

Appendix: Content Summaries of Selected Best Papers for the 2019 IMIA Yearbook, Section Sensors, Signals, and Imaging Informatics

Bandeira Diniz JO, Bandeira Diniz PH, Azevedo Valente TL, Corrêa Silva A, de Paiva AC, Gattass M

Detection of mass regions in mammograms by bilateral analysis adapted to breast density using similarity indexes and convolutional neural networks

Comput Methods Programs Biomed 2018 Mar;156:191-207

Mammographic imaging is a critical tool for detecting breast cancers early. However, 13% of cancers go undetected. Computational techniques are being explored to improve the early detection rate of suspicious findings associated with breast cancer. In this paper, the authors describe a computer-aided diagnosis methodology to assist breast radiologists with identifying mass regions in 2D mammographic images. The approach consists of several modules: a preprocessing step (*e.g.*, resizing, cropping), a density classification model (*i.e.*, classifying dense versus non-dense breasts), an image and segmentation step, the selection of candidate masses, and a final classifier to identify actual masses. Separate final classifiers are trained for dense and non-dense breast images. In addition, the authors utilize bilateral analysis to identify differences between contralateral breasts as a way to assist in identifying candidate masses. Their system was trained using a large public dataset called Digital Database for Screening Mammography, which contains more than 2,500 labeled exams of digitized film mammograms. They showed that the method achieved 91% accuracy in classifying masses in non-dense breast tissue and 95% accuracy in detecting masses in dense breasts. This work demonstrates the utility of incorporating domain knowledge such as breast density and bilateral analysis to improve the performance of mass detection. The authors provide a comprehensive expla-

nation of their methodology and comparison of their approach's performance with other state-of-the-art methods.

Lee H, Yune S, Mansouri M, Ki M, Tajmir SH, Guerrier CE, Ebert SA, Pomerantz SR, Romero JM, Kamalian S, Gonzalez RG, Lev MH, Do S

An explainable deep-learning algorithm for the detection of acute intracranial hemorrhage from small datasets

Nat Biomed Eng 2019 Mar;3(3):173-82

The rapid development of machine and deep learning algorithms have yielded systems that are capable of improving diagnostic accuracy and optimizing the delivery of healthcare. However, a major hurdle towards adoption is the increasing difficulty in fully explaining or understanding the basis of these models' predictions. The authors demonstrate an approach that assists human users with understanding the basis of a model's prediction using a combination of heatmap-based visualizations. The overarching goal of their work is to automatically predict whether intracranial hemorrhage (ICH) is present in a non-contrast head computed tomography (CT) study and if so, classify its subtype. The optimal model architecture was investigated by using multiple pre-trained deep convolutional neural network architectures (VGG16, ResNet-50, Inception-v3, Inception-ResNet-v2) and various data preprocessing techniques. The authors used the class activation mapping technique to highlight important regions of the image for the prediction to a target label. They also generated a radiology atlas of ICH by ranking all activation maps at all blocks for each label based on relevance count and selecting the top 5% of them to best represent the image assigned with the label. The authors had access to an imbalanced dataset of 625 ICH-positive and 279 ICH-negative cases of which 100 positive and 100 negative cases were set aside as a validation set. A retrospective dataset of 100 ICH-positive and 100 ICH-negative cases and a prospective dataset of 79 ICH-positive and 117 ICH-negative cases collected consecutively at a single institution were used to evaluate model performance. Model performance was compared against five radiologists per

dataset (three residents and two board-certified attendings). The system achieved similar performance to that of radiologists in both datasets, achieving a sensitivity of 98% and 92% in the retrospective and prospective datasets, respectively. The work is notable for its effective use of heatmaps to visualize model regions and its comprehensive evaluation of model performance across datasets and against human readers.

Samad MD, Ulloa A, Wehner GJ, Jing L, Hartzel D, Good CW, Williams BA, Haggerty CM, Fornwalt BK

Predicting survival from large echocardiography and electronic health record datasets: optimization with machine learning

JACC Cardiovasc Imaging 2019 Apr;12(4):681-9

Over 10 million echocardiograms are performed annually in United States Medicare patients alone. When interpreted alongside patient demographic information, these imaging data are useful in diagnosing and prognosticating a patient's status. However, physicians face the challenge of relating the sheer number of measurements that are generated from echocardiography to the dozens of plausible diagnoses that could be coded from these images. In this paper, the authors present an analytical framework for predicting one- and five-year mortality survival using electronic health record data, including echocardiograms, from a large patient population seen at their institution. Their approach combines clinical variables (age, sex, height, weight, heart rate, blood pressure, cholesterol, smoking status, and 90 ICD-10 codes (International Classification of Diseases 10th revision)) along with physician-reported left ventricular ejection fraction, and 516 different echocardiographic measurements. Missing data were imputed using multivariate imputation by chained equations (MICE). Linear (logistic regression) and nonlinear (random forest) models were trained and evaluated using a 10-fold nested cross-validation design. The authors had access to data from over 171,510 patients and 331,317 electrocardiograms. Machine learning models achieved higher prediction accuracy (AUC

> 0.82) over common clinical risk scores such as Framingham (AUC between 0.61 to 0.79). The nonlinear model outperformed the linear model. The most common informative features across all predictions were found to be age, tricuspid regurgitation jet maximum velocity, heart rate, and left ventricular ejection fraction. This work demonstrated the feasibility of harnessing a large clinical population and combining clinical and imaging information to significantly improve the performance of mortality predictions.

Vasilakakis MD, Iakovidis DK, Spyrou E, Koulaouzidis A

DINOSARC: color features based on selective aggregation of chromatic image components for wireless capsule endoscopy
Comput Math Methods Med 2018 Sep 3;2018:2026962

Wireless Capsule Endoscopy (WCE) is a swallowable pill containing a miniaturized camera that generates hundreds of thousands of color images of the digestive tract. While an innovative and non-invasive way to detect abnormalities such as ulcers, polyps, and bleeding, interpreting WCE images is time-consuming and error-prone. Salient point detection, an unsupervised process of extracting image features that are most associated with abnormalities, is increasingly being used to facilitate discrimination between normal and abnormal WCE images. In this study, the authors developed a salient point and region detection algorithm to estimate local and global image descriptors that are predictive of abnormalities. The algorithm consists of several components including a color-based salient point detector (based on a narrow color range that is usually located on the margins of the overall color range of an image), a salient region detector based on

superpixels (using the simple iterative linear clustering algorithm), and a generator of local and global color image descriptors that are extracted from the superpixel regions. A public dataset previously released by the authors consisting of images obtained from a MiroCam endoscope with 360x360 pixels was used to evaluate the approach. The developed algorithm, DINOSARC, achieved the highest percentage of true positive salient points over other state-of-the-art algorithms. It also demonstrated the highest overall area under the curve using its generated local and global image descriptors to classify normal versus abnormal WCE images. This work is important in its focus on salient feature detection and the identification of local and global image descriptors in an efficient and unsupervised manner. It is also demonstrating how novel sensors, combined with appropriate imaging informatics, can be made applicable for a clinical setting.