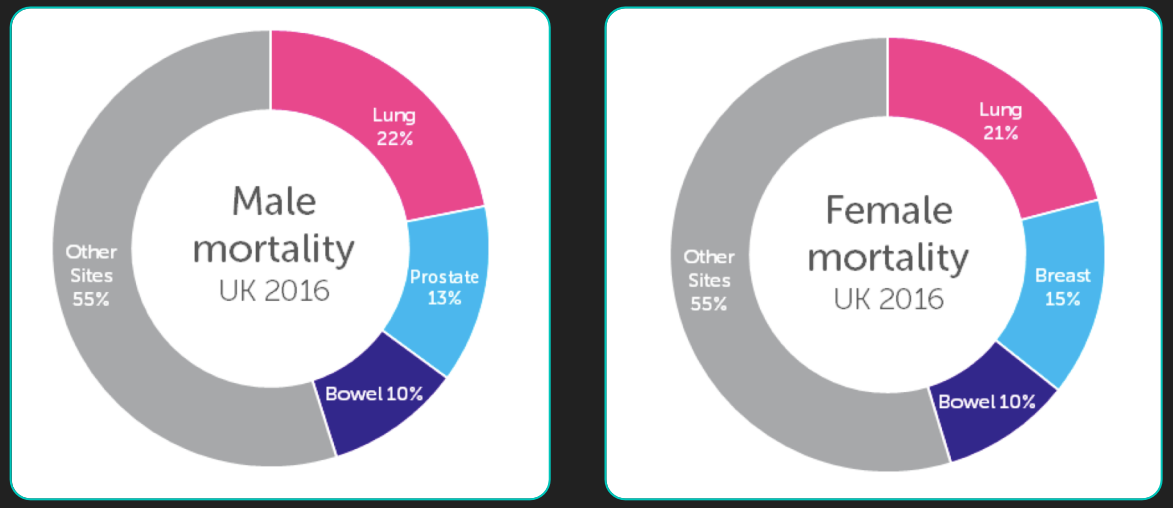
Review of deep learning models for cancer detection in computed tomography images

Dominika Bakalarz

University of Oxford, Hilary Term 2019

Lung cancer



(a) (b)

Fig.1. The most common causes of cancer deaths for men (a) and women (b). Source [1].

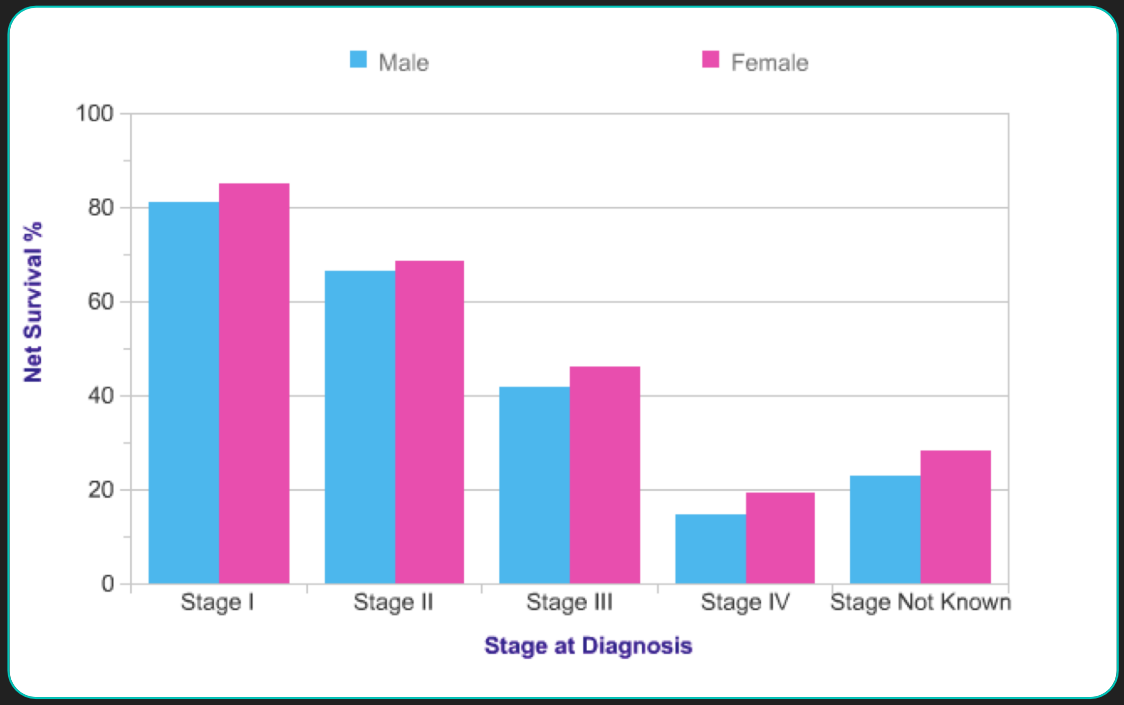


Fig. 2 1-year survival rate among patients diagnosed with lung cancer. Source [2].

Computed Tomography (CT)

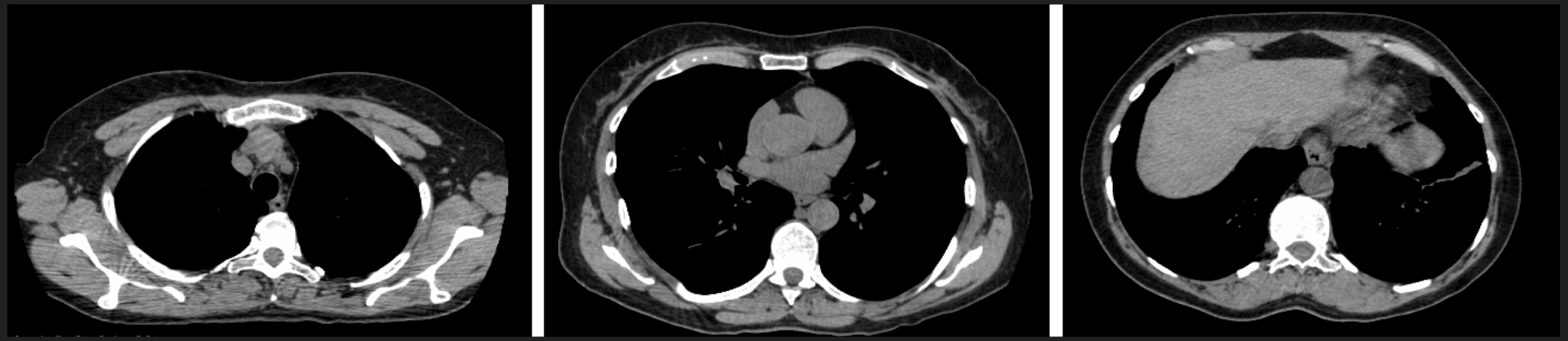


Fig.3 Sample images from LIDC database

- **CAD** (Computer-Aided Diagnosis) systems
- Lung Image Database Consortium (**LIDC**) – 1018 cases

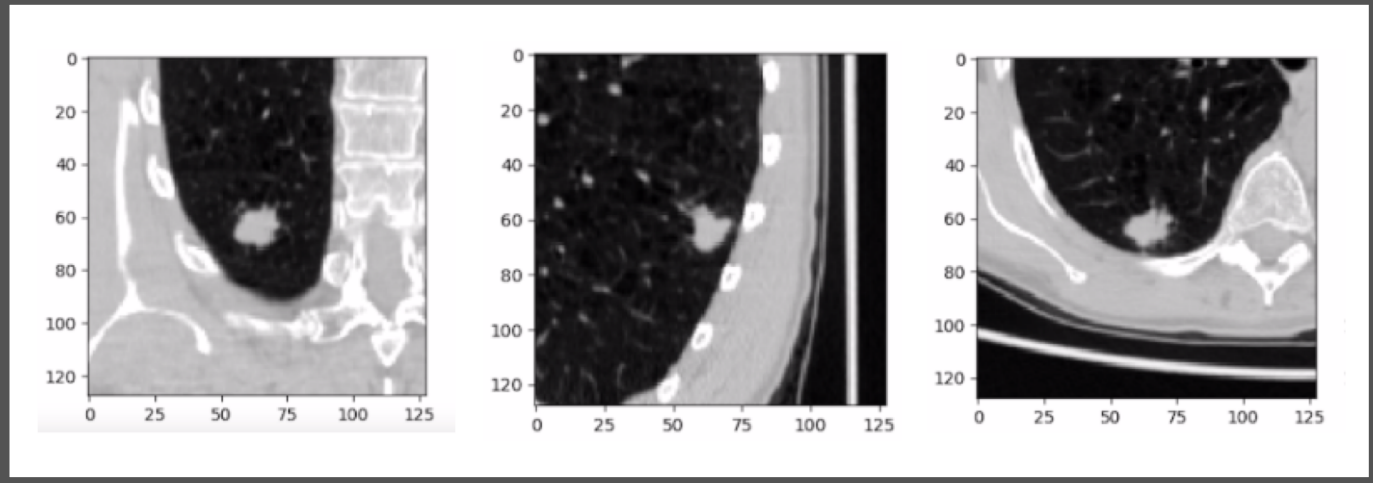


Fig.4 Sample malignant nodule (left - x-axis, centre - y-axis, right - z-axis). Source [3]

Deep learning methods

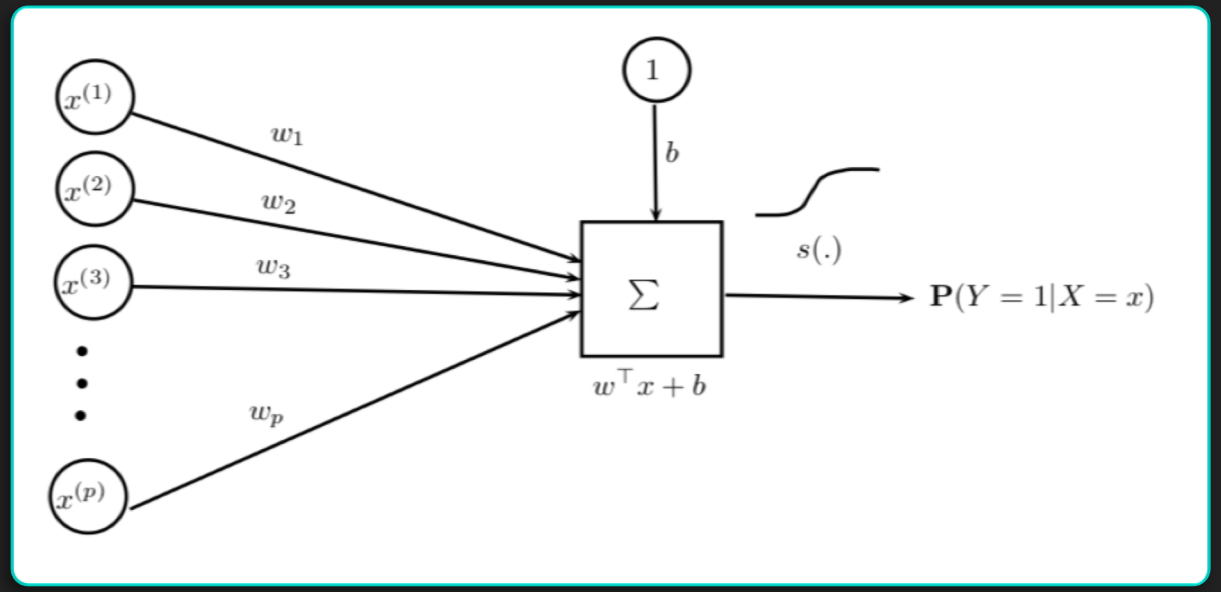


Fig. 5 Neural network with one neuron. Source [4]

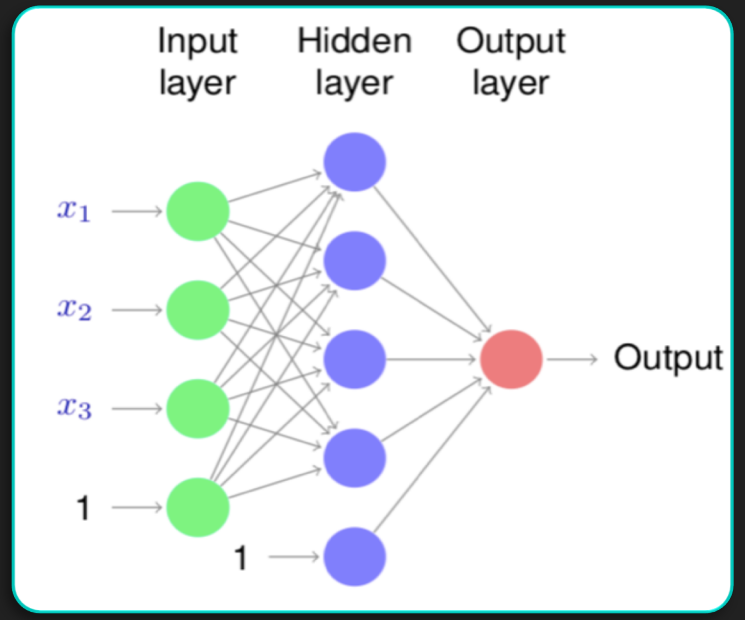
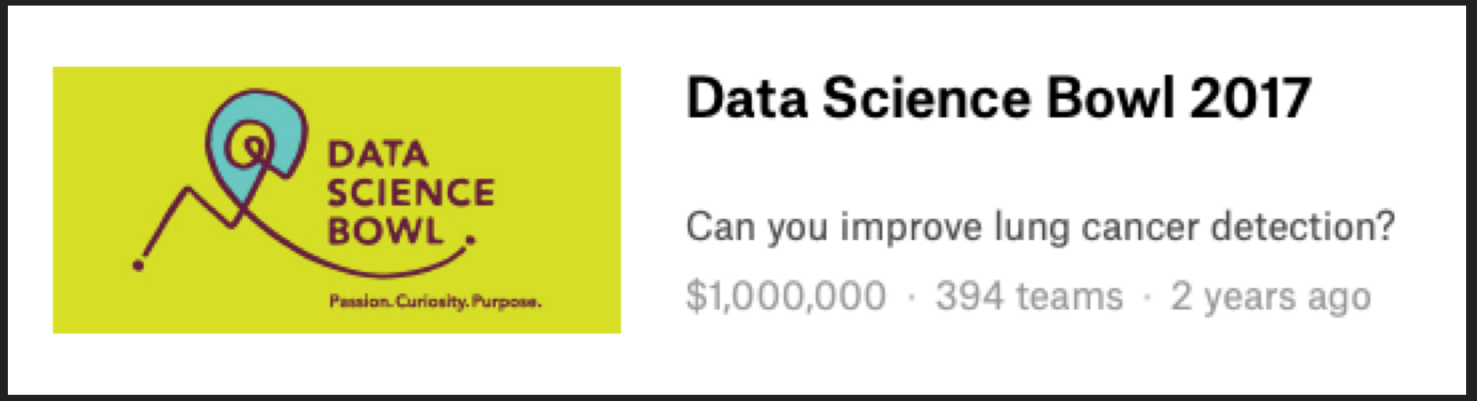


Fig. 6 Sample neural network. Source [4]

Data Science Bowl 2017

- **Challenge:** to predict the risk of lung cancer within one year from the CT screening
- **Training data:** CT scans, ground truth labels confirmed by pathology diagnosis



[8]

Data Science Bowl 2017

- "grt123" solution by Liao et al., **1st** place with a Log-Loss 0.39975,
- Solution by Julian de Wit and Daniel Hammack, **2nd** place with a Log-Loss 0.40117,
- "Deep Breath" solution by Verleysen et al., **9th** place and a Log-Loss 0.43872.

Solution by Verleysen et al., 9th place

Outline of the solution:

- 1. assign nodule probability to each voxel
- 2. identify group of high probability voxels as nodule candidates
- 3. reduce FP rate – expert network and complex lung segmentation

Solution by Julian de Wit and Daniel Hammack, 2nd place

- **Daniel Hammack -** 2-stage model:
- 1. nodule detection and feature prediction (4 features from LIDC data)
- 2. linear model for cancer prediction

- **Julian de Wit -** 1-stage model:
- nodule detection, malignancy estimation and cancer prediction with a tree model,
- No lung segmentation

Solution by Liao et al., 1st place

- 1. 3D region proposal network (RPN) for nodule detection
- 2. Leaky Noisy-OR network calculating cancer probability based on top five nodules

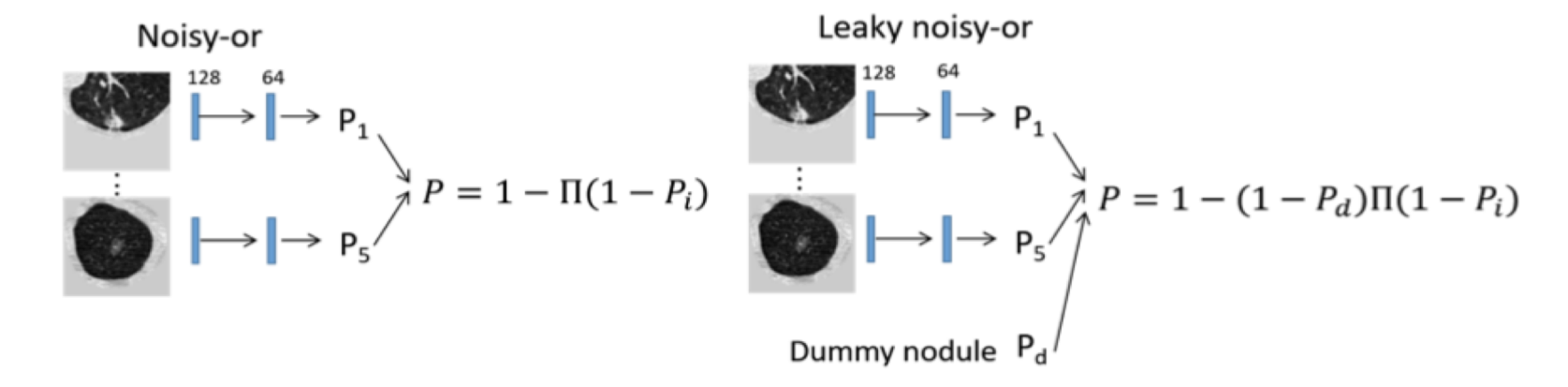


Fig 7. Noisy-or (left) and Leaky noisy-or (right) networks used by Liao et al. to calculate cancer probability. P_i - probability that i-th nodule is malignant, P – overall probability of cancer. Source [9]

State-of-the-art work

- Wang et al. (2018) achieved a sensitivity of 95.8% at 2 false positives per scan on the LIDC dataset in **nodule detection.** Comparison: LIDC radiologists achieved a sensitivity 55% - 79%, and 0.76 - 2.15 false positives per scan.

- Causey et al. (2018) achieved AUC of 0.99 in **nodule classification** task – result commensurate with experienced radiologists from LIDC cohort.

Commercial Applications

aidence
human sense in artificial intelligence [5]

veolity [6]

NUANCE [7]

Fig. 8 Veye Chest by Aidence. Source: [5]

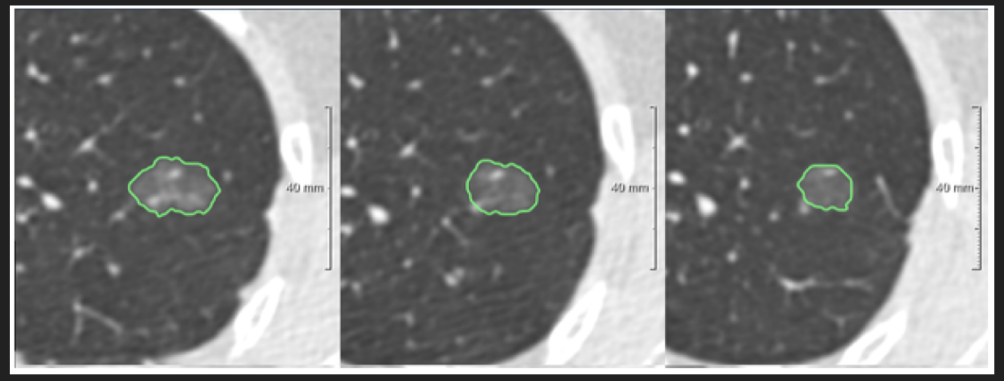


Fig. 9 One-click nodule segmentation – feature of Veolity product. Source: [6]

Future Challenges

- **Model interpretability:** LIME, Sharpley Values
- **Data limitation:** a need for bigger datasets
- **Data protection**

Conclusions

- Deep learning models are very promising methods in the field of cancer diagnosis
- State-of-the-art models achieve accuracy commensurate with experienced radiologists
- Despite several challenges, we expect such models to be more and more popular in the upcoming years
- However, it is still unclear if these models could ever completely substitute radiologists

Image Sources

- [1] https://www.cancerresearchuk.org/health-professional/cancer-statistics/mortality/common-cancers-compared#heading-One
- [2] https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/lung-cancer/survival#heading-Three
- [3] https://eliasvansteenkiste.github.io/machine\%20learning/lung-cancer-pred/
- [4] http://www.stats.ox.ac.uk/~palamara/teaching/SML19/
- [5] https://aidence.com/our-solution
- [6] https://www.veolity.com/about-veolity/product-information/
- [7] https://www.nuance.com/healthcare/diagnostics-solutions/ai-marketplace.html
- [8] https://www.kaggle.com/c/data-science-bowl-2017/
- [9] F. Liao, M. Liang, Z. Li, X. Hu, and S. Song. Evaluate the Malignancy of Pulmonary Nodules Using the 3D Deep Leaky Noisy-or Network. arXiv preprint, arXiv:1711.08324, 2017.