



Data fusion for better decision making

PhD student JAROSLAVA ARSENJEVA

Review for 2019/2020 study year

Thesis advisor: Viktor Medvedev, Dr.

Thesis consultant: Gintautas Dzemyda, Prof. Habil. Dr.

Informatics engineering field (T007)



Institute of Data Science
and Digital Technologies

The time of doctorate studies: 2018 – 2022 years



- Object of research: Data fusion methods.
- Research goal: To propose a data fusion algorithm (framework) to improve decision making in medicine using data mining methods.



Institute of Data Science
and Digital Technologies



Research tasks

- Perform an analytical review of data fusion existing methods;
- Select appropriate data for further;
- Identify main scientific problems (for example, in the medical field) in data fusion tasks;
- Select appropriate data fusion methods for previously selected data;
- Develop a data fusion algorithm to improve decision making;
- Evaluate the results; make necessary changes



Institute of Data Science
and Digital Technologies

Planned results



- Identification of proper data fusion methods to improve decision making
- Proposition and further application of a data fusion algorithm (framework), that would improve decision making using data obtained from mining techniques



Institute of Data Science
and Digital Technologies

Full work plan



Studijų metai	Egzaminai ¹		Dalyvavimas konferencijose ²		Publikacijos ³		
	Planas	Įvykdyta	Planas	Įvykdyta	Planas	Įvykdyta	Būklė ⁴
I (2018/2019)	2	2					
II (2019/2020)	2	2	1	1			
III (2020/2021)			1				
IV (2021/2022)			1		2		



Institute of Data Science
and Digital Technologies

Last semester work plan



Egzaminai		Dalyvavimas konferencijose		Publikacijos	
Planas	Įvykdyta	Planas	Įvykdyta	Planas	Įvykdyta
Visi egzaminai išlaikyti		Konferencijos pavadinimas	Suplanuotos ar kitos konferencijos pavadinimas	Žurnalo pavadinimas: Baltic Journal of Modern Computing	
				Pocė, Ingrida; Arsenjeva, Jaroslava; Kielaitė-Gulla, Aistė; Samuilis, Artūras; Strupas, Kęstutis; Dzemyda, Gintautas. Pancreas segmentation in CT images: state of the art in clinical practice // Baltic journal of modern computing. Riga : University of Latvia. ISSN 2255-8942. eISSN 2255-8950. 2021, vol. 9, no. 1, p. 25-34. DOI: 10.22364/bjmc.2021.9.1.02.	



Institute of Data Science
and Digital Technologies



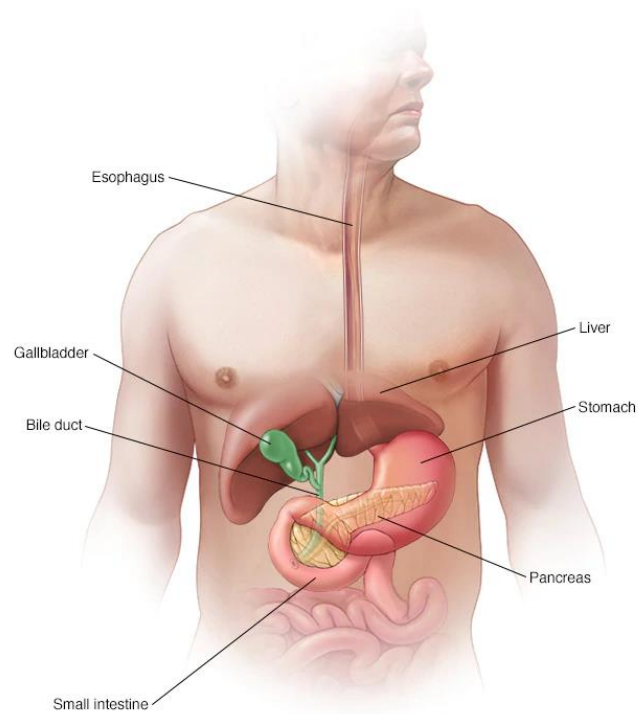
Next semester work plan

- Publish a paper in a peer-reviewed journal
- Participate in a conference
- Start preparing parts of the dissertation



Institute of Data Science
and Digital Technologies

Pancreatic cancer detection



© MAYO FOUNDATION FOR MEDICAL EDUCATION AND RESEARCH. ALL RIGHTS RESERVED.



Data available (portal venous phase)

- 20 cancerous and 17 healthy CTs from Santariskes
- Respectively 430 and 1119 slices
- Healthy dataset: 80 patients from CancerImagingArchive (18942 slices)
- Cancerous dataset: 281 people for Medical segmentation Decathlon (8796 slices)

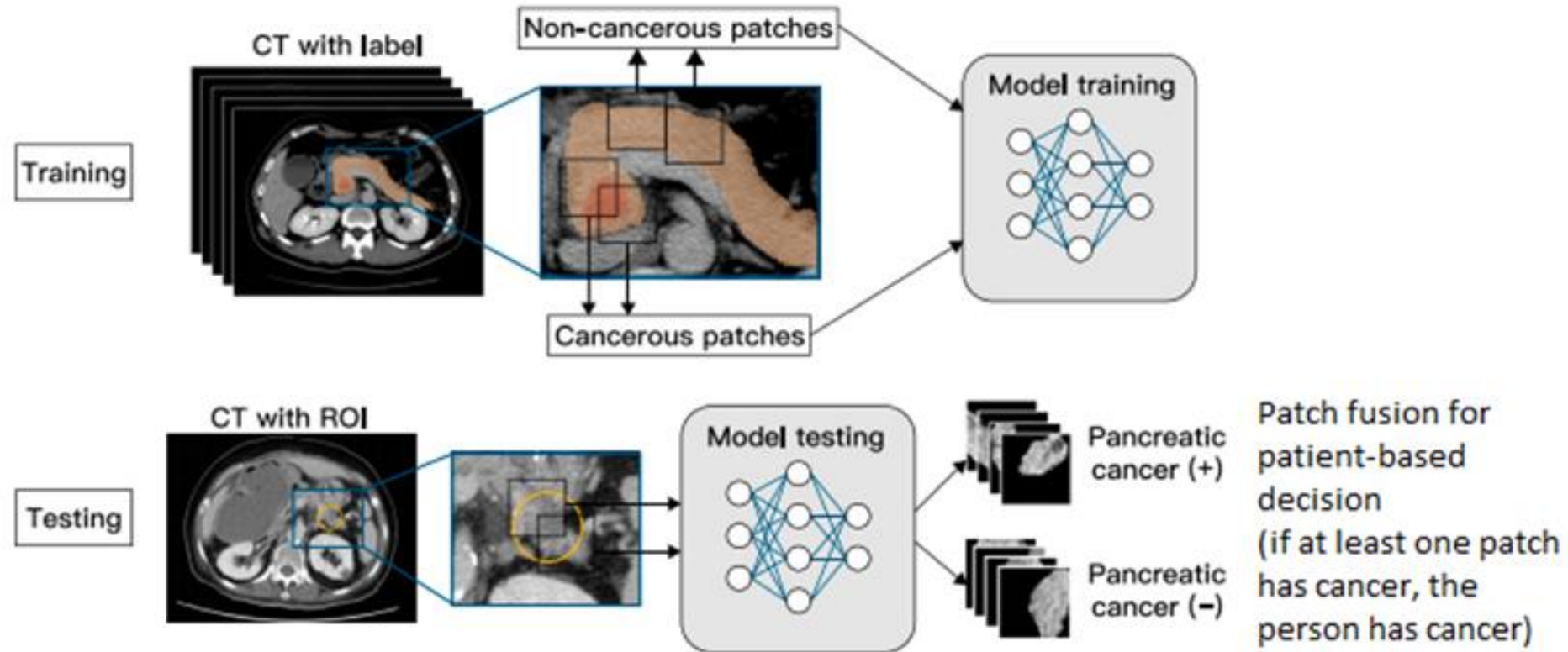
Results of CNN binary classification

- Training and validating on Santariskes data – 60% accuracy
- Training on mentioned datasets and validating on Santariskes data – 72% accuracy
- Structure: 2d layer (128 (3, 3), relu) + one more 2d layer (128, (3, 3), relu) + 1 dense layer (activation: sigmoid)
- Used Tensorflow, python version 3.7
- Images resized to 50x50 from original 512x512

Future work

- Pancreases sementation and splitting into 'patches' (from paper Deep learning to distinguish pancreatic cancer tissue from non-cancerous pancreatic tissue: a retrospective study with cross-racial external validation)
- Data fusion using patient statistics

Future work



Bibliography



- <https://www.mayoclinic.org/diseases-conditions/pancreatic-cancer/symptoms-causes/syc-20355421>
- <https://newsroom.uw.edu/news/scientists-activate-immune-cells-attack-pancreatic-cancer>
- <https://wiki.cancerimagingarchive.net/display/Public/Pancreas-CT#225140400c26eab54502412cbbd0e1c0fddd917b>
- <https://drive.google.com/drive/folders/1HqEgzS8BV2c7xYNrZdEAnrHk7osJJ--2>
- [https://doi.org/10.1016/S2589-7500\(20\)30078-9](https://doi.org/10.1016/S2589-7500(20)30078-9)



Institute of Data Science
and Digital Technologies