Anatomy & Physiology

Frenz and Jaffe, Anat Physiol 2014, 4:4
http://dx.doi.org/10.4172/2161-0940.1000e134

Editorial Open Access

Unlocking the Potential of PACS and VNA Data Stores

Christopher M Frenz[1*] **and Eric A Jaffe**[2]

[1]Department of Information Technology, Interfaith Medical Center, 1545 Atlantic Ave, Brooklyn, NY11213, USA

[2]Department of Medicine, Interfaith Medical Center, 1545 Atlantic Ave, Brooklyn, NY11213, USA

[*]**Corresponding author:** Frenz CM, Department of Information Technology, Interfaith Medical Center, 1545 Atlantic Ave, Brooklyn, NY11213, USA, Tel: 914 714-5635; E-mail: cfrenz@interfaithmedical.org

Rec date: Aug 14, 2014; **Acc date:** Aug 16, 2014; **Pub date:** Aug 18, 2014

Editorial

In recent years, the use in medicine of digital imaging technologies has grown at an extraordinary rate with the current use and application of these technologies far exceeding the early adoption of classical Radiology. Modern digital imaging techniques encompass technologies such as radiography, computerized transaxial tomography (CT scan), magnetic resonance imaging (MRI), ultrasound, and nuclear medicine-based techniques such as Positron Emission Tomography (PET). Moreover, concerted efforts are now being made to enhance the utility of these technologies by allowing multiple imaging technologies to be overlain and simultaneously analyzed, such as PET-CT scans, thereby further enhancing the diagnostic value of medical imaging.

This increased utilization of digital imaging technologies has lead most hospitals and medical centers to incorporate Picture Archiving and Communication Systems (PACS) into their IT infrastructure as a means of storage and retrieval of medical images pertaining to patient care. Modern PACS systems typically employ the Digital Imaging and Communications in Medicine (DICOM) standard as a means of defining the formats for file storage as well as the application protocols for the transfer of these images. Many of the original PACS systems were geared towards radiology images but often had some proprietary aspects which complicated image exchange between systems despite being based around DICOM standards. Non-radiological digital images were often stored in proprietary systems, which meant that medical images were often stored in multiple isolated silos within an organization's IT infrastructure. Over time, this has pushed most large healthcare organizations to adopt Vendor Neutral Archives (VNAs) as a preferred means of image storage in order to allow the images obtained from a variety of imaging sources to be archived in a single location in a more standardized way [1].

These vast collections of images and their associated reports represent a source of valuable but underutilized clinical data that can be used for medical data mining and imaging informatics. Analyses of images such as these can lead to a better understanding of the morphologies that contribute to disease and therefore the pathologies that are associated with a given condition. In turn, this knowledge will be of aid in the building of computer aided detection (CADe) and diagnosis (CADx) systems. CADx and CADe systems are a relatively recent phenomenon but they are beginning to see widespread adoption in the diagnosis of cancer. Such systems have been developed and successfully used to identify a variety of cancers from medical images including breast cancer [2] and lung cancer [3,4]. These CAD based techniques are often found to be better at detecting early stage malignancies than humans [4] and as such provide a source of valuable diagnostic aid. Moreover, there also exist CAD systems for the identification of coronary artery disease [5].

CADe and CADx systems are typically developed using machine learning algorithms that over time learn to perform the recognition of complex patterns, such as what geometries and dye localizations could be indicative of coronary artery disease when found in coronary angiography images. The development of accurate machine learning based CAD systems, however, is largely dependent on the availability and quality of the data used to train the system, in which the data desired to be used as input into the CAD system (e.g. medical images) is already associated with a known outcome (e.g. a diagnosis).

For example, consider a relatively common machine learning method such as an artificial neural network (ANN). Artificial neural networks are named after their biological counterparts in that they consist of a group of information processing units ("neurons") that operate in parallel. Signals can be passed between neurons through a series of weighted connections and neural networks are able to "learn" by adjusting the strengths of these connections until they can approximate a function that computes the proper output for a given input pattern. In order for an ANN to be trained to recognize coronary artery disease, the ANN would need to be exposed to multiple images of coronary arteries with and without signs of coronary artery disease. The ANN would take these images as input and make a prediction (output) as to whether or not the image was indicative of coronary artery disease. Since the diagnosis for each image is already known, the discrepancy between the prediction and the known outcome can be used to provide feedback to the ANN which will be used to change the strengths of the neural interconnections. After repeated exposure to the data in this training set and the resultant feedback, the ANN will begin to predict outcomes for data in the training set with a progressively higher degree of accuracy. Once training is complete and validation performed to ensure that the ANN can accurately evaluate data that was not used in the training set, the ANN can be considered as having predictive value [6]. While the machine learning algorithms are not limited to ANNs, it is these trained function approximations that perform the complex pattern recognition that is the hallmark of most CAD systems.

The value of PACS systems and VNAs is that they not only store a large number of medical images pertaining to a variety of morphologies and pathologies, but that the morphologies and pathologies are readily accessible in the reports that correspond to the images. This means that the images and reports contained in these archive systems provide an ideal source for obtaining the training data required to develop a CAD system. Moreover, the push for standardization on VNA architectures should also help promote the development of CADx and CADe systems. Better standardization of imaging formats and application protocols will provide the potential for reusable libraries to be developed that will allow for machine learning to be performed across data from a much greater range of medical imaging sources. It is time for collaboration between physicians and informatics professionals to become significantly more prevalent to help ensure that in the future collected clinical data can serve not only the source patient, but all patients as a source of training data for the development of improved diagnostic tools.

References

1. Agarwal TK, Sanjeev (2012) Vendor neutral archive in PACS. Indian J Radiol Imaging 22: 242-245.
2. Jung NY, Kang BJ, Kim HS, Cha ES, Lee JH, et al. (2014) Who could benefit the most from using a computer-aided detection system in full-field digital mammography? World J Surg Oncol 12: 168.
3. Bley TA, Baumann T, Saueressig U, Pache G, Treier M, et al. (2008) Comparison of radiologist and CAD performance in the detection of CT-confirmed subtle pulmonary nodules on digital chest radiographs. Invest Radiol 43: 343-348.
4. Marten K, Engelke C (2007) Computer-aided detection and automated CT volumetry of pulmonary nodules. Eur Radiol 17: 888-901.
5. Anders K, Achenbach S, Petit I, Daniel WG, Uder M, et al. (2013) Accuracy of automated software guided detection of significant coronary artery stenosis by CT angiography: comparison with invasive catheterization. Eur Radiol 23: 1218-25.
6. Frenz CM (2005) Give your .NET app brains and brawn with the intelligence of neural networks. MSDN Magazine.