

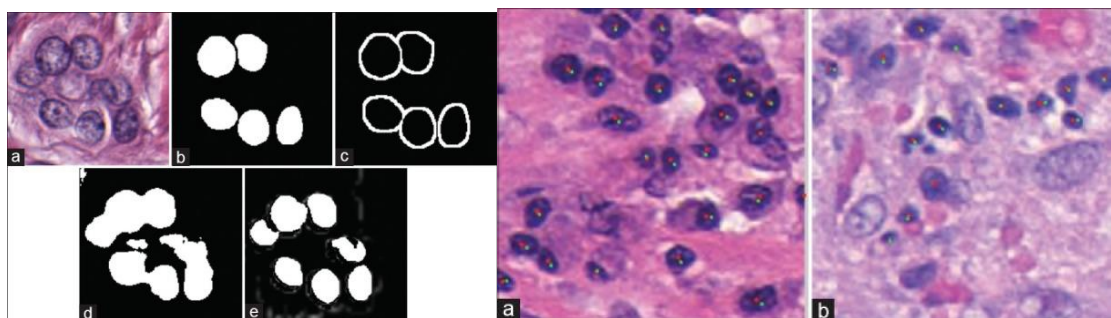
# AI-based Medical image analysis as an important tool for timely and accurate diagnostics

Blog post by Dejan Štepec ([XLAB](#))

Computer-aided diagnostic procedures have already become an important tool for the clinical routine, aiding clinicians at providing timely and accurate diagnostics results. The technological advances have made available various imaging modalities, which are opening the challenge of processing ever increasing volumes of imaging data. The increasing quantity and quality of the imaging data is opening an opportunity for more accurate diagnostics, but at the same time limiting the ability of the clinicians to directly process them, once again, due to increasing volumes and complexity of the imaging data.



Common image analysis tasks include classification, detection, and segmentation [1]. In the classification task, the algorithm aims to classify images into two or more classes. Examples include classification of lung nodules into benign or malignant or classification of pathology samples into different cancer types. In the detection task, the algorithm aims to localize structures in 2D or 3D space. For example, detecting lung nodules or liver metastases on CT scans or detecting cancerous regions in histology images. In the segmentation task, the algorithm tries to provide a pixel-wise delineation of an organ or pathology. For example, segmentation of the surface of lungs, a kidney, a spleen or tumors on CT, ultrasound, MRI or histology images (Figure 1).



(a) Nuclei segmentation

(b) Lymphocyte detection

Figure 1: Applications of AI based methods in the digital pathology domain [2].

Recently, with the advent of AI and, in particular, deep-learning techniques, there has been a significant methodological shift in how images (including medical images) are processed (Figure 2). Simple, manually engineered techniques for edge detection and feature extraction were replaced with data-driven, deep-learning approaches. Through feature learning instead of feature engineering, whereby systems automatically learn the representations needed for feature detection or classification from labelled training data, deep neural networks promise faster and more accurate results.

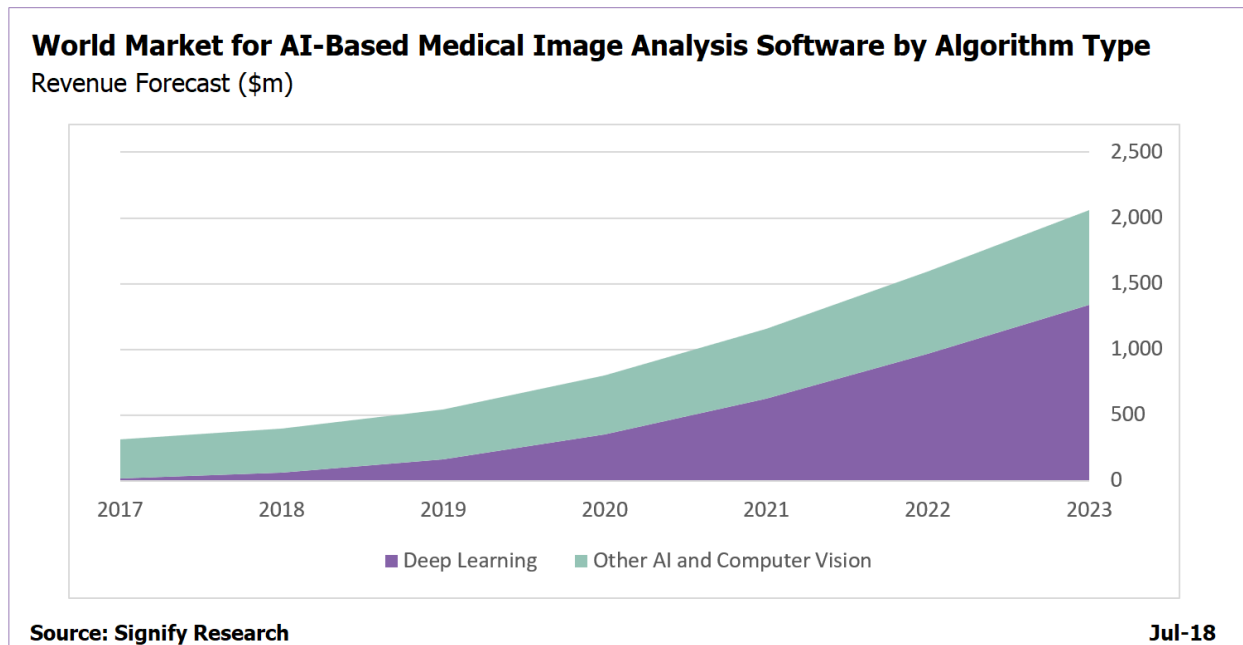


Figure 2: World market for AI-based medical image analysis solutions [5].

Deep-learning approaches require abundance of labeled data and are trained in a supervised fashion. Labeled data represents the original raw data, associated with metadata (e.g. tumor segmentation masks, cancer types, regions of interest), that is usually obtained manually by human experts. Labeled data usually represents a precondition for successful utilization of state-of-the-art AI-based methods. In the domains where labeled data can be obtained using automatic techniques (e.g. extraction of labels from different clinical reports using text mining techniques), deep-learning techniques have flourished and are performing with human-level performance. One such particular example are chest x-ray images. For this data type, large labeled datasets exist (e.g. CheXpert [6]) and, for these datasets, human-level performance was achieved for the detection of selected 5 pathologies (Figure 3).

Digital pathology is another domain where digitalization of the clinical process has opened the opportunity to apply machine learning techniques. In 2016, a competition was held [7] where automated solutions for detecting lymph nodes metastases were developed and compared against pathologists in terms of performance [4]. In comparison with classification tasks, where labeled data can be obtained (semi-) automatically from the reports, detection and segmentation tasks require enormous manual labour. Obtaining such detailed labels to learn supervised models is a costly and, in many cases, also an impossible process, due to the unknown set of all the disease biomarkers. Because of that, weakly-supervised and unsupervised approaches are getting traction. In the case of weakly-supervised approaches, labeled data for the detection

and segmentation tasks is only needed at the image level. For the unsupervised methods, there is no need for no labelled data at all.

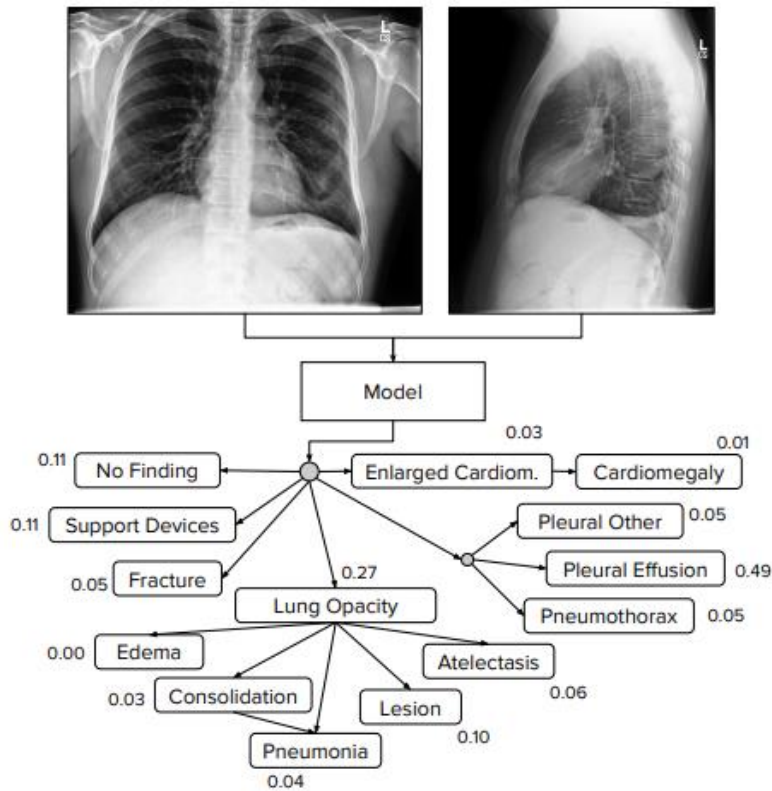


Figure 3: The CheXpert task is to predict the probability of different observations from multi-view chest radiographs [3].

Machine learning itself is just the enabler whereas compelling use-cases where AI can be shown to improve clinical outcomes and deliver a clear return of investment are the actual adoption drivers. Early AI-based methods have failed to take-off due to poor performance, but with the prevalence of deep-learning techniques and their superior performance and industrial adoption in different domains, AI has opened a completely new look. In some domains, considerations have been made to use such methods for the primary read applications, replacing the current dogma of having such tools solely as a “second pair of eyes”, thus completely automating certain diagnostic procedures.

Clinical Speciality	Key Uses for Machine Learning	Machine Learning Readiness
Radiology	Workflow efficiency and clinical decision support	High
Dermatology	Detect skin cancer and diagnose malignancy of moles	Low
Ophthalmology	Detect and diagnose disease, such as diabetic retinopathy	High
Pathology	Detect and diagnose disease from tissue samples	Low
Oncology	Cancer diagnosis, prognosis and treatment planning	Mid
Surgical	Endoscopic image analysis	Low

Machine learning readiness by imaging speciality.

Source: Signify Research [8]

In parallel with recent advancements in the research community, the first wave of commercialized deep-learning-based solutions are now entering the market and gaining initial momentum and acceptance in different clinical applications where they reduce manual time-consuming burden while increasing productivity and accuracy of different medical diagnostic procedures.

In [IPC](#), we are advancing the field of unsupervised anomaly detection, which is particularly suitable for the domains with limited amounts of data and, especially, labelled data. Such methods could significantly advance diagnostic procedures of paediatric cancers where data is of limited quantity, thus preventing the adoption of recent advancements in the broader medical image analysis domain.

#### References:

- [1] Lee, June-Goo, et al. "Deep learning in medical imaging: general overview." Korean journal of radiology 18.4 (2017): 570-584.
- [2] Janowczyk, Andrew, and Anant Madabhushi. "Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases." Journal of pathology informatics 7 (2016)
- [3] Irvin, Jeremy, et al. "Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 33. 2019.
- [4] Bejnordi, Babak Ehteshami, et al. "Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer." Jama 318.22 (2017): 2199-2210.
- [5] <https://www.signifyresearch.net/medical-imaging/ai-medical-imaging-top-2-billion-2023/>
- [6] <https://stanfordmlgroup.github.io/competitions/chexpert/>
- [7] <https://camelyon16.grand-challenge.org/>
- [8] <https://www.purestorage.com/content/dam/purestorage/pdf/whitepapers/signify-ai-in-medical-imaging-wp.pdf>