supports this training approach [11].

5. CONCLUSIONS

To date a grid based training system has been developed and is currently being assessed by radiologists in terms of the information used in the system. A trial of the system is underway with potential trainees. Future developments will examine providing further detailed feedback, including ROC data analysis, to participants as well as the potential to use case ontology information.

The system enables existing cases in the grid database to be audited as well as permitting new cases to be annotated and added to enlarge this dataset as required. Permitting an individual to use their confidential PERFORMS data to select particular cases to examine for training has been demonstrated to be feasible and other research has shown that the difficulties elaborated for an individual whilst examining the PERFORMS scheme cases do relate to difficulties which that individual has when examining real life screening cases. Thus the approach facilitates an individualised training system. Expansion of the approach can be to utilise other characteristics of an individual's performance data to select which cases to present to the trainee, as well as intelligently selecting cases based on the ongoing performance of the trainee during training.

ACKNOWLEDGEMENT

GIMI is funded by the Technology Strategy Board's Collaborative Research and Development programme, following an open competition. We acknowledge the continued support of our colleagues from the Oxford University Computing Department.

REFERENCES

- [1] Patnick, J., "Celebrating 20 years of screening.", NHSCSP (2008).
- [2] Patnick, J., "Saving lives through screening: NHS Breast Screening Programme Annual Review 2008," NHSBSP (2008).
- [3] Gale, A.G., "PERFORMS - a self assessment scheme for radiologists in breast screening", *Seminars in Breast Disease: Improving and Monitoring Mammographic Interpretative Skills*, 6:148-152, (2003).
- [4] Scott, H.J., & Gale, A.G., "Breast screening: PERFORMS identifies key mammographic training needs." *British Journal of Radiology*, 2006, 79: S127-S133.
- [5] Power, D.J., Politou, E.A., Slaymaker, M.A. and Simpson, A.C., "Towards secure Grid-enabled healthcare", *Softw. Pract. Exper.*, 35:857-871, (2005). DOI: 10.1002/spe.692.
- [6] Simpson, A.C., Power, D.J., Slaymaker, M.A., Russell, D. and Katarova, M., "On the development of secure service-oriented architectures to support medical research. *International Journal of Healthcare Information Systems and Informatics*, 2(2):75-89, 2007.
- [7] Yap, M.H., Gale, A.G., "Grid-enabled mammographic auditing and training system", *Proc. SPIE Vol. 6919*, 69190A (Mar. 11, 2008)
- [8] Yap, M.H., Gale, A.G., Scott, H.J., "Generic Infrastructure for Medical Informatics (GIMI): The Development of a Mammographic Training System", E.A. Krupinski (ed.): *IWDM 2008, LNCS 5116*, pp. 577-584, Springer-Verlag Berlin Heidelberg 2008.
- [9] eclipse, <http://www.eclipse.org> [08 April 2008]
- [10] Laboratory of Neuro Imaging, UCLA, <http://www.loni.ucla.edu/Software/> [02 June 2007]
- [11] Scott, H.J., Evans, A., Gale, A.G., Murphy, A. & Reed, J. "The relationship between real life breast screening and an annual self assessment scheme. In *Image Perception, Observer Performance, and Technology Assessment.*" D. Manning and B Sahiner (Eds.) *Proceedings of SPIE 2009* (in press)