DEEP LEARNING IN BIG MEDICAL IMAGE DATA

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

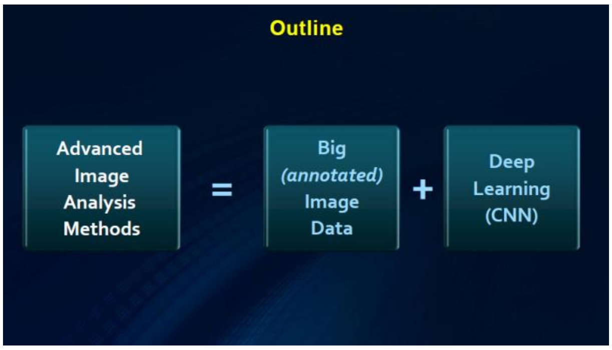
V. LIAUCHUK
Research Assistant of the Laboratory of Biomedical Images Analysis

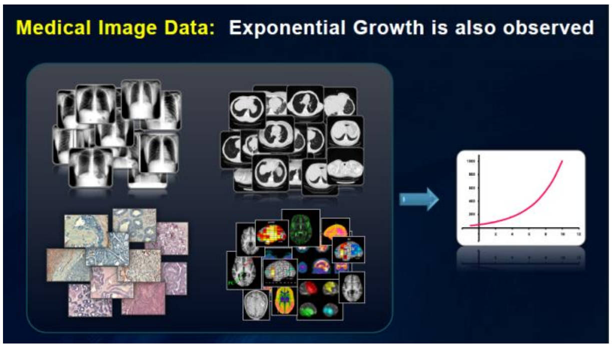
A. KALINOVSKY
Research Officer of the Laboratory of Biomedical Images Analysis

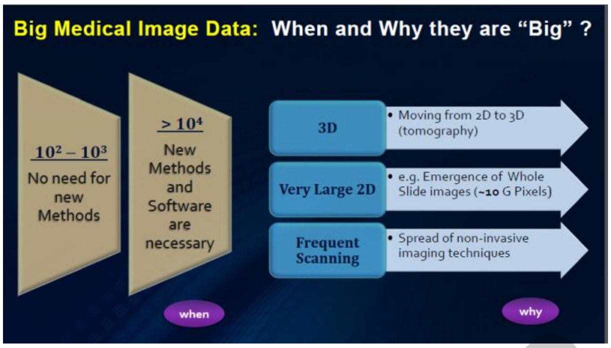
E. SNEZHKO
Senior research associate of laboratory of Mathematical cybernetics

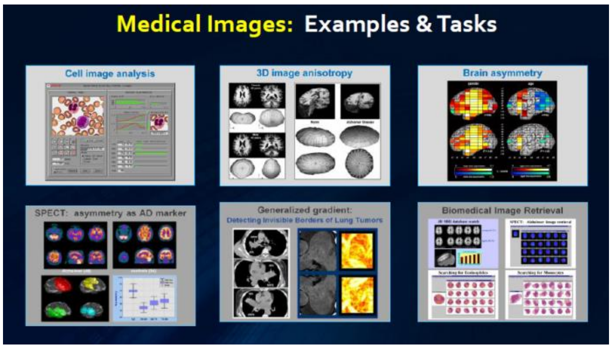
A. TUZIKOV, DSc.
General Director United Institute of Informatics Problems, National Academy of Sciences of Belarus, doctor of physical and mathematical sciences, professor, Corresponding member of the National Academy of Sciences of Belarus on "Informatics in Medicine and Biology"

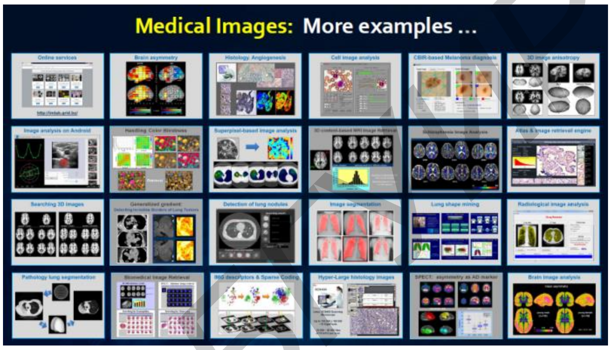
Biomedical Image Analysis Department, United Institute of Informatics Problems, National Academy of Sciences of Belarus, Republic of Belarus
E-mail: vassili.kovalev@gmail.com

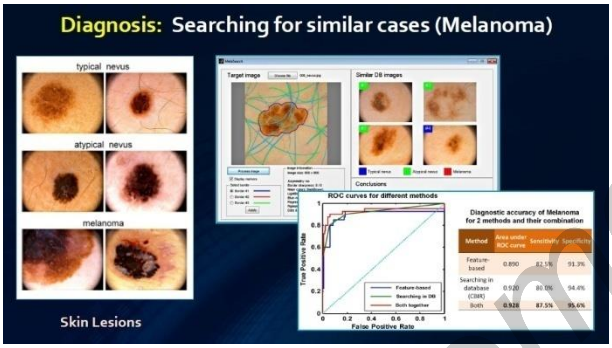


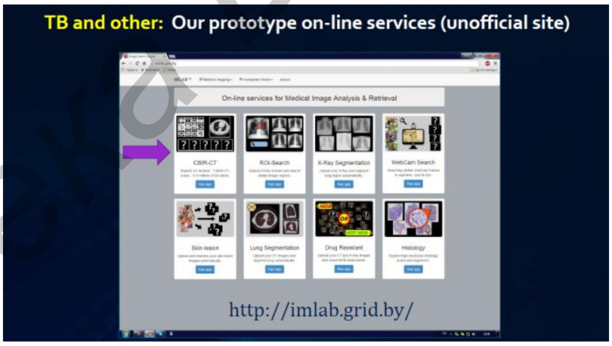

Medical Image Data: Exponential Growth is also observed

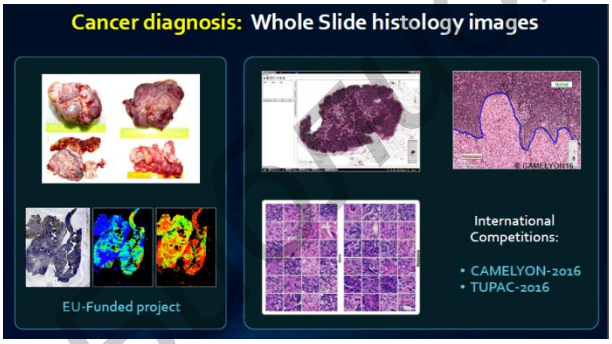

Big Medical Image Data: When and Why they are "Big" ?
10² – 10³
No need for new Methods
> 10⁴
New Methods and Software are necessary
3D
Moving from 2D to 3D (tomography)
Very Large 2D
e.g. Emergence of Whole Slide images (~10 G Pixels)
Frequent Scanning
Spread of non-invasive imaging techniques
when
why

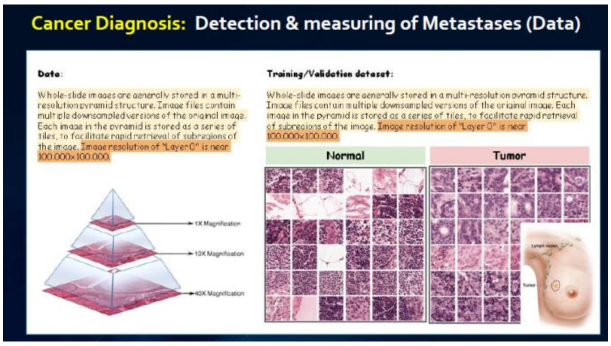

Medical Images: Examples & Tasks
Cell image analysis
3D image anisotropy
Brain asymmetry
SPECT: asymmetry as AD marker
Generalized gradient:
Detecting Invisible Borders of Lung Tumors
Biomedical Image Retrieval

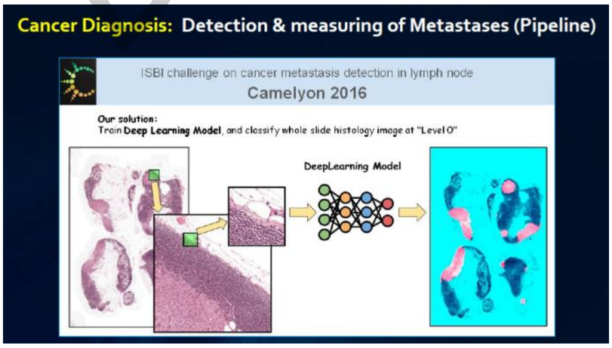

Medical Images: More examples ...

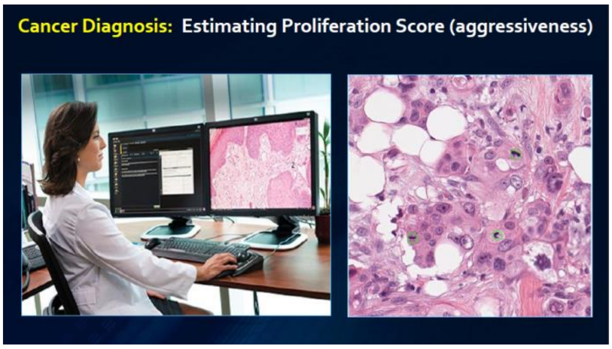

Diagnosis: Searching for similar cases (Melanoma)
typical nevus
atypical nevus
melanoma
Skin Lesions
ROC curves for different methods
Diagnostic accuracy of Melanoma for 2 methods and their combination
True Positive Rate
False Positive Rate

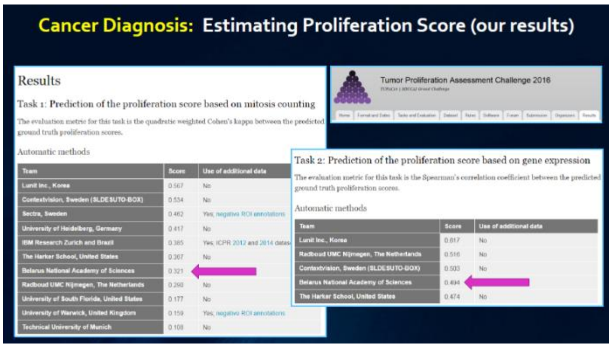

TB and other: Our prototype on-line services (unofficial site)
On-line services for Medical Image Analysis & Retrieval
http://imlab.grid.by/

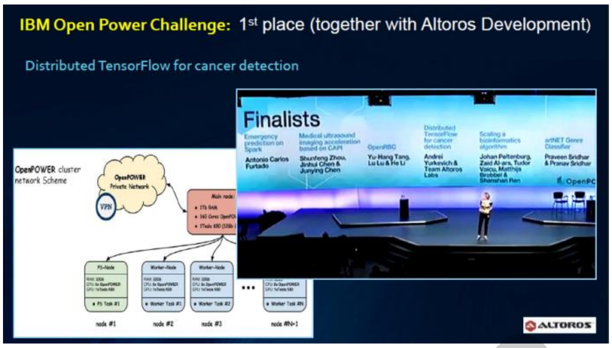

Cancer diagnosis: Whole Slide histology images
EU-Funded project
International Competitions:
• CAMELYON-2016
• TUPAC-2016

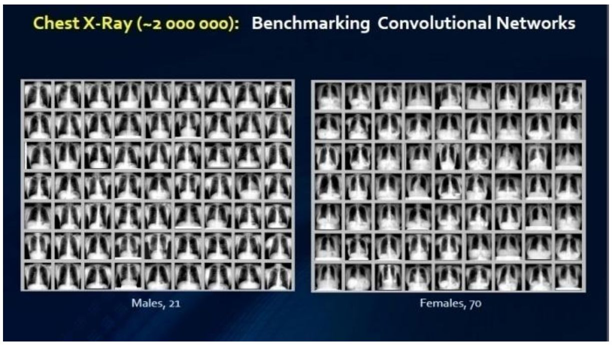

Cancer Diagnosis: Detection & measuring of Metastases (Data)
Data:
Training/Validation dataset:
Normal
Tumor

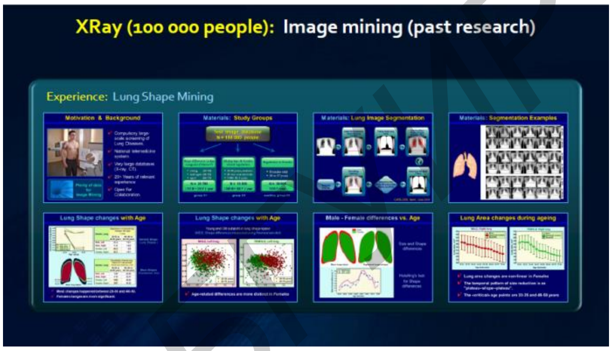

Cancer Diagnosis: Detection & measuring of Metastases (Pipeline)
ISBI challenge on cancer metastasis detection in lymph node
Camelyon 2016
Our solution:
Train Deep Learning Model, and classify whole slide histology image at "Level 0"
DeepLearning Model

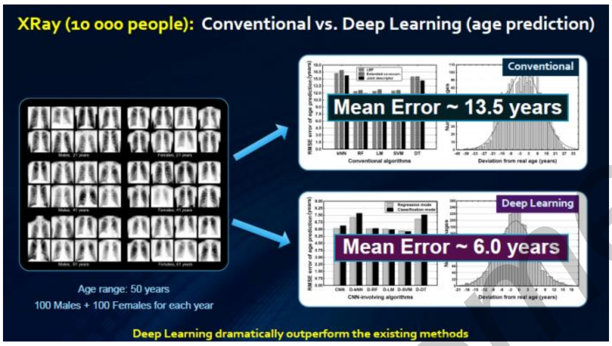

Cancer Diagnosis: Estimating Proliferation Score (aggressiveness)

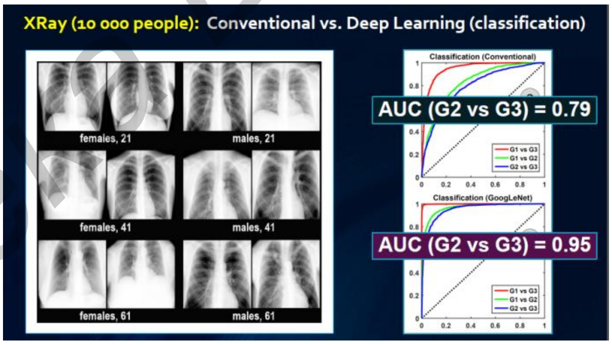

Cancer Diagnosis: Estimating Proliferation Score (our results)
Results
Task 1: Prediction of the proliferation score based on mitosis counting
Automatic methods
Tumor Proliferation Assessment Challenge 2016
Task 2: Prediction of the proliferation score based on gene expression
Automatic methods

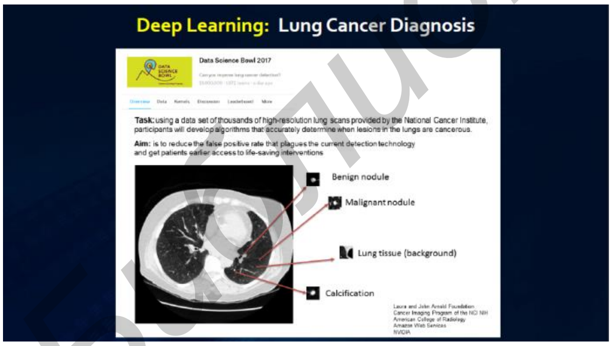

IBM Open Power Challenge: 1st place (together with Altoros Development)
Distributed TensorFlow for cancer detection
Finalists
OpenPOWER cluster network Scheme
ALTOROS

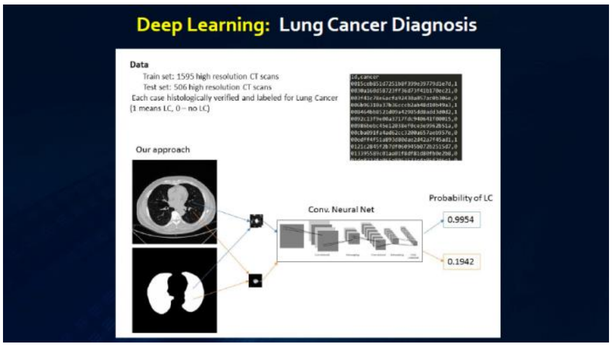

Chest X-Ray (~2 000 000): Benchmarking Convolutional Networks
Males, 21
Females, 70

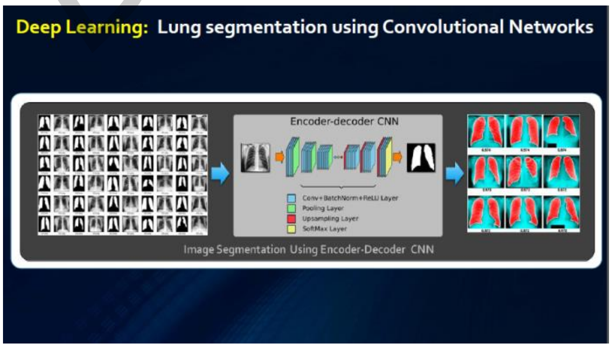

XRay (100 000 people): Image mining (past research)
Experience: Lung Shape Mining

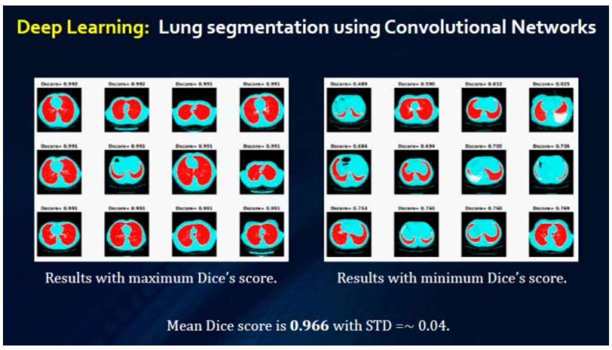

XRay (10 000 people): Conventional vs. Deep Learning (age prediction)
Conventional
Mean Error ~ 13.5 years
Deep Learning
Mean Error ~ 6.0 years
Age range: 50 years
100 Males + 100 Females for each year
Deep Learning dramatically outperform the existing methods

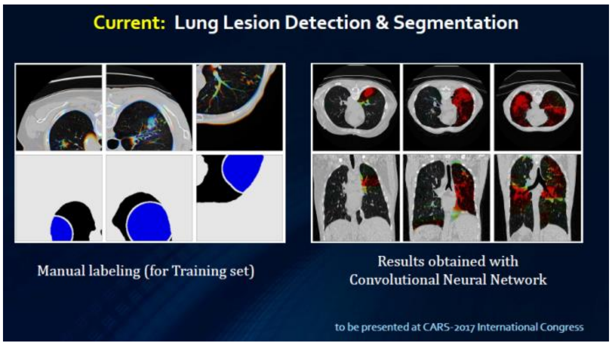

XRay (10 000 people): Conventional vs. Deep Learning (classification)
females, 21
males, 21
females, 41
males, 41
females, 61
males, 61
Classification (Conventional)
AUC (G2 vs G3) = 0.79
Classification (GoogLeNet)
AUC (G2 vs G3) = 0.95

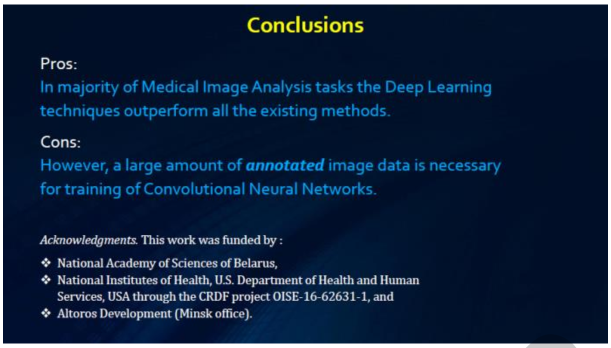

Deep Learning: Lung Cancer Diagnosis
Data Science Bowl 2017
Task: using a data set of thousands of high-resolution lung scans provided by the National Cancer Institute, participants will develop algorithms that accurately determine when lesions in the lungs are cancerous.
Aim: is to reduce the false positive rate that plagues the current detection technology and get patients earlier access to life-saving interventions
Benign nodule
Malignant nodule
Lung tissue (background)
Calcification


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
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).