Xiaosong Wang, PhD

10 Center Drive #1C224, Bethesda, MD 20892, USA
http://wangxiaosong.net, xiaosong.wang@live.com, xiaosong.wang@nih.gov

EMPLOYMENT

Jul. 2015 – present **Visiting Fellow**
CAD Lab, National Institutes of Health Clinical Center, Bethesda, USA
- Build large-scale medical image dataset via data mining and NLP (DeepLesion, ChestX-ray, KeyImage dataset)
- Develop deep learning based algorithms for medical image analysis and CAD (Lymph Node Segmentation, Prostate Cancer CAD, Common Chest Thorax Disease Classification and Localization)

Dec. 2013 – Jun. 2015 **Product Manager Post-processing Workstations**
HSW BU, Shanghai United Imaging Healthcare, Shanghai, CHINA
- Lead a product definition and upstream marketing team (product managers / owners) for mutli-modality post-processing workstations.
- Plan road-maps of advanced post-processing workstations for medical images. Analyze and prototype clinical requirements from clients
- Lead clinical collaboration with radiology departments in top-grade hospitals
- Manage product, project and human resource (CT/MR/PET/Mammo advanced workstations, group of product managers/specialists)
- Design and practice validation and clinical trial of advanced applications to obtain CFDA approval

Sep. 2011 – Nov. 2013 **CAD Dept. Manager / Tech Lead / Algorithm Engineer**
HSW BU, Shanghai United Imaging Healthcare, Shanghai, CHINA
- Start and develop a new algorithm department from square one
- Lead algorithm R&D in CADe / CADx and many other post-processing applications in medical images (more details in R&D PROJECTS section)
- Manage project and human resource (group of 8 algorithm engineers)
- Design and practice validation and evaluation of advanced applications in clinical settings

EDUCATION

Oct. 2006 – Sep. 2011 UNIVERSITY OF BRISTOL, Bristol, UK
Ph.D in Computer Vision

Oct. 2005 – Sep. 2006 UNIVERSITY OF BRISTL, Bristol, UK
M.Sc in Computer Science with Distinction

Sep. 2001 – Jun. 2005 HARBIN INSTITUTE OF TECHNOLOGY, Harbin, CHINA
B.Sc in Computational Mathematics with Distinction

DISTINCTIONS

Nov. 2017 2017 NIH Clinical Center Director Award
Nov. 2016 RSNA 2016 Fellow Research Award
Aug. 2016 NIH Fellows Award for Research Excellence (FARE)
Sep. 2010 ECCV Student Travel Award
Sep. 2009 Best Industrial Paper Prize, BMVC 2009
Oct. 2006 – Sep. 2009 Great Western Research PhD scholarship, UK
Sep. 2006 Msc thesis with Distinction (top 5%)
Jul. 2005 Bsc thesis with Distinction (100 among over 5600 graduates)

SERVICES

Reviewer	Pattern Recognition, Medial Image Analysis, IEEE Transactions on Biomedical Engineering, Journal of Biomedical Informatics, IET Computer Vision, IEEE Transactions on Medical Imaging, Neurocomputing, Knowledge-Based Systems, Machine Vision and Applications, Pattern Analysis and Application
Chief Judge	Informatics / Computational Biology section, NIH Fellows Award for Research Excellence Competition 2018
Lead Judge	Poster competition for NIH Postbac Poster Day 2017, 13th Annual Graduate Student Research Symposium

INVITED TALK

Jul. 2017	Big Data, Weak Label and True Clinical Impacts for Radiology Imaging Diagnosis. *Medical Computer Vision and Health Informatics Workshop, CVPR 2017*
Jun. 2017	Weakly-supervised Classification and Localization of Common Chest Disease in X-ray using Deep learning. *NCBI, National Library of Medicine.*
Feb. 2017	Unsupervised Image Categorization in a Large Scale Radiology Image Database. *Computer Science Department, John Hopkins University.*
Mar. 2016	Towards large-scale radiology image database auto-annotation: a Deep Pseudo-task Learning approach. *NIH Pi Day, PiCo Talks 2016*

MEDIA COVERAGE

Oct. 02 2017	NIH releases massive database of chest x-rays. *By Greg Slabodkin, Health Data Management*
Oct. 01 2017	NIH Releases Chest X-Rays + Data. *By Carolyn Bloch , Federal Telemedicine News*
Sep. 29 2017	How 100K Chest X-Ray Images Could Improve AI and Patient Health. *By Jack Murtha, Healthcare Analytics News*
Sep. 28 2017	NIH releases massive database of chest x-rays. *By Brian Casey, AuntMinnie staff writer*
Sep. 28 2017	NIH to provide one of largest chest x-ray datasets for research. *By Amy Wallace, United Press International*
Sep. 27 2017	NIH Clinical Center provides one of the largest publicly available chest x-ray datasets to scientific community. *By Molly Freimuth, NIH news release*
Sep. 27 2017	NIH Releases 100,000 Chest X-Rays. *By John Commins, HealthLeaders Media*
Sep. 27 2017	NIH Clinical Center issues 100K x-ray images for machine learning. *By Subrata Thakar, Radiology Business*
March 2017	Trainee Research Prizes from the 2016 RSNA Scientific Assembly and Annual Meeting. *By Radiology Journal, Special Communication, March 2017 Volume 282, Issue 3*
Oct. 31 2016	Deep learning can find, label images in PACS. *By Erik L. Ridley, AuntMinnie staff writer*
Dec. 2009	Summary of Prizes Presented at BMVC 2009, *By Simon J.D. Prince, British Machine Vision Association News, Volume 20 Number 1*

PEER-REVIEWED PUBLICATIONS

[12] **X. Wang**, Y. Peng, L. Lu, Z. Lu, R. M. Summers. TieNet: Text-Image Embedding Network for Common Thorax Disease Classification and Reporting in Chest X-rays. IEEE CVPR 2018.

[11] K. Yan, **X. Wang**, L. Lu, L. Zhang, A. Harrison, M. Bagheri, R. M. Summers. Deep Lesion Graphs in the Wild: Relationship Learning and Organization of Significant Radiology Image Findings in a Diverse Large-scale Lesion Database. IEEE CVPR 2018.

[10] Y. Peng, **X. Wang**, L. Lu, M. Bagheri, R. Summers, Z. Lu. NegBio: a high-performance tool for negation and uncertainty detection in radiology reports. AMIA 2018 Informatics Summit, 2018 (Oral)

[9] **X. Wang**, Y. Peng, L. Lu, Z. Lu, M. Bagheri, R. Summers. ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases. IEEE CVPR, 2017 (spotlight)

[8] Y. Tsehay, N. Lay, **X. Wang**, B. Turkbey, J. T. Kwak, P. Choyke, P. Pinto, B. Wood, R. Summers. Biopsy-guided Learning with Deep Convolutional Neural Networks for Prostate Cancer Detection on Multiparametric MRI. IEEE ISBI, 2017

[7] **X. Wang**, L. Lu, H. Shin, L. Kim, M. Bagheri, I. Nogues, J. Yao, R. M. Summers. Unsupervised Joint Mining of Deep Features and Image Labels for Large-scale Radiology Image Annotation and Scene Recognition. IEEE WACV, 2017

[6] Y. Tsehay, N. Lay, H. Roth, **X. Wang**, J. T. Kwak, B. Turkbey, P. Pinto, B. Wood, R. Summers. Convolutional neural network based deep-learning architecture for prostate cancer detection on multiparametric magnetic resonance images. SPIE Medical Imaging, 2017

[5] I. Nogues, L. Lu, **X. Wang**, H. Roth, G. Bertasius, N. Lay, J. Shi, Y. Tsehay, R. M. Summers. Automatic Lymph Node Cluster Segmentation using Holistically-Nested Networks and Structured Optimization. MICCAI, 2016

[4] **X. Wang** and M. Mirmehdi. Archive Film Defect Detection and Removal: an Automatic Restoration Framework. IEEE Transactions on Imaging Processing (T-IP 2012), 21(8):3757-3769, March 2012.

[3] **X. Wang** and M. Mirmehdi. Archive Film Restoration based on Spatiotemporal Random Walks. In Proceedings of the 11th European Conference on Computer Vision (ECCV 2010), September 2010.

[2] **X. Wang** and M. Mirmehdi. HMM based Archive Film Defect Detection with Spatial and Temporal constraints. In Proceedings of the 20th British Machine Vision Conference (BMVC 2009), **Winner of the Best Industrial Paper Prize**, September 2009.

[1] **X. Wang** and M. Mirmehdi. Archive Film Defect Detection based on a Hidden Markov Model. In Proceedings of the 10th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS 2009), May 2009.

ABSTRACTS

[4] **X. Wang***, K. Yan*, L. Lu, R. Summers. Detection of Radiology Image Findings Using Large-scale Clinical Lesion Annotations, RSNA 2017.

[3] Y. Peng, **X. Wang**, L. Lu, M. Bagheri, R.Summers, Z. Lu: DeepText Mining Radiology Reports for Deep Learning Radiology Images. American Medical Informatics Association (AMIA) Annual Symposium (Oral), Nov. 2017

[2] I. Nogues, L. Lu, **X. Wang**, H. Roth, G. Bertasius, N. Lay, J. Shi, Y. Tsehay, R. M. Summers: Automatic Lymph Node Cluster Segmentation Using Holistically-Nested Deep Convolutional Neural Networks and Structured Optimization in CT Images. RSNA 2016

[1] **X. Wang**, L. Lu, H. Shin, L. Kim, I. Nogues, J. Yao, R. M. Summers: Automated Annotation of a Large Scale Radiology Image Database Using Deep Learning. RSNA 2016, **Winner of RSNA Fellow Research Award**, **Best paper in Imaging Informatics section**

PATENTS

[11] **X. Wang**, K. Yan, L. Lu, R. M. Summers: Detection of Radiology Image Findings Using Large-scale Clinical Lesion Annotations. U.S. Patent Application No. 62/514,223, 2017

[10] **X. Wang**, Y. Peng, L. Lu, Z. Lu, R. M. Summers: ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases. U.S. Patent Application No. 62/476,029, 2017

[9] L. Lu, **X. Wang**, R. M. Summers: Category Discovery and Image Auto-annotation via looped Deep Pseudo-task Optimization. US Patent Application, 62/302,096, 2016

[8] L. Lu, H. Roth, I. Nogues, R. M. Summers, **X. Wang**: Integrating Deep Boundary and Appearance Convolutional Neural Networks for Bottom-up Organ Segmentation. US Patent Application, 62/345,606, 2016

[7] L. Wang, **X. Wang**: Heart model building method, heart model registration and heart multi-plane reconstruction method. Chinese Patent Application: CN103839249A, 2014

[6] H. Li, H. Shi, **X. Wang**: Pectoralis segmentation method in breast image. Chinese Patent: CN104182965B, 2014

[5] H. Li, **X. Wang**: Region real-time segmentation method for medical image. Chinese Patent Application: CN104299217A, 2013

[4] Z. Chen, **X. Wang**, C. Qiu: Image processing method and image processing device. Chinese Patent Application: CN104462149A, 2013

[3] **X. Wang**: Simulation method and quantitative test method of liver perfusion. Chinese Patent Application: CN103902801A, 2012

[2] **X. Wang**: CT liver-perfusion image post-processing method and CT liver-perfusion method. Chinese Patent: CN103839249B, 2012

[1] Y. Fei, **X. Wang**, C. Li: Automatic window width and window level extraction method based on neural network. Chinese Patent Application: CN103310227A, 2012

OTHER PUBLICATIONS

[5] K. Yan*, **X. Wang***, L. Lu, R. M. Summers: DeepLesion: Automated Deep Mining, Categorization and Detection of Significant Radiology Image Findings using Large-Scale Clinical Lesion Annotations. arXiv:1710.01766, 2017

[4] **X. Wang**, Y. Peng, L. Lu, Z. Lu, M. Bagheri, R. M. Summers. ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases. arXiv:1705.02315, 2017

[3] **X. Wang**, L. Lu, H.-C. Shin, L. Kim, M. Bagheri, I. Nogues, J. Yao, R. M. Summers. Unsupervised Joint Mining of Deep Features and Image Labels for Large-scale Radiology Image Categorization and Scene Recognition. arXiv:1701.06599, 2016

[2] **X. Wang**, L. Lu, H.-C. Shin, L. Kim, I. Nogues, J. Yao, R. Summers. Unsupervised Category Discovery via Looped Deep Pseudo-Task Optimization Using a Large Scale Radiology Image Database. arXiv:1603.07965, 2016

[1] **X. Wang**. Automatic Quality Improvement of Archive Film. Ph.D. Thesis, University of Bristol, UK, 2011

R&D PROJECTS

- **Automatic Detection and Reporting for Common Thorax Diseases (2017- Ongoing)** By utilizing the constructed ChestX-ray14 dataset, we try to achieve better classification and localization performance using not only the images but also the raw text report. We create an end-to-end trainable framework by introducing multi-level attentions into the model and combining the CNN and RNN into a unified architecture. Given an image, the system can output detected disease keywords, generated report and location info of detected disease pattern in one go.

- **DeepLesion: Large-Scale Clinical Lesion Annotation and Detection (2017- Ongoing)** Vast amounts of clinical annotations (usually associated with disease image findings and marked using arrows, lines, lesion diameters, segmentation, etc.) have been collected over several decades and stored in hospitals' Picture Archiving and Communication Systems. We mine and harvest one major type of clinical annotation data - lesion diameters annotated on bookmarked images - to learn an effective multi-class lesion detector via unsupervised and supervised deep Convolutional Neural Networks (CNN). Our dataset is composed of 33,688 bookmarked radiology images from 10,825 studies of 4,477 unique patients.

- **BIOPSY-GUIDED MRI PROSTATE CANCER DETECTION (2016-2017)** Prostate Cancer is highly prevalent and is the second most common cause of cancer-related deaths in men. Multiparametric MRI is robust in detecting PCa. We developed a weakly supervised computer-aided detection system that uses biopsy points to learn to identify PCa on mpMRI. Our CAD system, which is based on a deep convolutional neural network architecture, yielded an AUC of 0.903±0.009 on a receiver operation characteristic curve computed on 10 different models in a 10 fold cross-validation.

- **Weakly-Supervised Classification and Localization of Common Thorax Diseases (2016-2017)** Construct a new chest X-ray database, namely "ChestX-ray8", which comprises 108,948 frontal-view X-ray images of 32,717 unique patients with the text-mined eight disease image labels (where each image can have multi-labels), from the associated radiological reports using natural language processing. Demonstrate that these commonly occurring thoracic diseases can be detected and even spatially-located via a unified weakly-supervised multi-label image classification and disease localization framework.

- **Lymph Node Segmentation (2016)** Lymph node segmentation is an important yet challenging problem in medical image analysis. The presence of enlarged lymph nodes (LNs) signals the onset or progression of a malignant disease or infection. We presents a novel approach to TA LNC segmentation that combines holistically-nested neural networks (HNNs) and structured optimization (SO).

- **Unsupervised Image Category Discovery (2015-2016)** Unsupervised image categorization (i.e., without the ground-truth labeling) is critically important and difficult when annotations are extremely hard to obtain in the conventional way of "Google Search" and crowd sourcing. We address this problem by presenting a looped deep pseudo-task optimization (LDPO) framework for joint mining of deep CNN features and image labels. Our method is conceptually simple and rests upon the hypothesized "convergence" of better labels leading to better trained CNN models which in turn feed more discriminative image representations to facilitate more meaningful clusters/labels.

- **Clinical Validation of post-processing applications (2012 - 2015)** Design and develop methods for validating the post-processing workstation produced in HSW BU, e.g. blinded validation in clinical studies, internal quantitative evaluation of the accuracy of algorithms. Workstations cross all modality are certificated by CFDA.

- **MR Breast CAD (2012-2013)** Develop a CADx module for MR breast Mass, including an interactive mass segmentation tool and mass kinetic analysis tools.

- **CT Lung Nodule CAD (2012-2013)** Develop a CADe module for automatic CT lung Nodule detection, involving nodule enhancement, nodule segmentation, feature calculation / selection, LDA based nodule classification, for both solid and GGO nodules.

- **CT Colonoscopy (2012-2013)** Develop a tool for processing CT colonoscopy images, involving colon segmentation, cleansing, central line extraction, polyp segmentation, feature calculation/extraction, SVM based polyp classification

- **FFDM Mass/Calcification CAD (2013)** Develop a tool for automatic detection of mass and calcification in FFDM, involving breast segmentation, pectoralis removal, background suppression, mass/calcification segmentation, feature calculation/selection, LDA based mass/calcification classification

- **MI Cardiac Perfusion (2013)** Develop a tool for processing PET/SPECT Cardiac Perfusion images, involving SA view reconstruction, LV segmentation, Bull's eye polar map calculation.

- **Resting-state fMRI (2013)** Develop a tool for processing resting-state fMRI data with ICA and seeded region based analysis methods. Collaboration with Shanghai Xuhui District Hospital in a clinical migraine study (data mining the relationship between migraine and fMRI resting-state components)

- **DBT / CBCT / Synthesis 2D Image Reconstruction (2013)** Develop FBP and iterative algorithms for DBT/CBCT tomography reconstruction from projection data, including FBP, BPF, MLEM, etc.

- **CT Liver Perfusion (2012)** Develop a method for processing CT liver perfusion data, involving artery localization, bone removal, dual-input single-compartment modeling (kinetic analysis), computing functional maps.

- **MR BOLD (2012)** C++ implementation of SPM and an inline version for real-time scan controlling.

- **MR Auto-Windowing (2011)** Develop an automatic windowing method for MR images, involving feature calculation, SOM/adaptive k-means classification, RBF windowing training, online training.

- **Automatic Archive Film Restoration (2006-2011)** Develop a unified framework for automatic archive film restoration, which is composed of three parts, i.e. defect detection, false alarm elimination and defect removal. First, we propose a novel probabilistic approach to detect defects in digitized archive film, by combining temporal and spatial information across a number of frames. A two-stage false alarm elimination process is then applied on the resulting defect maps, comprising MRF modelling and localized feature tracking. Given the resulting defect maps, restoration is performed for defects and missing regions in archive films. The proposed statistical framework is based on random walks to examine the spatiotemporal path of a degraded pixel, and uses texture features in addition to intensity and motion information traditionally used in previous restoration works.

- **A SIFT Tool (2006)** The aim of this project is to produce a fully comprehensive software library which implements previous work by David G. Lowe on the subject of the Scale Invariant Feature Transform (SIFT). The SIFT library is critically evaluated and demonstrated by using images under different transformations and an object recognition system.