

Overview of ImageCLEFtuberculosis 2020 — Automatic CT-based Report Generation

Serge Kozlovski¹, Vitali Liauchuk¹, Yashin Dicente Cid², Aleh Tarasau³,
Vassili Kovalev¹, and Henning Müller^{4,5}

¹ United Institute of Informatics Problems, Minsk, Belarus;

² University of Warwick, Coventry, England, UK

³ Republican Research and Practical Centre for Pulmonology and TB, Minsk,
Belarus;

⁴ University of Applied Sciences Western Switzerland (HES-SO), Sierre, Switzerland;

⁵ University of Geneva, Switzerland

kozlovski.serge@gmail.com

Abstract. ImageCLEF is a part of the Conference and Labs of the Evaluation Forum (CLEF) initiative and presents a set of image information retrieval tasks. ImageCLEF was historically focused on the variety of multimodal image classification, retrieval and annotation tasks. The tuberculosis task started in ImageCLEF in 2017 and changed from year to year. This year's edition was dedicated to the automatic generation of a lung-wise CT report (CTR) based on three relevant CT findings. This year 9 groups from 8 countries participated in the task and submitted results. This year's task is similar to the CTR (CT Report) subtask from the previous year, so it is possible to compare the results almost directly. Impressive improvement of the results was obtained with 0.92 (+0.10) average Area Under ROC-curve (AUC) and 0.89 (+0.20) minimum AUC for the three CT findings proposed.

Keywords: Tuberculosis, Computed Tomography, Image Classification, Automatic Report Generation, 3D Data Analysis

1 Introduction

ImageCLEF⁶ is a part of the the CLEF⁷ initiative and presents a set of image information retrieval tasks. Medical tasks were included in the 2nd edition of the ImageCLEF in 2004 and have been held every year since then [1–4]. The tuberculosis task is one of the medical tasks this year. More information on the other tasks organized in 2020 can be found in [5] and the past editions are described in [6–12].

Copyright © 2020 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0). CLEF 2020, 22-25 September 2020, Thessaloniki, Greece.

⁶ <http://www.imageclef.org/>

⁷ <http://www.clef-initiative.eu/>

Tuberculosis (TB) is a bacterial infection caused by a germ called *Mycobacterium tuberculosis*. After more than 130 years since its discovery, the disease remains a persistent threat and one of the top 10 causes of death worldwide according to the WHO [13]. The bacteria usually attack the lungs and generally TB can be cured with antibiotics. However, the different types of TB require different treatments, and therefore detection of the specific case characteristics is important. In particular, detection of the TB type and presence of different lesion types are important real-world tasks.

In the previous editions of this task, the setup evolved from year to year. In the first two editions [8, 9] the participants had to detect Multi-drug resistant patients (MDR subtask) and to classify the TB type (TBT subtask) both based only on the CT image. After 2 editions it was concluded to drop the MDR subtask because it seemed impossible to solve it well based only on the CT image, and the TBT subtask was also discontinued because of a very little improvement in the results between the 1st and the 2nd editions. At the same time, most of the participants obtained good results in the severity scoring (SVR) subtask introduced in 2018. In the third edition, the SVR subtask was included again for the updated data set, and a new subtask based on providing an automatic report (CT Report) for the TB case was added [7].

In this year’s edition, we decided to skip the SVR subtask and concentrate on the automated CT report generation task, since it has an important outcome that can have a major impact in the real-world clinical routine. To make the task both more attractive for participants and practically valuable, this year’s report generation was lung-based rather than CT-based, which means that labels for the left and right lungs were provided independently. The set of target labels in the CT Report was updated in accordance with the opinion of medical experts.

This article first describes the task proposed for TB in 2020. Then, details on the data sets, evaluation methodology and participation are given. The results section describes the submitted runs and the results obtained. A discussion and conclusion section ends the paper.

2 Task, Data Sets, Evaluation, Participation

2.1 The Task in 2020

In this task, the participants had to generate automatic lung-wise reports based on the CT image data. Each report should include the probability scores (ranging from 0 to 1) for each of the three labels and for each of the lungs. Two labels indicated the presence of a specific lesion in the lung - caverns and pleurisy, the third label indicated that the lung is affected by any lesion (not limited to the mentioned two).

The resulting list of entries for each CT included six entries: “left lung affected”, “right lung affected”, “caverns in the left lung”, “caverns in the right lung”, “pleurisy in the left lung”, “pleurisy in the right lung”.

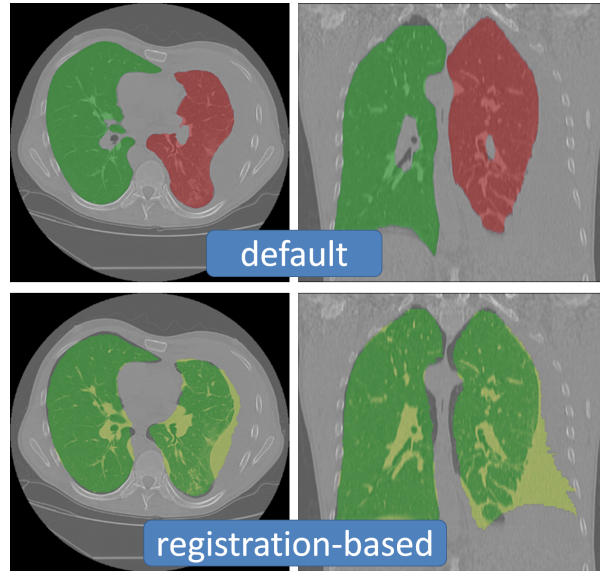


Fig. 1. CT image of a TB patient having pleurisy with the default lung masks (top) and the lung masks obtained via registration-based approach (bottom).

2.2 Data Sets

In this edition, a data set containing chest CT scans of 403 TB patients was used. The data set was divided into 283 patients for training and 120 for testing subsets.

For every patient, a 3D CT image series was provided with a size of 512×512 pixels and median number of slices equal to 128. All the CT images were stored in NIFTI file format with .nii.gz file extension (g-zipped .nii files). This file format stores raw voxel intensities in Hounsfield units (HU) as well as the corresponding image meta-data such as image dimensions, voxel size in physical units, slice thickness, etc.

In addition, for each CT image two versions of automatically extracted masks of the lungs were provided. The first version of segmentation [14] (default) was retrieved using the same technique as the previous years and provides accurate masks but it tends to miss large abnormal regions of lungs in the most severe TB cases. The second version of the segmentation [15] was retrieved using a non-rigid image registration scheme, which on the contrary provides more rough bounds, but behaves more stable in terms of including lesion areas. Fig. 1 illustrates both versions of lung masks, default and registration-based, on a CT image of a patient with pleurisy. It can be seen that the default lung masks tend to leave parts of large lesions outside of the segmentation.

All the data were provided by the Republican Research and Practical Center for Pulmonology and Tuberculosis which is located in Minsk, Belarus. The data

were collected and labelled in the framework of several projects that aim at the creation of information resources on the lung TB and drug resistance challenges.

The projects were conducted by a multi-disciplinary team and funded by the National Institute of Allergy and Infectious Diseases, National Institutes of Health (NIH), U.S. Department of Health and Human Services, USA, through the Civilian Research and Development Foundation (CRDF).

The dedicated web-portal⁸ developed in the framework of the projects stores information of more than 3000 TB cases patients from 15 countries. The information includes CT scans, X-ray images, genome data, clinical and social data.

2.3 Labels

Pathological changes in the lungs which are affected by tuberculosis may be represented by a large variety of findings. There are common findings that appear in most of the cases, such finding normally include aggregations of foci and infiltrations of different sizes. In more rare cases one can observe such kind of lesions as fibrosis, atelectasis, pneumothorax, etc.

In the 2020 edition of the task, three labels were assigned for each lung individually: "lung affected", "presence of pleurisy", "presence of caverns".

The "Left lung affected" and "right lung affected" labels indicated presence of any kind of TB-associated lesions in the left and right lung, respectively.

Presence of pleurisy and caverns were considered separately from the other types of lesions.

Pleurisy is known as inflammation of the membranes that surrounds the lungs and line the chest cavity⁹. Caverns, also known as pulmonary cavities, are gas-filled areas of the lung in the center of nodules or areas of consolidation [16].

Typical examples of CT findings are shown in Fig. 2. Table 1 details the distribution of patients within each label.

Table 1. Distribution of CT images with each label.

	Left lung affected	Right lung affected	Left lung pleurisy	Right lung pleurisy	Left lung caverns	Right lung caverns
Train	211 (75%)	233 (82%)	7 (2%)	14 (4%)	66 (23%)	79 (28%)
Test	75 (63%)	99 (83%)	3 (3%)	5 (4%)	28 (23%)	46 (38%)

2.4 Evaluation Measures and Scenario

Similarly to the previous editions, each participating group could submit up to 10 runs.

⁸ <http://tbportals.niaid.nih.gov/>

⁹ <https://www.nhlbi.nih.gov/health-topics/pleurisy-and-other-pleural-disorders>

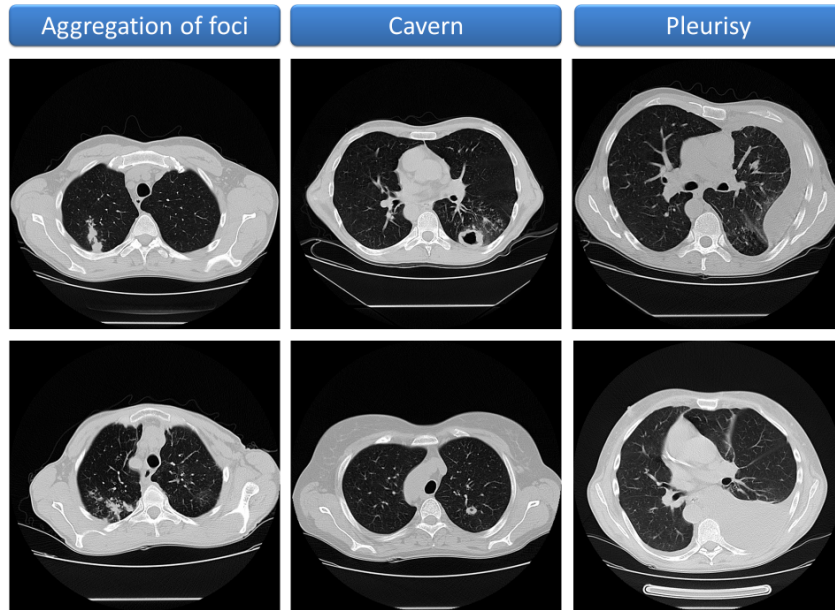


Fig. 2. Slices of typical CT images with three types of the TB-related findings.

The participants had to provide the probabilities for each of the three CT finding for each patient in the lung-wise manner, i.e. for each patient they had to provide a 6-dimensional vector with the probabilities. This task was considered a multi-binary classification problem and standard binary classification metrics were provided. The runs submitted by the participants were ranked based on the average ROC AUC and the min ROC AUC obtained.

The ROC AUCs were calculated separately for each of three target findings. That means the lung-wise predictions for the left and right lungs were concatenated, then ROC curve was created and the score was calculated.

The main purpose of the lung-wise labelling was to encourage the participants to switch from per-CT to per-lung analysis, which showed its effectiveness in the previous year edition of the task. Table 1 shows the numbers of CTs having each label.

2.5 Participation

This year there were 38 registered teams and 25 signed the end user agreement. 9 groups from 8 countries participated and submitted results. The number of submissions is a bit lower than in 2019 (13 for both subtasks). Table 2 shows the list of participants and their institutions.

Table 2. List of participants who submitted at least one run.

Group name	Main institution	Country
chejiao	Yunnan University	China
CompElecEngCU	Cukurova University	Turkey
FAST_NU_DS	National University of Computer and Emerging Sciences	Pakistan
JBTTM	SSN College of Engineering	India
KDE-lab	Toyohashi University	Japan
SenticLab.UAIC	SenticLab, Alexandru Ioan Cuza University of Iasi	Romania
SDVA-UCSD	San Diego VA/UCSD	USA
sztaki_dsd	Institute for Computer Science and Control	Hungary
uaic2020	Alexandru Ioan Cuza University of Iasi	Romania

3 Results

This section provides a detailed description of the results obtained by the participants.

To perform ranking in this task we used the mean ROC AUC and minimum ROC AUC values calculated over the three binary CT-findings proposed. Table 3 shows these two measures calculated for the best runs submitted by each participating groups. For each best run and for each CT-finding, Figures 3, 4, and 5 show the corresponding ROC curves. In addition, precision-recall (PR) plots are presented in Figures 6, 7, and 8. AUC values for all PR plots can be found in Table 4.

Table 3. Summary on the participant submissions and their results.

Group rank	Group name	# of runs	Mean ROC AUC	Min ROC AUC	Rank of the best run
1	SenticLab.UAIC	9	0.924	0.885	1
2	SDVA-UCSD	10	0.875	0.811	6
3	chejiao	7	0.791	0.682	16
4	CompElecEngCU	10	0.767	0.733	21
5	KDE-lab	10	0.753	0.698	28
6	FAST_NU_DS	3	0.705	0.644	37
7	uaic2020	8	0.659	0.562	40
8	JBTTM	2	0.601	0.432	49
9	sztaki_dsd	8	0.595	0.546	50

SenticLab.UAIC [17] is the winner of the task with mean ROC AUC of 0.924 and min ROC AUC of 0.885. In terms of ROC AUC their approach outperformed

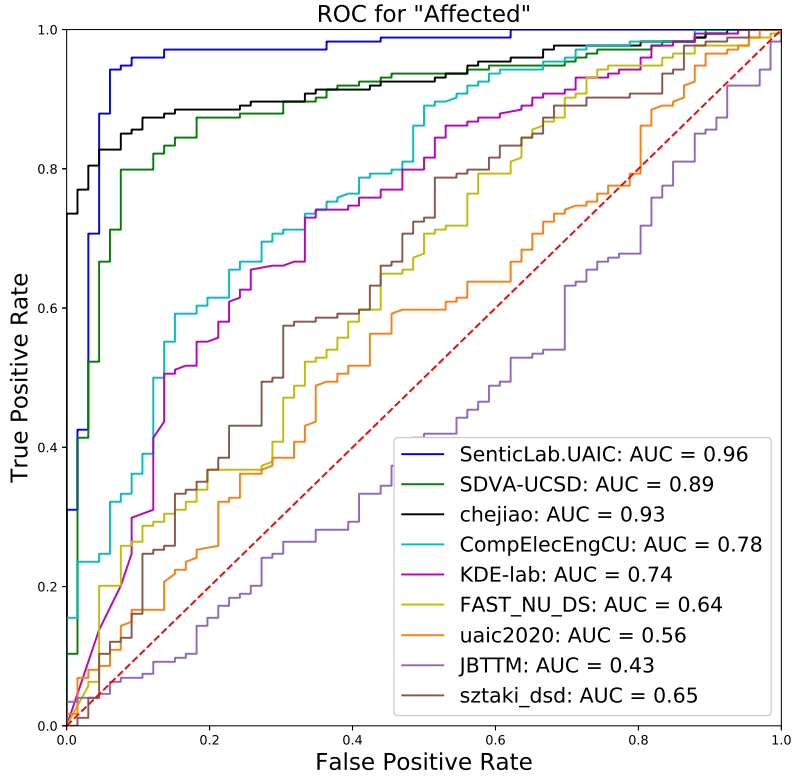


Fig. 3. ROC curve and AUC value obtained by the best run of each group for the "Affected" finding. The dashed line marks the ROC of a random classifier.

all the other methods for "Affected" and "Caverns" labels, although placed 2nd for "Pleurisy". In their experiments, the SenticLab.UAIC team compared several 2D and 3D CNNs. The 2D CNNs worked better in their validation phase and are the only approaches submitted for testing. They used the 2D slices and 2D projections of the 3D volumes in the different axes as input for the 2D CNNs. In addition, they used 3 different lung segmentations, the two provided and one more based on U-Net.

The SDVA-UCSD [18] team ranked 2nd place in terms of mean and min ROC AUC and achieved the best score for "Pleurisy" label. The team approach was based on the usage of 3D CNN with a convolutional block attention module (CBAM) and a customized loss functions. The team employed a laterality-neglection procedure for full utilization of lung-wise labelling advantages and

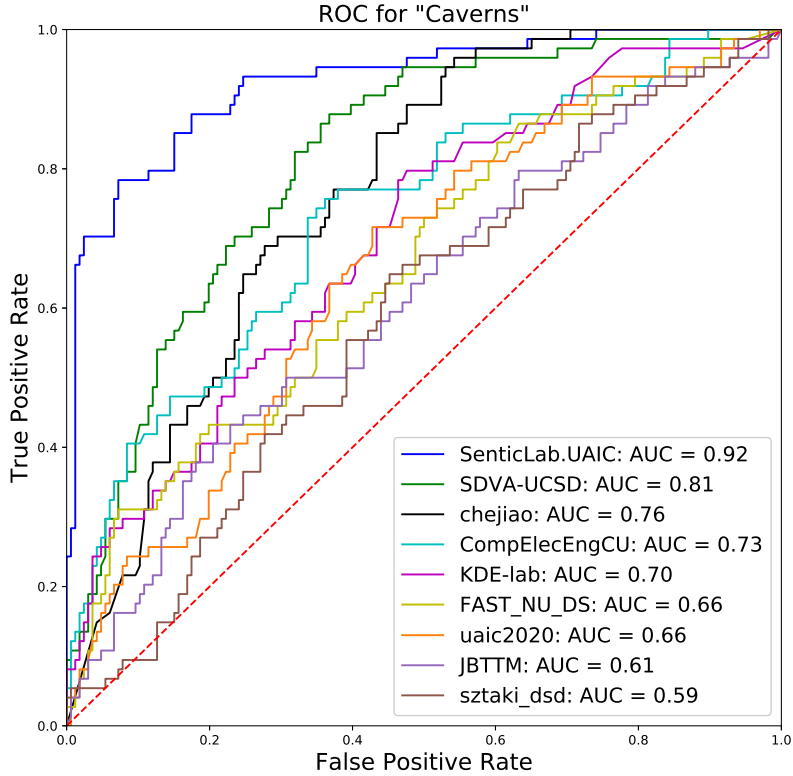


Fig. 4. ROC curve and AUC value obtained by the best run of each group for the "Caverns" finding. The dashed line marks the ROC of a random classifier.

gave attention to the accurate lung extraction based on the both provided lung masks.

The Chejiao [19] team treated the task as multiple binary classification tasks where left and right lung images were considered as independent data samples. They extended the projection-based image processing approach [20] by utilizing the ShuffleNet architecture and employing a Mixup data enhancement technique [21].

The CompElecEngCU team in their work [22] extracted 2D slices at all the three axes and used an ensemble of neural networks to predict the target features in a patient-wise manner.

The KDE-lab [23] used a neural network model that takes inputs from several 2D CNN networks trained on a large dataset of a general use images rather than

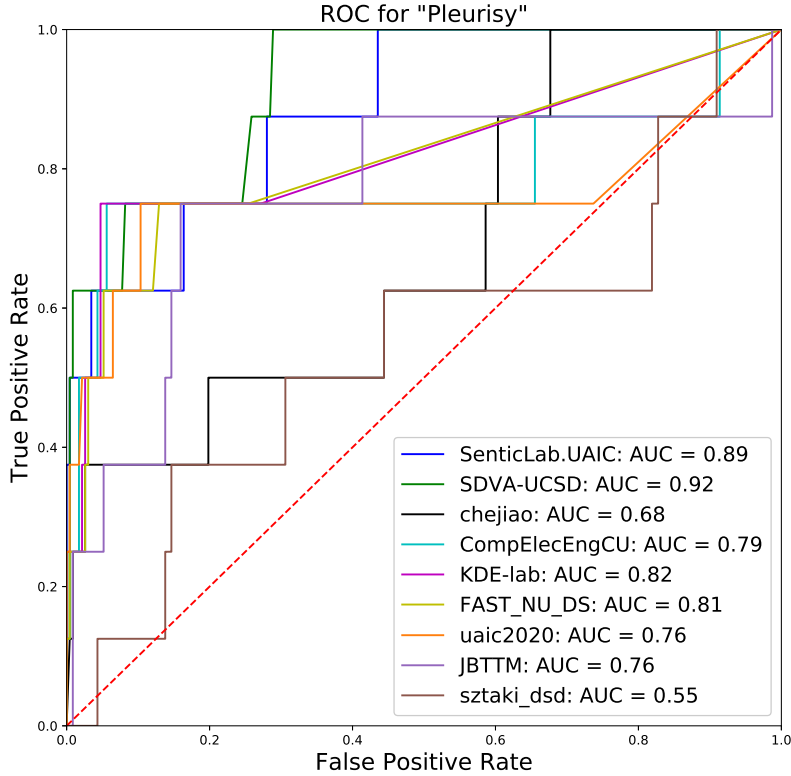


Fig. 5. ROC curve and AUC value obtained by the best run of each group for the "Pleurisy" finding. The dashed line marks the ROC of a random classifier.

the task dataset. The team used multi-axis projections as CNN input in their experiments.

The FAST_NU_DS [24] team tested several approaches based on training a classifier on a mixture of image features of different sort including conventional features such as Local Binary Patterns, Haralick features, intensity histograms and image features derived from a trained VGG19 neural network.

The UAIC [25] team used SVMs and a CNN for lung-wise processing. For each lung they selected the slices containing lung and discarded those slices which were too similar to neighbouring ones. Each slice was masked using the provided mask. Finally, the images are fed directly into the SVM and the CNN.

The JBTTM [26] team used a projection-based approach and utilized different lung masks for different labels, similarly to the previous year's winning

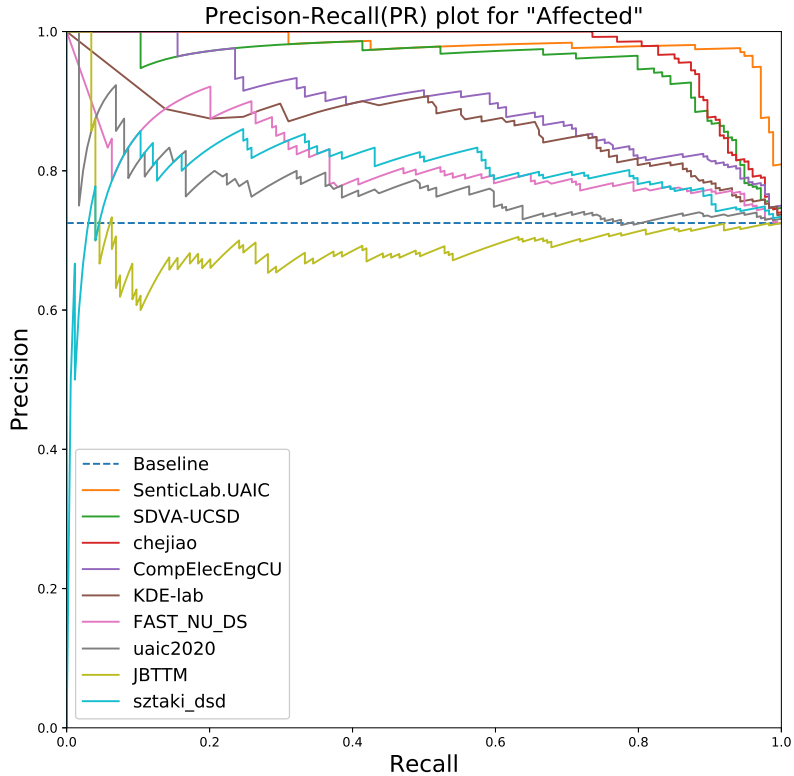


Fig. 6. PR curve obtained by the best run of each group for the "Affected" finding.

method [20]. The team also performed a series of experiments with 3D CNNs but the results were not used in the final submission.

The Sztaki_dsd [27] team used 2D CNN for a per-slice analysis of the CT images. Large parts of their experiment were dedicated to the aggregation of slice-based predictions to the CT level.

4 Discussion and Conclusions

The results obtained in this year's task have improved with respect to the similar CTR subtask presented in the 2019 edition. The SenticLab.UAIC team achieved 0.92 mean AUC, which is an improvement compared to the results achieved last year by the UIIP_BioMed team. Although this and previous year's CTR subtask

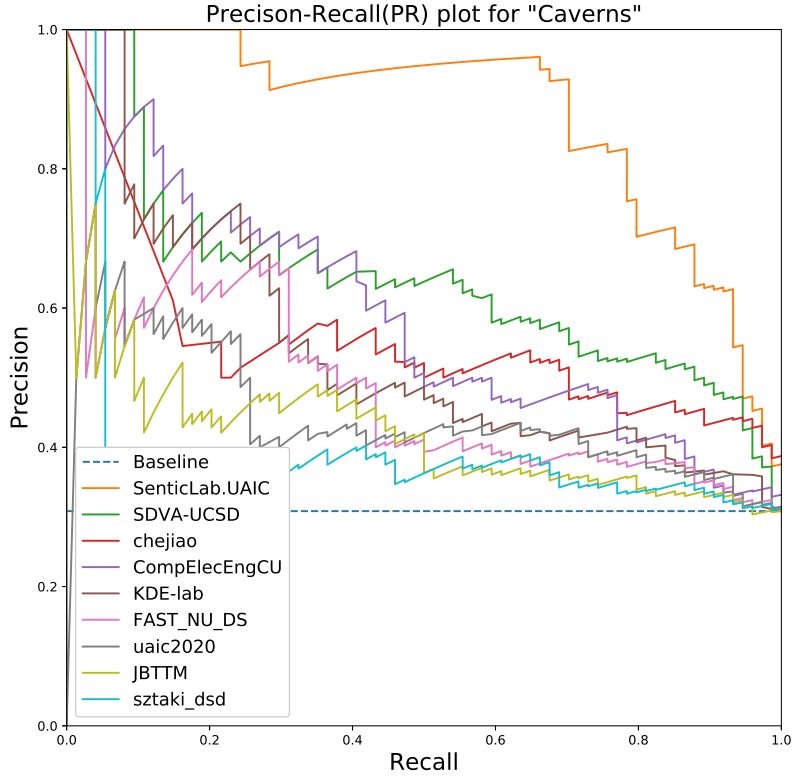


Fig. 7. PR curve obtained by the best run of each group for the "Caverns" finding.

results can not be directly compared due to the different labelling logic, we can still compare results for the three CT findings proposed in both this and the previous year, and observe increase in both mean ROC AUC (0.92 vs 0.82) and min ROC AUC scores (0.89 vs 0.69).

At the same time, we should note several points that make a comparison of the results a bit controversial. First, at least partially the improvement may be related to the more precise labelling in this year rather than more effective approaches. Second, due to the natural reasons there is a misalignment between CT-finding distribution in the training and test sets in both editions of the tasks and this misalignment is different. Third, the "pleurisy" finding is very rare, therefore the retrieved scores are not very reliable.

In any case, the review of the participants working notes demonstrates valuable experiments and extension of previously used methods. In particular, the

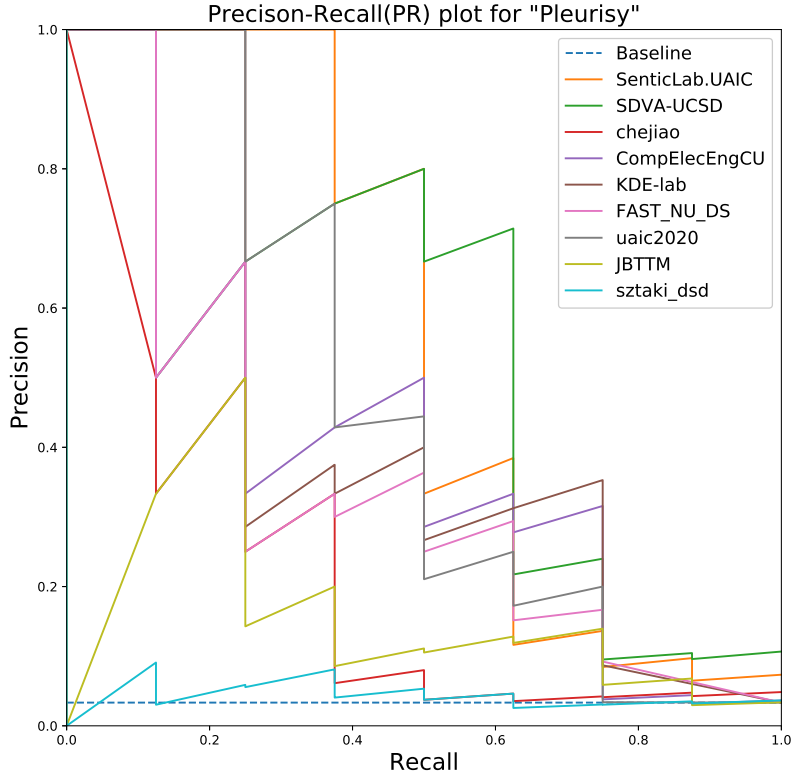


Fig. 8. PR curve obtained by the best run of each group for the "Pleurisy" finding.

analysis of the winning approach allows to conclude that the huge amount of experiments and the described method extensions definitely play an important role in achieving superior results.

In addition to ROC AUC metric we provide Precision-Recall (PR) curves and PR AUC values for better understanding of the results retrieved for naturally imbalanced test data. Accordingly to the PR AUC metric (Table 4), SenticLab.UAIC outperformed the other participants in "Caverns" (with a significant gain) and "Pleurisy" labels prediction and shared the best result with chejjiao team in the case of "Affected" label.

This year, only one group applied differing techniques for different findings, the others used a uniform approach to detect each of the CT-findings in a multi-binary classification setup. All the participants treated labels independently, without attempts to find a relation between the findings.

Table 4. PR AUC obtained by the best run of each group.

	Affected Caverns Pleurisy		
Baseline Classifier	0.72	0.31	0.03
SenticLab.UAIC	0.98	0.88	0.55
SDVA-UCSD	0.95	0.64	0.52
chejiao	0.98	0.55	0.21
CompElecEngCU	0.90	0.59	0.39
KDE-lab	0.87	0.54	0.43
FAST_NU_DS	0.81	0.48	0.35
uaic2020	0.77	0.44	0.43
JBTTM	0.70	0.42	0.15
sztaki_dsd	0.80	0.39	0.04

The trend toward using convolutional neural networks is strong again. Last year, 10 out of the 12 groups used CNNs at least in one of their attempts, and this year all groups used CNNs for their submissions. Several groups tried a few different methods during their experiments, all reported approaches are listed below.

The majority of the participants (six groups) used some variations of the projection-based approach [20]. These groups extracted axial, coronal and sagittal projections from the CT image and executed further analysis using the 2D CNNs. Different CNN architectures and model training tweaks were used. Two groups used conventional methods or handcrafted features in addition to the 2D CNNs for analysis of the projection images. Four groups tried 3D CNNs for direct analysis of the CT volumetric data. Two groups used per-slice analysis, and one of the groups performed additional partially manual adaptation of the lung-based labeling to the slice-based labeling.

All participants used some techniques for artificial data set enlargement and a few pre-processing steps, such as resizing, cropping, normalization, slice filtering or concatenations. Many groups used both of the provided lung masks, and the winning group used an additional custom lung segmentation to make the data pre-processing even more accurate. It should be noted that some groups did not utilize the lung-wise labelling advantage and processed the entire slices (or projections) containing both left and right lungs.

The overall improvement of the results, appearance of the new effective approaches, variability in network architectures and training schemes suggest that the future development and extension of the proposed task is reasonable and may introduce new valuable results. Possible updates for the future editions may include: (i) extending the number of lesion classes; (ii) inclusion of some kind of lesion location information, up to switching from binary classification to a detection/segmentation task; (iii) inclusion of some kind of lesion characteristic information, such as lesion size.

Acknowledgements

Data collection for the Tuberculosis task was supported by the National Institute of Allergy and Infectious Diseases, National Institutes of Health, US Department of Health and Human Services, CRDF project DAA9-19-65987-1 "Year 8: Belarus TB Database and TB Portals".

References

1. Kalpathy-Cramer, J., García Seco de Herrera, A., Demner-Fushman, D., Antani, S., Bedrick, S., Müller, H.: Evaluating performance of biomedical image retrieval systems: Overview of the medical image retrieval task at ImageCLEF 2004–2014. *Computerized Medical Imaging and Graphics* **39**(0) (2015) 55 – 61
2. Müller, H., Clough, P., Deselaers, T., Caputo, B., eds.: *ImageCLEF – Experimental Evaluation in Visual Information Retrieval*. Volume 32 of *The Springer International Series On Information Retrieval*. Springer, Berlin Heidelberg (2010)
3. García Seco de Herrera, A., Schaer, R., Bromuri, S., Müller, H.: Overview of the ImageCLEF 2016 medical task. In: *Working Notes of CLEF 2016 (Cross Language Evaluation Forum)*. (September 2016)
4. Müller, H., Clough, P., Hersh, W., Geissbuhler, A.: *ImageCLEF 2004–2005: Results experiences and new ideas for image retrieval evaluation*. In: *International Conference on Content-Based Multimedia Indexing (CBMI 2005)*, Riga, Latvia, IEEE (June 2005)
5. Ionescu, B., Müller, H., Péteri, R., Abacha, A.B., Datla, V., Hasan, S.A., Demner-Fushman, D., Kozlovski, S., Liauchuk, V., Cid, Y.D., Kovalev, V., Pelka, O., Friedrich, C.M., de Herrera, A.G.S., Ninh, V.T., Le, T.K., Zhou, L., Piras, L., Riegler, M., Halvorsen, P., Tran, M.T., Lux, M., Gurrin, C., Dang-Nguyen, D.T., Chamberlain, J., Clark, A., Campello, A., Fichou, D., Berari, R., Brie, P., Dogariu, M., Ştefan, L.D., Constantin, M.G.: Overview of the ImageCLEF 2020: Multimedia retrieval in medical, lifelogging, nature, and internet applications. In: *Experimental IR Meets Multilinguality, Multimodality, and Interaction*. Volume 12260 of *Proceedings of the 11th International Conference of the CLEF Association (CLEF 2020)*, Thessaloniki, Greece, LNCS Lecture Notes in Computer Science, Springer (September 22-25 2020)
6. Ionescu, B., Müller, H., Péteri, R., Dicente Cid, Y., 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*. Volume 2380 of *Proceedings of the 10th International Conference of the CLEF Association (CLEF 2019)*, Lugano, Switzerland, LNCS Lecture Notes in Computer Science, Springer (September 9-12 2019)
7. 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, Lugano, Switzerland, CEUR-WS.org <<http://ceur-ws.org>> (September 9-12 2019)
8. Ionescu, B., Müller, H., Villegas, M., de Herrera, A.G.S., Eickhoff, C., Andrearczyk, V., Dicente Cid, Y., Liauchuk, V., Kovalev, V., Hasan, S.A., Ling, Y., Farri, O., Liu, J., Lungren, M., Dang-Nguyen, D.T., Piras, L., Riegler, M., Zhou, L., Lux, M., Gurrin, C.: Overview of ImageCLEF 2018: Challenges, datasets and evaluation. In: Experimental IR Meets Multilinguality, Multimodality, and Interaction. Proceedings of the Ninth International Conference of the CLEF Association (CLEF 2018), Avignon, France, LNCS Lecture Notes in Computer Science, Springer (September 10-14 2018)
 9. Ionescu, B., Müller, H., Villegas, M., Arenas, H., Boato, G., Dang-Nguyen, D.T., Dicente Cid, Y., Eickhoff, C., Garcia Seco de Herrera, A., Gurrin, C., Islam, B., Kovalev, V., Liauchuk, V., Mothe, J., Piras, L., Riegler, M., Schwall, I.: Overview of ImageCLEF 2017: Information extraction from images. In: Experimental IR Meets Multilinguality, Multimodality, and Interaction 8th International Conference of the CLEF Association, CLEF 2017. Volume 10456 of Lecture Notes in Computer Science., Dublin, Ireland, Springer (September 11-14 2017)
 10. Villegas, M., Müller, H., Garcia Seco de Herrera, A., Schaer, R., Bromuri, S., Gilbert, A., Piras, L., Wang, J., Yan, F., Ramisa, A., Dellandrea, A., Gaizauskas, R., Mikolajczyk, K., Puigcerver, J., Toselli, A.H., Sanchez, J.A., Vidal, E.: General overview of ImageCLEF at the CLEF 2016 labs. In: CLEF 2016 Proceedings. Lecture Notes in Computer Science, Evora. Portugal, Springer (September 2016)
 11. Villegas, M., Müller, H., Gilbert, A., Piras, L., Wang, J., Mikolajczyk, K., Garcia Seco de Herrera, A., Bromuri, S., Amin, M.A., Kazi Mohammed, M., Acar, B., Uskudarli, S., Marvasti, N.B., Aldana, J.F., Roldán García, M.d.M.: General overview of ImageCLEF at the CLEF 2015 labs. In: Working Notes of CLEF 2015. Lecture Notes in Computer Science. Springer International Publishing (2015)
 12. Caputo, B., Müller, H., Thomee, B., Villegas, M., Paredes, R., Zellhofer, D., Goeau, H., Joly, A., Bonnet, P., Martinez Gomez, J., Garcia Varea, I., Cazorla, C.: ImageCLEF 2013: the vision, the data and the open challenges. In: Working Notes of CLEF 2013 (Cross Language Evaluation Forum). (September 2013)
 13. World Health Organization, et al.: Global tuberculosis report 2019. (2019)
 14. Dicente Cid, Y., Jimenez-del-Toro, O., Depeursinge, A., Müller, H.: Efficient and fully automatic segmentation of the lungs in CT volumes. In Orcun Goksel, Jimenez-del-Toro, O., Foncubierta-Rodriguez, A., Müller, H., eds.: Proceedings of the VISCERAL Challenge at ISBI. Number 1390 in CEUR Workshop Proceedings (Apr 2015) 31–35
 15. Liauchuk, V., Kovalev, V.: Imageclef 2017: Supervoxels and co-occurrence for tuberculosis CT image classification. In: CLEF2017 Working Notes. CEUR Workshop Proceedings, Dublin, Ireland, CEUR-WS.org <<http://ceur-ws.org>> (September 11-14 2017)
 16. Gadkowski, L.B., Stout, J.E.: Cavitory pulmonary disease. *Clinical Microbiology Reviews* **21**(2) (2008) 305—333
 17. Miron, R., Moisii, C., Breaban, M.E.: Revealing Lung Affections from CTs. A Comparative Analysis of Various Deep Learning Approaches for Dealing with Volumetric Data. In: CLEF2020 Working Notes. CEUR Workshop Proceedings, Thessaloniki, Greece, CEUR-WS.org <<http://ceur-ws.org>> (September 22-25 2020)
 18. Lu, X., Chang, E.Y., Liu, Z., Hsu, C.n., Du, J., Gentili, A.: ImageCLEF2020: Laterality-Reduction Three-Dimensional CBAM-Resnet with Balanced Sampler for Multi-Binary Classification of Tuberculosis and CT Auto Reports. In:

- CLEF2020 Working Notes. CEUR Workshop Proceedings, Thessaloniki, Greece, CEUR-WS.org <<http://ceur-ws.org>> (September 22-25 2020)
19. Che, J., Ding, H., Zhou, X.: Chejiao at ImageCLEFmed Tuberculosis 2020: CT report generation based on transfer learning. In: CLEF2020 Working Notes. CEUR Workshop Proceedings, Thessaloniki, Greece, CEUR-WS.org <<http://ceur-ws.org>> (September 22-25 2020)
 20. Liauchuk, V.: ImageCLEF 2019: Projection-based CT Image Analysis for TB Severity Scoring and CT Report Generation. In: CLEF2019 Working Notes. Volume 2380 of CEUR Workshop Proceedings., Lugano, Switzerland, CEUR-WS.org <<http://ceur-ws.org/Vol-2380>> (September 9-12 2019)
 21. He, T., Zhang, Z., Zhang, H., Zhang, Z., Xie, J., Li, M.: Bag of tricks for image classification with convolutional neural networks. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). (2019) 558–567
 22. Mossa, A.A., Eriş, H., Çevik, U.: Ensemble of Deep Learning Models for Automatic Tuberculosis Diagnosis Using Chest CT Scans: Contribution to the ImageCLEF-2020 Challenges. In: CLEF2020 Working Notes. CEUR Workshop Proceedings, Thessaloniki, Greece, CEUR-WS.org <<http://ceur-ws.org>> (September 22-25 2020)
 23. Asakawa, T., Aono, M.: ImageCLEF 2020: Deep Learning for Tuberculosis in Chest CT Image Analysis based on multi-axis projections. In: CLEF2020 Working Notes. CEUR Workshop Proceedings, Thessaloniki, Greece, CEUR-WS.org <<http://ceur-ws.org>> (September 22-25 2020)
 24. Waqas, M., Khan, Z., Anjum, S., Tahir, M.A.: Lung-Wise Tuberculosis Analysis and Automatic CT Report Generation with Hybrid Feature and Ensemble Learning. In: CLEF2020 Working Notes. CEUR Workshop Proceedings, Thessaloniki, Greece, CEUR-WS.org <<http://ceur-ws.org>> (September 22-25 2020)
 25. Coca, L.G., Hanganu, A., Cusmuluc, C.G., Iftene, A.: UAIC2020: Lung Analysis for Tuberculosis Detection. In: CLEF2020 Working Notes. CEUR Workshop Proceedings, Thessaloniki, Greece, CEUR-WS.org <<http://ceur-ws.org>> (September 22-25 2020)
 26. Balwal, U., Yeragudipati, S.A., Bhuvana, J., Mirnalinee, T.T.: Deep Learning based TB Severity Prediction. In: CLEF2020 Working Notes. CEUR Workshop Proceedings, Thessaloniki, Greece, CEUR-WS.org <<http://ceur-ws.org>> (September 22-25 2020)
 27. Lestyan, B., Benczúr, A.A., Daróczy, B.: SZTAKI @ ImageCLEFmed 2020 Tuberculosis Task. In: CLEF2020 Working Notes. CEUR Workshop Proceedings, Thessaloniki, Greece, CEUR-WS.org <<http://ceur-ws.org>> (September 22-25 2020)