Method and System of Building Hospital-Scale Medical Image Database

Summary (1024-character limit)
Hospital Picture Archiving and Communication Systems (PACS) contain vast amounts of underutilized informatics about disease conditions. As computer image processing and systems advance, PACS informatics may form the foundation for precision automated computer-aided diagnostics for a wide range of disease conditions. Development of such systems may improve diagnostic accuracy and better inform treatment, but creating systems and algorithms capable of “learning” to recognize and locate the image patterns of disease and associated labels is a difficult problem. Researchers at the National Institutes of Health Clinical Center (NIHCC) have developed a technology that applies deep learning to PACS images to produce a database where certain disease features are identified and spatially located. Researchers at the NIHCC seek licensing of the PACS.

NIH Reference Number
E-076-2017

Product Type
• Software

Keywords
• Computer Vision
• Deep Learning
• Medical Imaging Informatics
• Computer Assisted Diagnostics
• Picture Archiving and Communication Systems
• PACS
• National Institutes of Health Clinical Center
• NIHCC

Collaboration Opportunity
This invention is available for licensing.

Contact
• John D. Hewes
  NCI - National Cancer Institute
  240-276-5515
  John.Hewes@nih.gov

Description of Technology
Developing computer systems that can recognize and locate image features associated with disease is a challenge for developing fully-automated and high precision computer assisted diagnostics. Joint learning of language tasks in association with vision tasks (association of image features with text annotation) adds an additional level of challenge. Furthermore, scaling-up approaches from small to large datasets presents additional issues, particularly related to medical images. In this case, identifying such features requires specialized skill for even a human and the text descriptions from trained physicians may be variable, complex, and abstract. The application of deep learning to medical image feature detection in association with language recognition may aid in the development of precision automated computer-aided diagnostics for a wide range of disease conditions, based on large scale PACS datasets.

The technology developed by researchers at the National Institutes of Health Clinical Center (NIHCC), applies natural language processing techniques and deep learning methods to mine PACS images and generate large-scale image databases. Diseases can be detected and spatially-located within the dataset generated by this method. The generation of such datasets is an important step toward utilizing PACS informatics and development of fully-automated high precision computer diagnostic systems.

Potential Commercial Applications
- Computer Assisted Diagnostics
- Medical Image Informatics

Competitive Advantages
- Ability to create large-scale labeled medical image database

Inventor(s)
Xiaosong Wang (CC), Yifan Peng (NLM), Le Lu (CC), Zhiyong Lu (NLM), Ronald Summers (CC)

Development Stage
- Discovery (Lead Identification)

Publications
Wang X et al. ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases [arxiv abstract 1705.02315]

Patent Status