

# ImageCLEF 2019: Projection-based CT Image Analysis for TB Severity Scoring and CT Report Generation

Vitali Liauchuk

United Institute of Informatics Problems, Minsk, Belarus  
vitali.liauchuk@gmail.com

**Abstract.** This paper presents an approach for automated analysis of 3D Computed Tomography (CT) images based on representing the 3D CT data as a set of 2D projection images along all three axes. Such approach reduces the dimensionality of the input data and therefore allows using less complicated models for image classification tasks. Deep Learning methods were used to predict most of the features of CT images of patients with lung tuberculosis (TB). For part of the features, conventional methods were used. Two different methods of segmentation of lungs were employed including the registration-based scheme. The proposed image analysis approach was utilized by United Institute of Informatics Problems (UIIP\_BioMed) participant for accomplishing the two subtasks of ImageCLEF Tuberculosis task of ImageCLEF 2019 international competition. Employing the proposed approach allowed achieving the best performance in both CT Report Generation and TB Severity Scoring subtasks. Source codes implementing the proposed methods are available on Github<sup>1</sup>.

**Keywords:** Computed Tomography, Tuberculosis, Deep Learning, Projections

## 1 Introduction

Automated analysis of 3D CT images is an important step in many tasks connected with the development of Computer-Aided Diagnosis systems, screening of lung diseases, early detection of pathology and development of the dedicated web-portals<sup>2</sup> [10]. However, the task of CT image analysis nowadays remains challenging due to a number of factors. Partly this is caused by the complexity of 3D CT image data as well as by the diversity of representations of such data. A CT image can often be represented as a set of 2D slices with inter-slice

---

Copyright © 2019 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0). CLEF 2019, 9-12 September 2019, Lugano, Switzerland.

<sup>1</sup> <https://github.com/skliff13/CompetitionsParticipation>

<sup>2</sup> <http://tbportals.niaid.nih.gov/>

distance varying between 0.5 and 5 mm. Variable size and shape of CT image voxels make it difficult to correctly apply many image analysis algorithms. The other problem comes from low availability of CT imaging data compared to some other biomedical image modalities such as X-ray [8, 14], histology [4, 13] and microscopy images [1]. This strengthens with the complexity and high costs of manual labeling of 3D CT images. Availability of good quality labeling significantly eases the task of CT image analysis [6, 9].

The approach for analysis of 3D CT images proposed with this study employs the idea of representing each 3D CT scan as a set of 2D projection images along all three ( $X$ ,  $Y$  and  $Z$ ) axes: sagittal, frontal and axial projections. The advantage of such an approach over the conventional slice-wise representation consists in the fact that each projection obtained via averaging of voxel intensity values contains information about all the slices present in the image. On the one hand, such approach significantly reduces the complexity of the input data which eases the task of training the Convolution Neural Networks on a limited amount of training data. On the other hand, generating projections along different axes ( $X$ ,  $Y$ ,  $Z$ ) provides additional "native" augmentation of the input data.

## 2 Subtasks and datasets

The tuberculosis task [2] of ImageCLEF 2019 Challenge [5] included two subtasks all dealing with 3D CT images. Both subtasks shared the same CT imaging data which included 218 images in the Training (also referred as Development) dataset and 117 in the Test dataset.

The subtask #1 (SVR subtask) was dedicated to the problem of categorizing TB cases into one of the two classes: high severity and low severity. In contrast to the previous year's challenge, the task was only to predict TB Severity class ("HIGH"/"LOW") rather than to provide a Severity score from 1 to 5. Also, along with the CT images of TB patients an additional information about the patients was provided. The metadata included information about presence of disability, relapse, presence of TB symptoms, co-morbidity, bacillarity, drug resistance status, patient's education, being ex-prisoner, smoking status and alcohol addiction. The task was aimed at automatic classification of TB Severity into "HIGH" and "LOW" classes. The frequencies of occurrence of each metadata label are listed in Table 1.

The subtask #2 (CTR subtask) was the newly-introduces subtask which was dedicated to automated generation of CT reports which indicated presence of several types of abnormalities in lungs. Such automated annotation of CT scans is important for development of the dedicated image databases. The task was to predict the presence of six types of findings in CT scans. Information about the corresponding labels is listed in Table 2.

**Table 1.** Presence of metadata labels for SVR subtask in the datasets.

<b>Label</b>	<b>In Training set</b>	<b>In Test set</b>
Disability	34	15
Relapse	76	42
SymptomsOfTB	117	47
Comorbidity	122	56
Bacillary	185	108
DrugResistance	149	91
HigherEducation	28	19
ExPrisoner	27	12
Alcoholic	49	29
Smoking	114	70

**Table 2.** Presence of labels for CTR subtask in the Training dataset.

<b>Finding</b>	<b>In Training set</b>
LeftLungAffected	156
RightLungAffected	177
LungCapacityDecrease	64
Calcification	28
Pleurisy	16
Caverns	89

### 3 Methods

This section contains a description of the methods used with the current study. Some results obtained for the CTR subtask were used for completion of the SVR subtask.

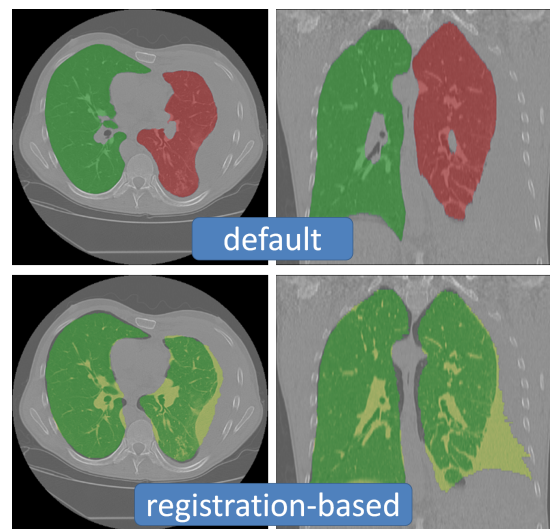
#### 3.1 Data pre-processing

The key idea of the approach proposed with this study consists in converting 3D CT scans into 2D projections followed by analysis of the obtained 2D projection images with use of Deep Learning and conventional methods.

For generation of the projections, two different versions of automatically extracted lung masks were used: the lung masks provided by the competition organizers [3] ("default") and the masks obtained via a conventional segmentation-through-registration scheme [12]. A short description of implementation of the registration-based lung segmentation method can be found in [7]. Source codes implementing this method are available on Github<sup>3</sup>. The lung masks obtained via registration-based approach are in general less accurate compared to the default ones but they appear useful in the cases of presence of large lesions in lungs. Fig. 1 illustrates both versions of lung masks on a CT image of a patient with pleurisy (Patient-ID: "CTR\_TRN\_013"). It can be seen that the default lung masks tend to leave parts of large lesions outside of the segmentation.

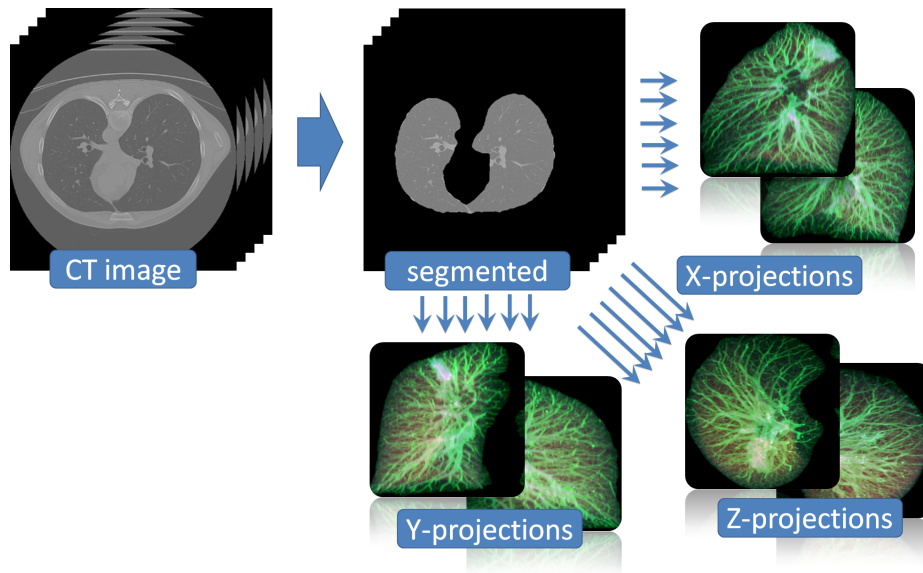
<sup>3</sup> [https://github.com/skliff13/CT\\_RegSegm](https://github.com/skliff13/CT_RegSegm)

The process of generation of CT projection images considered performing the following major procedures. The CT image voxel intensity values were increased by the value of 1024 Hounsfield Units (HU) to ensure only positive intensity values. To exclude the influence of image segmentation faults on the lung borders, the lung masks were eroded using ellipsoidal structure elements of radius 10 along  $XY$  plane. The radius along  $Z$ -axis was calculated respectively to the inter-slice distance of the specific CT image. Image intensity values outside of the eroded lung masks were zeroed. Optionally, an intensity threshold can be applied to filter noisy voxels with low intensities.



**Fig. 1.** CT image of a TB patient having pleurisy with the default lung masks (top) and the lung masks obtained via registration-based approach (bottom).

Each 2D projection image was represented as a pseudo-RGB image and had three channels. The first (red) channel contained mean values of CT image intensities along the specified axis. The mean values were finally divided on their maximum value along the projection image. The second (green) channel contained maximum intensity values along the specified axis divided by the value of 1500 which corresponds to  $1500 - 1024 = 476$  HU in terms of the original voxel intensities. The third (blue) channel was composed of the corresponding Standard Deviation intensity values. As in the case of red channel, blue channel values were divided by their maximum. Finally, the resultant 2D projection images were cropped using the bounding boxes of non-zero regions and resized to  $256 \times 256$  pixels size. Projections were generated for each lung separately. The general scheme of generation of 2D projections from 3D CT images is shown in Fig. 2



**Fig. 2.** General scheme of generation of 2D projections from 3D CT images.

Using this scheme, six 2D projection images were generated for each CT scan in the dataset: sagittal, frontal and axial projections for each of the two lungs. Three versions of projection images were generated with use of different lung masks and intensity thresholds. Details on the versions are shown in Table 3.

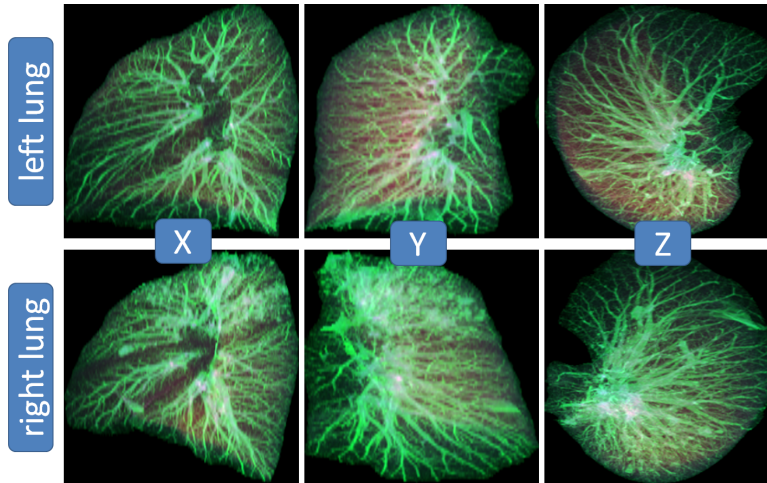
**Table 3.** Versions of 2D projections used with this study.

Version	Lung masks used	Intensity threshold
v1.0	Default	No threshold
v1.1	Registration-based	No threshold
v1.2	Default	+1000 HU

Fig. 3 illustrates the v1.0 projection images along the three axes of left and right lung of a CT scan which is labeled as having only right lung affected with presence of caverns (Patient-ID: "CTR\_TRN\_037"). It can be seen from the projection images that the upper lobe of right lung is affected by the disease.

### 3.2 Training and validation subsets

For correct validation of the developed image analysis algorithms, the Development dataset provided by the organizers was split into training and validation subsets. With this study, each 4-th CT image from the Development dataset was labeled as validation image (54 cases), whereas all other CT cases were used for



**Fig. 3.** Examples of 2D sagittal ( $X$ ), frontal ( $Y$ ) and axial ( $Z$ ) projections of left and right lungs.

training the algorithms (164 cases). This training/validation data split was used at every stage for both subtasks without changes.

### 3.3 Utilized neural network model

Convolutional Neural Networks (CNNs) were used to predict most of CT image characteristics. A deeply modified version of a popular VGG16 [11] neural network was used with this study.

Conventional VGG16 architecture includes 13 convolutional layers in five blocks and three fully-connected layers ( $13 + 3 = 16$ ). Considering the limited amount of the available data, the employed neural network model was simplified by (i) reducing the number of convolutional layers to one per block (5 layers), (ii) using global maximum pooling after the last convolutional block instead of flattening, and (iii) reducing the number of nodes in the fully-connected layers to 128. All this significantly reduces the number of trained parameters of the network model which minimizes the effect of over-fitting. The resultant network model included eight trainable layers: 5 convolutional and 3 fully-connected.

### 3.4 Prediction of labels for the CTR subtask

**Training CNN for lung-wise detection of abnormalities.** Most of the metadata labels provided with this challenge (TB Severity class, presence of certain lesions) were assigned at the level of CT scans. On the other hand, *LeftLungAffected* and *RightLungAffected* labels from the CTR subtask metadata can be used to train image classification models to detect presence of abnormalities at single lung level.

At this step, a Convolutional Neural Network was trained to classify a 2D lung projection image as either healthy ("normal") or affected by TB ("abnormal"). The training and validation data was composed of lung projection images along  $X$ ,  $Y$  and  $Z$  axes, left and right lungs were considered separately. Thus, the total number of input samples was  $164 \times 2$  (lungs)  $\times 3$  (projections) = 984 for training and  $54 \times 2 \times 3 = 324$  for validation. The size of the network input was  $256 \times 256 \times 3$  which corresponded to a projection image of a single lung. Categorical cross-entropy was used as the cost function. Additional on-the-fly data augmentation was applied which included random rotations with 10 degree range, width and height shifts and random re-scaling.

The network training for lung-wise "normal"/"abnormal" classification was performed with use of lung projections v1.0, Adam optimizer and learning rate set to  $10^{-5}$  within 120 training epochs. Network weights were initialized randomly. Here and further the training was performed with use of Keras framework with Tensorflow backend on a personal computer equipped with a GPU of Nvidia TITAN X type with 3072 CUDA Cores and 12 GB of GDDR5 on-board memory.

The trained CNN was evaluated on the lung projection images generated from the validation CT scans to produce the confidence scores. Projection-wise comparison to the ground truth data resulted in 0.865 Area Under ROC-Curve (AUC). However, 2D projections generated along different axes from the same 3D image must have the same label ("normal" or "abnormal"). Therefore, the confidence scores obtained for the three projections of each lung were combined to increase the lung-wise classification quality. The best performance of lung-wise "normal"/"abnormal" classification was achieved by means of using maximum among the three confidence values which corresponded to  $X$ ,  $Y$  and  $Z$  axes (AUC = 0.920).

**Detection of caverns and lung capacity decrease.** Similar approach was used for detection of caverns and lung capacity decrease. In this case, the corresponding labels were specified at CT image level, therefore the input data consisted of projections of both lungs. The corresponding neural network model had the input of size  $256 \times 512 \times 3$  which corresponded to two concatenated projection images. Lung projections v1.1 were used.

Weights of the convolutional layers were initialized with the corresponding weights of the CNN previously trained for lung-wise "normal"/"abnormal" classification. The networks for detection of lung capacity decrease and caverns were trained for 50 and 120 epochs respectively. Evaluation of the trained CNNs on the validation data resulted in 0.832 AUC for lung capacity decrease and 0.809 AUC for caverns detection at projection level. Averaging the confidence values obtained for  $X$ ,  $Y$  and  $Z$  projections gave the CT-wise detection performance of 0.856 AUC for lung capacity decrease and 0.879 AUC for caverns.

**Detection of calcification and pleurisy.** Scores for prediction of calcification were calculated using the mean intensity values of v1.2 projection images. These

values reflect the number of voxels in lung regions of the original CT images with intensities exceeding 1000 HU threshold. Evaluation of such scores on the Development dataset resulted in 0.726 AUC value.

Since the lung masks provided by the organizers tend to exclude pleurisy from lung segmentation (see Fig.1), the scores for detection of pleurisy were calculated as difference between the volume of lung masks obtained via registration-based scheme and the volume of lung masks provided by the organizers. Evaluation of such scores gave 0.776 AUC value on the Development dataset.

**Summary on the CTR subtask.** The results of evaluation of the prediction algorithms for the CTR subtask are shown in Table 4. All the AUC values were assessed on the validation dataset.

**Table 4.** Prediction performance for the CTR subtask evaluated on the validation data.

<b>Label</b>	<b>AUC</b>
LeftLungAffected	0.906
RightLungAffected	0.957
Calcification	0.765
Caverns	0.879
Pleurisy	0.824
LungCapacityDecrease	0.856
Mean AUC	0.864
Min AUC	0.765

### 3.5 Prediction of TB Severity in the SVR subtask

The approach used with this study for prediction of TB severity consisted of two major stages.

The first stage considered training a CNN for prediction of severity class with use of only CT projection images. At this stage, the VGG16 network was trained to classify TB severity into "HIGH" and "LOW" classes. Similarly to the detection of caverns and lung capacity decrease in the CTR subtask, input for the CNN for severity classification was composed of projections of both left and right lungs. Weights of all trainable layers (convolutional and fully-connected) of the network were initialized with the corresponding weights of the CNN previously trained for lung-wise "normal"/"abnormal" classification. The network was trained on v1.0 projection images during 60 epochs. Evaluation of the trained network on the validation set resulted in 0.768 AUC value.

At the second stage, the output of a trained CNN (confidence score) was combined with the available metadata features for prediction of TB severity with use of Linear Regression classifier. Several more classifiers were tested including Random Forests, Logistic Regression and Support Vector Machine, but Linear



Regression gave the best AUC on the validation subset. Two sets of the metadata features were used: all available metadata features and the subset of four selected features which gave the same AUC on the validation subset. The subset of four features included *DrugResistance*, *HigherEducation*, *ExPrisoner* and *Alcoholic* labels.

AUC values assessed on the validation data using different combinations of metadata features and CNN outputs obtained from different projection images are shown in Table 5. The testing results presented with the table suggest that 2D projection images generated along *X* axis were the most informative for prediction of TB Severity class by the neural network. Utilizing projections along *Z* axis did not improve the results compared to the prediction using only *X* and *Y* projections.

**Table 5.** Prediction performance for the SVR subtask using different combinations of the metadata features and trained CNN outputs.

Metadata	CNN outputs (projections used)	AUC
All	-	0.785
Four	-	0.785
-	X only	0.804
-	Y only	0.801
-	Z only	0.707
-	X and Y only	0.818
-	X, Y and Z	0.793
All	X only	<b>0.844</b>
All	X and Y only	0.830
All	X, Y and Z	0.820
Four	X only	0.852
Four	X and Y only	<b>0.876</b>
Four	X, Y and Z	0.849

Among the combinations which used all available metadata features, the one which used CNN outputs from *X* projections showed the best performance with 0.844 AUC on the validation subset. The overall best performance on the validation data was achieved using the subset of four metadata features along with CNN outputs from *X* and *Y* projections. These two combinations were used in the two submitted runs for the SVR subtask.

## 4 Submissions and results

As the result of this study, three runs were submitted by UIIP\_BioMed for the CTR subtask and two for the SVR subtask. Full list of the submitted results for both subtasks is available at the task web page<sup>4</sup>. Network models and classifiers

<sup>4</sup> <https://www.imageclef.org/2019/medical/tuberculosis/>

used for predicting the Test data were trained on the training subset of Development data (see subsection 3.2). All the computed prediction scores were scaled to the range from 0 to 1 to fit the submission requirements.

Subsection 3.4 describes the methods used for the last ("CTR\_run3...") submission for the CTR subtask. The previous submissions differed with the ways of predicting *LungCapacityDecrease*, *Pleurisy* and *Caverns* and demonstrated poorer performance on the validation data. In "CTR\_run1...", these three findings were predicted by a single neural network trained for multi-class multi-label classification (0.844 mean AUC on validation). In "CTR\_run2...", *LungCapacityDecrease* and *Caverns* labels were predicted using separate CNNs trained for binary classification which gave better results (0.852 mean AUC on validation). Table 6 shows results achieved by the participants in the CTR subtask, one best run for each participating group. The run submitted by UIIP\_BioMed achieved the highest values of both mean and minimum AUC for prediction of presence of lung abnormalities.

**Table 6.** The best participants' runs submitted for the CTR subtask.

Group Name	Mean AUC	Min AUC	Rank
UIIP_BioMed	0.7968	0.6860	1
CompElecEngCU	0.7066	0.5739	2
MedGIFT	0.6795	0.5626	3
San Diego VA HCS/UCSD	0.6631	0.5541	4
HHU	0.6591	0.5159	5
UIIP	0.6464	0.4099	6
MostaganemFSEI	0.6273	0.4877	7
UniversityAlicante	0.6190	0.5366	8
PwC	0.6002	0.4724	9
LIST	0.5523	0.4317	10

The methods used for submissions in the SVR subtask are described in subsection 3.5. The first submitted run ("SRV\_run1...") used all metadata features and  $X$  projections for TB Severity classification and achieved the overall best performance in both AUC (0.7877) and Accuracy (0.7179). The second submission "SRV\_run2..." used the subset of four metadata features along with  $X$  and  $Y$  projections. This approach demonstrated better performance on the validation data. However, on the Test data it resulted in lower AUC (0.7636 vs. 0.7877) but higher Accuracy (0.7350 vs. 0.7179). Results of the participants' submissions with the highest AUC values are shown in Table 7.

## 5 Conclusions

The results of this study allow to draw the following conclusions:

**Table 7.** The best participants’ runs submitted for the SVR subtask.

<b>Group Name</b>	<b>AUC</b>	<b>Accuracy</b>	<b>Rank</b>
<u>UIIP_BioMed</u>	<u>0.7877</u>	<u>0.7179</u>	<u>1</u>
UIIP	0.7754	0.7179	2
HHU	0.7695	0.6923	3
CompElecEngCU	0.7629	0.6581	4
San Diego VA HCS/UCSD	0.7214	0.6838	5
MedGIFT	0.7196	0.6410	6
UniversityAlicante	0.7013	0.7009	7
MostaganemFSEI	0.6510	0.6154	8
SSN College of Engineering	0.6264	0.6068	9
University of Asia Pacific	0.6111	0.6154	10
FIIAugt	0.5692	0.5556	11

- 2D projections of lungs generated from 3D CT scans preserve basic information about the lungs structure which is sufficient for detection of abnormalities of different sort.
- TB Severity can be assessed based on the analysis of CT scans, however additional information on clinical data, drug resistance and patient’s social status are helpful for more accurate scoring of TB Severity.
- Usefulness of different methods of lungs segmentation may vary depending of the specific task being solved. Combining multiple lung segmentation methods may provide additional information which can be useful for detection of certain abnormalities in lungs.

## Acknowledgements

This study was partly supported by the National Institute of Allergy and Infectious Diseases, National Institutes of Health, U.S. Department of Health and Human Services, USA through the CRDF project DAA3-18-64818-1 ”Year 7: Belarus TB Database and TB Portals”.

## References

1. Al-Kofahi, Y., Zaltsman, A., Graves, R., Marshall, W., Rusu, M.: A deep learning-based algorithm for 2-D cell segmentation in microscopy images. *BMC Bioinformatics* **19**(1), 365 (Oct 2018). <https://doi.org/10.1186/s12859-018-2375-z>, <https://doi.org/10.1186/s12859-018-2375-z>
2. Dicente Cid, Y., Liauchuk, V., Klimuk, D., Tarasau, A., Kovalev, V., Müller, H.: Overview of ImageCLEFtuberculosis 2019 - automatic ct-based report generation and tuberculosis severity assessment. In: CLEF2019 Working Notes. CEUR Workshop Proceedings, CEUR-WS.org <<http://ceur-ws.org>>, Lugano, Switzerland (September 9-12 2019)
3. Dicente Cid, Y., Jiménez del Toro, O.A., Depeursinge, A., Müller, H.: Efficient and fully automatic segmentation of the lungs in ct volumes. In: Goksel, O., Jiménez del

- Toro, O.A., Foncubierta-Rodríguez, A., Müller, H. (eds.) Proceedings of the VIS-CERAL Anatomy Grand Challenge at the 2015 IEEE ISBI. pp. 31–35. CEUR Workshop Proceedings, CEUR-WS (May 2015)
4. Ehteshami Bejnordi, B., Veta, M., Johannes van Diest, P., van Ginneken, B., Karssemeijer, N., Litjens, G., van der Laak, J.A.W.M., , the CAMELYON16 Consortium: Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. *JAMA* **318**(22), 2199–2210 (12 2017). <https://doi.org/10.1001/jama.2017.14585>, <https://doi.org/10.1001/jama.2017.14585>
  5. Ionescu, B., Müller, H., Péteri, R., Cid, Y.D., Liauchuk, V., Kovalev, V., Klimuk, D., Tarasau, A., Abacha, A.B., Hasan, S.A., Datla, V., Liu, J., Demner-Fushman, D., Dang-Nguyen, D.T., Piras, L., Riegler, M., Tran, M.T., Lux, M., Gurrin, C., Pelka, O., Friedrich, C.M., de Herrera, A.G.S., Garcia, N., Kavallieratou, E., del Blanco, C.R., Rodríguez, C.C., Vasilopoulos, N., Karampidis, K., Chamberlain, J., Clark, A., Campello, A.: ImageCLEF 2019: Multimedia retrieval in medicine, lifelogging, security and nature. In: Experimental IR Meets Multilinguality, Multimodality, and Interaction. Proceedings of the 10th International Conference of the CLEF Association (CLEF 2019), vol. 2380. LNCS Lecture Notes in Computer Science, Springer, Lugano, Switzerland (September 9-12 2019)
  6. Kalinovsky, A., Liauchuk, V., Tarasau, A.: Lesion detection in CT images using Deep Learning semantic segmentation technique. In: International Workshop "Photogrammetric and computer vision techniques for video surveillance, biometrics and biomedicine". The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XLII, pp. 13–17. Moscow, Russia (May 2017). <https://doi.org/10.5194/isprs-archives-XLII-2-W4-13-2017>, <http://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLII-2-W4/13/2017/>
  7. Liauchuk, V., Kovalev, V.: ImageCLEF 2017: Supervoxels and co-occurrence for tuberculosis CT image classification. In: CLEF2017 Working Notes. CEUR Workshop Proceedings, CEUR-WS.org <<http://ceur-ws.org>>, Dublin, Ireland (September 11-14 2017)
  8. Liauchuk, V., Kovalev, V.: Detection of lung pathologies using deep convolutional networks trained on large X-ray chest screening database. In: Proceedings of the 14th international conference on Pattern Recognition and Information Processing (PRIP'2019). Minsk, Belarus (May 21-23 2019)
  9. Liauchuk, V., Tarasau, A., Snezhko, E., Kovalev, V.: ImageCLEF 2018: Lesion-based TB-descriptor for CT image analysis. In: CLEF2018 Working Notes. CEUR Workshop Proceedings, CEUR-WS.org <<http://ceur-ws.org>>, Avignon, France (September 10-14 2018)
  10. Rosenthal, A., Gabrielian, A., Engle, E., Hurt, D.E., Alexandru, S., Crudu, V., Sergueev, E., Kirichenko, V., Lapitskii, V., Snezhko, E., Kovalev, V., Astrovko, A., Skrahina, A., Taaffe, J., Harris, M., Long, A., Wollenberg, K., Akhundova, I., Ismayilova, S., Skrahin, A., Mammadbayov, E., Gadirova, H., Abuzarov, R., Seyfaddinova, M., Avaliani, Z., Strambu, I., Zaharia, D., Muntean, A., Ghita, E., Bogdan, M., Mindru, R., Spinu, V., Sora, A., Ene, C., Vashakidze, S., Shubladze, N., Nanava, U., Tuzikov, A., Tartakovskiy, M.: The TB Portals: an open-access, web-based platform for global drug-resistant-tuberculosis data sharing and analysis. *Journal of clinical microbiology* **55**(11), 3267–3282 (2017)
  11. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition (2014)

12. Sluimer, I., Prokop, M., van Ginneken, B.: Toward automated segmentation of the pathological lung in ct. *IEEE Transactions on Medical Imaging* **24**(8), 1025–1038 (Aug 2005). <https://doi.org/10.1109/TMI.2005.851757>
13. Veta, M., Heng, Y.J., Stathonikos, N., Bejnordi, B.E., Beca, F., Wollmann, T., Rohr, K., Shah, M.A., Wang, D., Rousson, M., Hedlund, M., Tellez, D., Ciompi, F., Zerhouni, E., Lanyi, D., Viana, M., Kovalev, V., Liauchuk, V., Phoulady, H.A., Qaiser, T., Graham, S., Rajpoot, N., Sjöblom, E., Molin, J., Paeng, K., Hwang, S., Park, S., Jia, Z., Chang, E.I.C., Xu, Y., Beck, A.H., van Diest, P.J., Pluim, J.P.: Predicting breast tumor proliferation from whole-slide images: The TUPAC16 challenge. *Medical Image Analysis* **54**, 111 – 121 (2019). <https://doi.org/https://doi.org/10.1016/j.media.2019.02.012>, <http://www.sciencedirect.com/science/article/pii/S1361841518305231>
14. Zaidi, S.M.A., Habib, S.S., Van Ginneken, B., Ferrand, R.A., Creswell, J., Khowaja, S., Khan, A.: Evaluation of the diagnostic accuracy of computer-aided detection of tuberculosis on chest radiography among private sector patients in Pakistan. *Scientific reports* **8**(1), 12339 (2018). <https://doi.org/10.1038/s41598-018-30810-1>