

## DEEP LEARNING IN BIG MEDICAL IMAGE DATA



**V. KOVALEV, PhD**  
*Head of the Laboratory of Biomedical Images Analysis*



**V. LIAUCHUK**  
*Research Assistant of the Laboratory of Biomedical Images Analysis*



**A. KALINOVSKY**  
*Research Officer of the Laboratory of Biomedical Images Analysis*



**E. SNEZHKO**  
*Senior research associate of laboratory of Mathematical cybernetics*



**A. TUZIKOV, DSc.**  
*General Director United Institute of Informatics Problems, National Academy of Sciences of Belarus, doctor of physical and mathematical sciences, professor, Corresponding member of the National Academy of Sciences of Belarus on "Informatics in Medicine and Biology"*

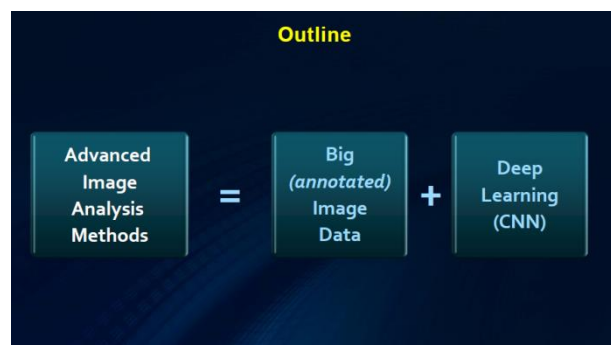
*Biomedical Image Analysis Department, United Institute of Informatics Problems, National Academy of Sciences of Belarus, Republic of Belarus  
E-mail: vassili.kovalev@gmail.com*

**Deep Learning in Big Medical Image Data**

*Kovalev V.A., Kalinovsky A.A., Liauchuk V.A., Snezhko E.V., and Tuzikov A.V.*

United Institute of Informatics Problems  
Belarus National Academy of Sciences

vassili.kovalev@gmail.com



### Medical Image Data: Exponential Growth is also observed

The slide shows a collection of diverse medical images on the left, including CT scans of the chest, MRI slices of the brain, and various histology slides. An arrow points from this collection to a line graph on the right, which illustrates exponential growth over time.

### Big Medical Image Data: When and Why they are "Big" ?

**when**  $10^2 - 10^3$  No need for new Methods

**when**  $> 10^4$  New Methods and Software are necessary

**why** 3D • Moving from 2D to 3D (tomography)

**why** Very Large 2D • e.g. Emergence of Whole Slide images (~10 G Pixels)

**why** Frequent Scanning • Spread of non-invasive imaging techniques

### Medical Images: Examples & Tasks

- Cell image analysis
- 3D image anisotropy
- Brain asymmetry
- SPECT: asymmetry as AD marker
- Generalized gradient: Detecting possible basis of Lung Tumors
- Biomedical Image Retrieval

### Medical Images: More examples ...

A grid of 20 small thumbnail images, each representing a different medical image analysis task or application, such as tumor segmentation, organ detection, and disease classification.

### Diagnosis: Searching for similar cases (Melanoma)

The slide displays a skin lesion diagnosis interface. On the left, there are three rows of skin lesions labeled 'typical nevus', 'atypical nevus', and 'melanoma'. In the center, there is a search interface for similar cases. On the right, there is an ROC curve and a table of diagnostic accuracy for different methods.

Method	Area under ROC curve	Sensitivity	Specificity
Pixel-based	0.850	87.7%	93.2%
Search in database	0.92	80.1%	94.4%
Combination	0.94	87.5%	95.4%

### TB and other: Our prototype on-line services (unofficial site)

A screenshot of the website 'http://imlab.grid.by/'. The website features a navigation menu and several service buttons: CBIR/CT, ROI Search, X-Ray Segmentation, WebCam Search, Skin Lesion, Lung Segmentation, Drug Resistance, and Histology.

### Cancer diagnosis: Whole Slide histology images

The slide shows whole slide histology images and a diagram of a multi-resolution pyramid. The pyramid has three levels labeled '1X Magnification', '10X Magnification', and '100X Magnification'. Below the pyramid, there are several histology images at different magnifications.

International Competitions:

- CAMELYON-2016
- TUPAC-2016

EU-Funded project

### Cancer Diagnosis: Detection & measuring of Metastases (Data)

**Data:** Whole-slide images are generally stored in a multi-resolution pyramid structure. Case files contain multiple downsampled versions of the original image. Each image in the pyramid is stored as a series of tiles, to facilitate rapid retrieval of sub-regions of the image. Image resolution of "Level 0" is near 100,000x100,000.

**Training/Validation dataset:** Whole-slide images are generally stored in a multi-resolution pyramid structure. Image files contain multiple downsampled versions of the original image. Each image in the pyramid is stored as a series of tiles, to facilitate rapid retrieval of sub-regions of the image. Image resolution of "Level 0" is near 100,000x100,000.

A grid of histology images showing normal tissue and tumor tissue. The images are arranged in a grid with columns labeled 'Normal' and 'Tumor'. A small inset image shows a whole slide image with a tumor region highlighted.

### Cancer Diagnosis: Detection & measuring of Metastases (Pipeline)

ISBI challenge on cancer metastasis detection in lymph node  
Camelyon 2016

Our solution: Train Deep Learning Model, and classify whole slide histology image at "Level 0"

The diagram shows a pipeline starting with a whole slide histology image, which is processed by a 'Deep Learning Model' to produce a heatmap of detected metastases.

### Cancer Diagnosis: Estimating Proliferation Score (aggressiveness)

A person in a white lab coat is sitting at a desk with a computer workstation. The monitor displays a histology image. To the right of the person, there is a large histology image showing a proliferation score.

### Cancer Diagnosis: Estimating Proliferation Score (our results)

**Results**

Task 1: Prediction of the proliferation score based on mitosis counting  
The evaluation metric for this task is the quadratic weighted Cohen's kappa between the predicted and ground truth proliferation scores.

Automatic methods

Team	Score	Use of additional data
Land Inc., Korea	0.567	No
Contribution, Sweden (SLC&UFD BCR)	0.534	No
Medica, Sweden	0.462	Yes, negative ROI annotations
University of Heidelberg, Germany	0.417	No
MIT Research (Dutch and Greek)	0.365	Yes, ICFN 2012 and 2014 data
The Harrier School, United States	0.367	No
Radboud National Academy of Sciences	0.321	Yes, negative ROI annotations
Radboud UMC Nijmegen, The Netherlands	0.266	No
University of South Florida, United States	0.177	No
University of Warwick, United Kingdom	0.159	Yes, negative ROI annotations
Technical University of Munich	0.109	No

Task 2: Prediction of the proliferation score based on gene expression  
The evaluation metric for this task is the Spearman's correlation coefficient between the predicted and ground truth proliferation scores.

Automatic methods

Team	Score	Use of additional data
Land Inc., Korea	0.677	No
Radboud UMC Nijmegen, The Netherlands	0.516	No
Contribution, Sweden (SLC&UFD BCR)	0.503	No
Radboud National Academy of Sciences	0.488	Yes, negative ROI annotations
The Harrier School, United States	0.474	No

### IBM Open Power Challenge: 1st place (together with Altoros Development)

Distributed TensorFlow for cancer detection

### Chest X-Ray (~2 000 000): Benchmarking Convolutional Networks

Males, 21

Females, 70

### XRy (100 000 people): Image mining (past research)

Experience: Lung Shape Mining

### XRy (10 000 people): Conventional vs. Deep Learning (age prediction)

Conventional  
Mean Error ~ 13.5 years

Deep Learning  
Mean Error ~ 6.0 years

Age range: 50 years  
100 Males + 100 Females for each year

Deep Learning dramatically outperform the existing methods

### XRy (10 000 people): Conventional vs. Deep Learning (classification)

Classification (Conventional)  
AUC (G2 vs G3) = 0.79

Classification (GoogLeNet)  
AUC (G2 vs G3) = 0.95

### Deep Learning: Lung Cancer Diagnosis

Data Science Bowl 2017

Task: Using a data set of thousands of high-resolution lung scans provided by the National Cancer Institute, participants will develop algorithms that accurately determine when lesions in the lungs are cancerous.

Goal: is to reduce the false positive rate that plagues the current detection technology and get patients earlier access to life-saving operations.

### Deep Learning: Lung Cancer Diagnosis

Data

Train set: 1595 high resolution CT scans  
Test set: 506 high resolution CT scans  
Each case histologically verified and labeled for Lung Cancer (1 means LC, 0 = no LC)

Our approach

Conv. Neural Net

Probability of LC

0.9954

0.1942

### Deep Learning: Lung segmentation using Convolutional Networks

Encoder-decoder CNN

Image Segmentation Using Encoder-Decoder CNN

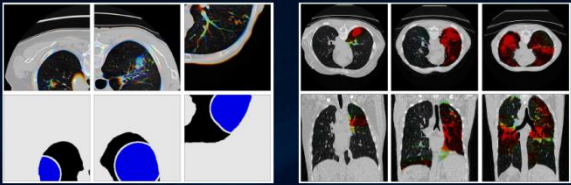
### Deep Learning: Lung segmentation using Convolutional Networks

Results with maximum Dice's score.

Results with minimum Dice's score.

Mean Dice score is 0.966 with STD ~ 0.04.

**Current: Lung Lesion Detection & Segmentation**



The image displays two sets of CT scan slices of the lungs. The left set, labeled 'Manual labeling (for Training set)', shows several slices with blue and red regions indicating manually segmented lesions. The right set, labeled 'Results obtained with Convolutional Neural Network', shows the same slices with the CNN's output, where lesions are highlighted in red and green, demonstrating high accuracy in detection and segmentation.

Manual labeling (for Training set)

Results obtained with Convolutional Neural Network

to be presented at CARS-2017 International Congress

**Conclusions**

**Pros:**  
In majority of Medical Image Analysis tasks the Deep Learning techniques outperform all the existing methods.

**Cons:**  
However, a large amount of *annotated* image data is necessary for training of Convolutional Neural Networks.

**Acknowledgments.** This work was funded by :

- ❖ National Academy of Sciences of Belarus,
- ❖ National Institutes of Health, U.S. Department of Health and Human Services, USA through the CRDF project OISE-16-62631-1, and
- ❖ Altoros Development (Minsk office).

Библиотека БГУИР