Overview of ImageCLEFtuberculosis 2021 — CT-based Tuberculosis Type Classification

Serge Kozlovski[1], Vitali Liauchuk[1], Yashin Dicente Cid[2], Vassili Kovalev[1] and Henning Müller[3,4]

[1]*United Institute of Informatics Problems, Minsk, Belarus*

[2]*University of Warwick, Coventry, UK*

[3]*University of Applied Sciences Western Switzerland (HES–SO), Sierre, Switzerland*

[4]*University of Geneva, Switzerland*

Abstract

ImageCLEF is a part of the Conference and Labs of the Evaluation Forum (CLEF) initiative and includes a variety of tasks dedicated to multimodal image information retrieval, including image classification and annotation. The tuberculosis (TB) task is one of the ImageCLEF tasks which started in 2017 and changed from year to year. The 2021 edition was dedicated to the automatic classification of five TB types: Infiltrative, Focal, Tuberculoma, Miliary, Fibro-cavernous. The task itself repeated one of the original subtasks from 2017 but the dataset was significantly changed. In 2021, 11 groups from 9 countries participated in the task and submitted at least one successful run. The task results can be compared to the TB type classification task results in the 2017 and 2018 editions. Although top scores were not improved compared to the previous editions, the participants' results allow us to analyze the effectiveness of applying recent deep learning approaches to the task.

Keywords

Tuberculosis, Computed Tomography, Image Classification, Tuberculosis Type, 3D Data Analysis

1. Introduction

ImageCLEF[1] is a part of the the CLEF[2] initiative and presents a set of image information retrieval tasks. Medical tasks were included in the 2nd edition of ImageCLEF in 2004 and have been held every year since then [1, 2, 3, 4, 5]. The tuberculosis task is one of the medical tasks this year. More information on the other tasks organized in 2021 can be found in [6] and the past editions of ImageCLEF are described in [7, 8, 9, 10, 11, 12, 13].

Tuberculosis (TB) is a bacterial infection caused by a germ called Mycobacterium tuberculosis. About 130 years after its discovery, the disease remains a persistent threat and one of the top 10 causes of death worldwide according to the WHO [14]. The bacteria usually attack the lungs and generally TB can be cured with antibiotics. However, different types of TB require different treatments and therefore detection of the specific case characteristics is an important real-world task.

[1]http://www.imageclef.org/

[2]http://www.clef-initiative.eu/

In the previous editions of this task, the setup evolved from year to year. In the first two editions [15, 16] participants had to detect multi-drug resistant patients (MDR subtask) and to classify the TB type (TBT subtask) both based only on the computed tomography (CT) image. After 2 editions it was concluded to drop the MDR subtask because it seemed impossible to solve based only on the image, and the TBT subtask was also suspended 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 3d edition, the Tuberculosis task [17] was restructured to allow usage of the uniform dataset, and included two subtasks – a continued severity score (SVR) prediction subtask and a new subtask based on providing an automatic CT report on the TB case (CTR subtask).

In the 4th edition [18], the SVR subtask was dropped and the automated CT report generation task was modified to be lung-based rather than CT-based.

Because of the fairly high results achieved by the participants in the CTR task in 2020, we decided to discontinue the CTR task at the moment and switch to the task which was not yet solved with sufficiently high quality. So in this year's edition, it was decided to bring back to life the Tuberculosis Type classification task from the 1st and 2nd ImageCLEFmed Tuberculosis editions. The dataset was updated, extended in size and some additional information was added for a part of the CT scans.

We hoped that utilizing the newest deep learning approaches together with the available pre-trained models and additional data sets would allow the participants to achieve better results for the TB Type classification compared to the early editions of the task.

This article first describes the task proposed for TB in 2021. 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 Set, Evaluation, Participation

2.1. The Task in 2021

In this task, participants had to automatically categorize each TB case into one of the following five types: (1) Infiltrative, (2) Focal, (3) Tuberculoma, (4) Miliary, (5) Fibro-cavernous. So the task is a multi-class classification problem.

2.2. Data Set

In this edition, a data set containing chest CT scans of 1,338 TB patients was used: 917 images for the training (development) data set and 421 for the test set. Each CT image corresponded to only one TB type and to one unique patient.

Additional meta-information containing CT-report in the 2020 edition format was provided for 243 cases. Since this meta-information may be potentially used as a target label in future task editions, it was provided only for a subset of train images. We expected participants would use the best approaches from the previous year task edition to generate CT report for all train and test cases, and then use this meta-information as an additional feature for TB type prediction.

For every patient, a 3D CT image series was provided with a slice 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.

Same as in the previous year, for all patients, we provided two versions of automatically extracted masks of the lungs obtained using the methods described in [19, 20, 18].

Typical examples of CTs with different TB types are shown in Fig. 1. Table 1 details the distribution of patients within each TB type. One can note an important unbalance in the label numbers caused by natural reasons. During the data split, we tried to achieve distribution similarity between the training and the testing data.

Table 1
Distribution of CT images within each class.

Set	Infiltrative	Focal	Tuberculoma	Miliary	Fibro-cavernous
Train	420 (46%)	226 (24%)	101 (11%)	100 (11%)	70 (8%)
Test	169 (40%)	86 (21%)	64 (15%)	60 (14%)	42 (10%)

2.3. Evaluation Measures and Scenario

Similar to the previous editions, each participating group could submit up to 10 runs in total.

The task was evaluated as a multi-class classification problem and scored using unweighted Cohen's Kappa coefficient and accuracy metrics. The ranking of this task was done first by Kappa and then by accuracy.

2.4. Participation

In 2021, there were 78 registered teams and 29 signed the end-user agreement. 11 groups from 9 countries participated and submitted results. The number of submissions is a bit higher than in 2020. Table 2 shows the list of participants and their institutions.

3. Results

To perform a ranking for this task we used the Cohen's Kappa coefficient as primary score and accuracy as secondary score. Table 3 shows these two measures calculated for the best run submitted by each participating group. For the best run of each group, Figures 2 and 3 show the corresponding confusion matrices and true positive rate for each TB type.

SenticLab.UAIC [21] is the winner of the task with a Kappa score of 0.221 and an accuracy of 0.466. In their experiments, the SenticLab.UAIC team compared several approaches based on 2D and 3D CNNs. The winning method was based on the slice-wise application of an EfficientNet-B4 2D CNN.
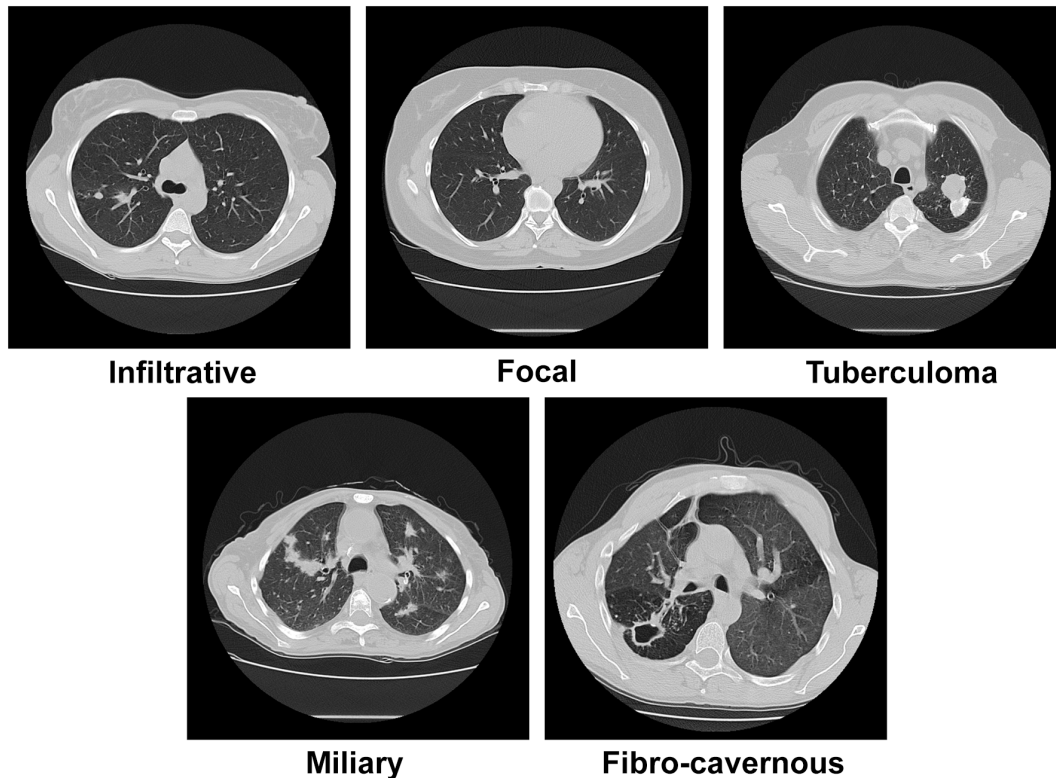
Figure 1: Slices of typical CT images with different TB types.

The hasibzunair [22] team ranked 2nd in terms of both Kappa and accuracy. The team approach was based on the usage of a hybrid 2D CNN-RNN model. The team experiments included extensive usage of transfer learning techniques and a custom loss function.

The SDVA-UCSD [23] team selected volumetric analysis and used 3D CNN models for the CT analysis. The team's best solution was found using a 3D ResNet34 with convolutional block attention.

The Emad_Aghajanzadeh team in their work [24] experimented with different approaches, including slice-wise analysis using a 2D CNN and a hybrid 2D CNN + RNN model, volume-based analysis using a 3D CNN and a hybrid 3D CNN + RNN model, and also a hybrid use of 2D + 3D CNN + RNN models. The team's best result was achieved using a custom 3D CNN.

The MIDL-NCAI-CUI team used slice-wise analysis using a pre-trained 2D CNN and ended up with the EfficientNet-B0 model.

The uaic2021 [25] team tried both 2D projection-based and volumetric approaches. They finished up with 3D MedicalNet10 for the best run.

The IALab_PUC [26] team tested several pipelines based on different 2D CNNs, including custom one and pre-trained DenseNet121.

The KDE-lab [27] team used slice-wise analysis using the 2D CNN model (EfficientNet-B5).

The JBTTM [28] team tried to use 3D CNNs but failed and used a simple shallow neural

Table 2
List of participating teams that submitted at least one run.

Group name	Main institution	Country
Emad_Aghajanzadeh	Ferdowsi University of Mashhad	Iran
hasibzunair	Concordia University	Canada
IALab_PUC	IALab PUC	Chili
JBTTM	SSN College of Engineering	India
KDE-lab	Toyohashi University	Japan
MIDL-NCAI-CUI	COMSATS University Islamabad	Pakistan
SenticLab.UAIC	Alexandru Ioan Cuza University of Iasi	Romania
SDVA-UCSD	San Diego VA/UCSD	USA
uaic2020	Alexandru Ioan Cuza University of Iasi	Romania
YNUZHOU	Yunnan University	China
Zhao_Shi_	Yunnan University	China

Table 3
Summary on the participant submissions and their results.

Group rank	Group name	# of runs	Kappa	Accuracy	Rank of the best run
1	SenticLab.UAIC	10	**0.221**	**0.466**	1
2	hasibzunair	8	0.200	0.423	4
3	SDVA-UCSD	7	0.190	0.371	8
4	Emad_Aghajanzadeh	10	0.181	0.404	11
5	MIDL-NCAI-CUI	5	0.140	0.333	23
6	uaic2020	3	0.129	0.333	28
7	IALab_PUC	3	0.120	0.401	30
8	KDE-lab	10	0.117	0.382	31
9	JBTTM	1	0.038	0.221	42
10	Zhao_Shi_	5	0.015	0.380	47
11	YNUZHOU	1	-0.008	0.385	55

network for the final submission.

The Zhao_Shi_ [29] team used slice-wise analysis using pre-trained EfficientNet-B0 2D CNN and ghost modules.

4. Discussion and Conclusions

The results obtained in the task can be compared to the original TBT subtasks presented in the 2017 [15] and 2018 [16] edition. Before comparison, we should note, that although the task setup is the same in both editions, the data set was significantly changed, which means participants needed to deal with different images and label distributions, so the achieved scores can not be compared directly.

Top scores in the 2017, 2018 and 2021 editions are fairly close (Table 4). The best score of 2021

Table 4
Kappa and accuracy scores achieved in 2017, 2018 and 2021 task editions. Colored entries correspond to groups that participated in more than one edition.

Group rank	2017	2018	2021
1	0.244 (0.403)	0.231 (0.423)	0.221 (0.466)
2	0.233 (0.387)	0.173 (0.353)	0.200 (0.423)
3	0.219 (0.407)	0.171 (0.385)	0.190 (0.371)
4	0.196 (0.390)	0.166 (0.379)	0.181 (0.404)
5	0.153 (0.343)	0.147 (0.338)	0.140 (0.333)
6	0.022 (0.240)	0.063 (0.274)	0.129 (0.333)
7	-	0.020 (0.259)	0.120 (0.401)
8	-	-0.002 (0.237)	0.117 (0.382)
9	-	-	0.038 (0.221)
10	-	-	0.015 (0.380)
11	-	-	-0.008 (0.385)

was achieved by SenticLab.UAIC group and is slightly worse than the best result of 2018 and 2017 in terms of Kappa score - 0.221 vs 0.231 (-0.01 drop) and 0.244 (-0.022 drop). On the other hand, four groups overcome the 2nd best result from 2018. The top-ranked groups in the 2017 edition achieved better scores than in 2018 and this year but the difference can be explained by a decrease in TB type balance, rather than a drop in the performance of the approaches effectiveness [15, 16]. Additionally, it is worth mentioning that the group SDVA-UCSD that participated in the 2018 and 2021 editions was able to improve their Kappa score from 0.147 to 0.190.

A detailed analysis of the participant predictions in Figures 2 and 3 demonstrate that many participants struggled with natural TB type unbalance, which resulted in observed overfitting to the most frequently presented classes. Only SDVA-UCSD and MIDL-NCAI-CUI were able to achieve better than random recall for all TB types. SenticLab.UAIC, hasibzunair and Emad_Aghajanzadeh were able to achieve better than random recall for all TB types except Tuberculoma.

Analyzing the participants' working notes papers we mentioned the variability of approaches and usage of modern machine learning techniques and methods. Thus, the top-3 groups used completely different approaches. In contrast to the 2017 and 2018 task editions, all participants used deep learning methods for CT analysis. The majority of the participants (eight groups) used 2D CNN to analyze either selected projections of CT images or all slices. Two of these groups further used the slice-wise features extracted by 2D CNN to train an RNN in order to extract inter-slice information. Four groups successfully tried to utilize 3D CNNs for whole CT analysis. Different neural network architectures and model training tweaks were used by the participants. The majority of the participants also used transfer learning techniques. All participants used some approaches for artificial data set enlargement and a few pre-processing steps, such as resizing, normalization, slice filtering etc.

We should mention that analysis of the participant approaches demonstrates that in some cases results can be improved if more attention is paid to common machine learning routines like

careful treatment of train-validation label distribution and accurate CT data pre-processing (for example, a few groups ignored the provided lung masks, which may affect model effectiveness). Unfortunately, none of the groups tried to utilize the provided partial CT report metadata, so its importance in this task remains an open question.

Possible updates for future editions of the TBT task should consider: (i) extending the additional meta-information for CT scans; (ii) including some kind of lesion location information to the data set.

Acknowledgements

References

[1] J. Kalpathy-Cramer, A. García Seco de Herrera, D. Demner-Fushman, S. Antani, S. Bedrick, H. Müller, Evaluating performance of biomedical image retrieval systems: Overview of the medical image retrieval task at ImageCLEF 2004–2014, Computerized Medical Imaging and Graphics 39 (2015) 55 – 61.

[2] H. Müller, P. Clough, T. Deselaers, B. Caputo (Eds.), ImageCLEF – Experimental Evaluation in Visual Information Retrieval, volume 32 of *The Springer International Series On Information Retrieval*, Springer, Berlin Heidelberg, 2010.

[3] A. García Seco de Herrera, R. Schaer, S. Bromuri, H. Müller, Overview of the ImageCLEF 2016 medical task, in: Working Notes of CLEF 2016 (Cross Language Evaluation Forum), 2016.

[4] H. Müller, P. Clough, W. Hersh, A. Geissbuhler, ImageCLEF 2004–2005: Results experiences and new ideas for image retrieval evaluation, in: International Conference on Content–Based Multimedia Indexing (CBMI 2005), IEEE, Riga, Latvia, 2005.

[5] T. Deselaers, T. M. Deserno, H. Müller, Automatic medical image annotation in ImageCLEF 2007: Overview, results, and discussion, Pattern Recognition Letters 29 (2008) 1988–1995.

---

[3]http://tbportals.niaid.nih.gov/

[6] B. Ionescu, H. Müller, R. Péteri, A. Ben Abacha, M. Sarrouti, D. Demner-Fushman, S. A. Hasan, S. Kozlovski, V. Liauchuk, Y. D. Cid, V. Kovalev, O. Pelka, A. G. S. de Herrera, J. Jacutprakart, C. M. Friedrich, R. Berari, A. Tauteanu, D. Fichou, P. Brie, M. Dogariu, L. D. Ştefan, M. G. Constantin, J. Chamberlain, A. Campello, A. Clark, T. A. Oliver, H. Moustahfid, A. Popescu, J. Deshayes-Chossart, Overview of the ImageCLEF 2021: Multimedia retrieval in medical, nature, internet and social media applications, in: Experimental IR Meets Multilinguality, Multimodality, and Interaction, Proceedings of the 12th International Conference of the CLEF Association (CLEF 2021), LNCS Lecture Notes in Computer Science, Springer, Bucharest, Romania, 2021.

[7] B. Ionescu, H. Müller, R. Péteri, A. B. Abacha, V. Datla, S. A. Hasan, D. Demner-Fushman, S. Kozlovski, V. Liauchuk, Y. D. Cid, V. Kovalev, O. Pelka, C. M. Friedrich, A. G. S. de Herrera, V.-T. Ninh, T.-K. Le, L. Zhou, L. Piras, M. Riegler, P. l Halvorsen, M.-T. Tran, M. Lux, C. Gurrin, D.-T. Dang-Nguyen, J. Chamberlain, A. Clark, A. Campello, D. Fichou, R. Berari, P. Brie, M. Dogariu, L. D. Ştefan, M. G. Constantin, 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)*, LNCS Lecture Notes in Computer Science, Springer, Thessaloniki, Greece, 2020.

[8] B. Ionescu, H. Müller, R. Péteri, Y. Dicente Cid, V. Liauchuk, V. Kovalev, D. Klimuk, A. Tarasau, A. B. Abacha, S. A. Hasan, V. Datla, J. Liu, D. Demner-Fushman, D.-T. Dang-Nguyen, L. Piras, M. Riegler, M.-T. Tran, M. Lux, C. Gurrin, O. Pelka, C. M. Friedrich, A. G. S. de Herrera, N. Garcia, E. Kavallieratou, C. R. del Blanco, C. C. Rodríguez, N. Vasillopoulos, K. Karampidis, J. Chamberlain, A. Clark, A. Campello, 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)*, LNCS Lecture Notes in Computer Science, Springer, Lugano, Switzerland, 2019.

[9] B. Ionescu, H. Müller, M. Villegas, A. G. S. de Herrera, C. Eickhoff, V. Andrearczyk, Y. Dicente Cid, V. Liauchuk, V. Kovalev, S. A. Hasan, Y. Ling, O. Farri, J. Liu, M. Lungren, D.-T. Dang-Nguyen, L. Piras, M. Riegler, L. Zhou, M. Lux, C. Gurrin, 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), LNCS Lecture Notes in Computer Science, Springer, Avignon, France, 2018.

[10] B. Ionescu, H. Müller, M. Villegas, H. Arenas, G. Boato, D.-T. Dang-Nguyen, Y. Dicente Cid, C. Eickhoff, A. Garcia Seco de Herrera, C. Gurrin, B. Islam, V. Kovalev, V. Liauchuk, J. Mothe, L. Piras, M. Riegler, I. Schwall, 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*, Springer, Dublin, Ireland, 2017.

[11] M. Villegas, H. Müller, A. Garcia Seco de Herrera, R. Schaer, S. Bromuri, A. Gilbert, L. Piras, J. Wang, F. Yan, A. Ramisa, A. Dellandrea, R. Gaizauskas, K. Mikolajczyk, J. Puigcerver, A. H. Toselli, J.-A. Sanchez, E. Vidal, General overview of ImageCLEF at the CLEF 2016 labs, in: CLEF 2016 Proceedings, Lecture Notes in Computer Science, Springer, Evora.

Portugal, 2016.

[12] M. Villegas, H. Müller, A. Gilbert, L. Piras, J. Wang, K. Mikolajczyk, A. García Seco de Herrera, S. Bromuri, M. A. Amin, M. Kazi Mohammed, B. Acar, S. Uskudarli, N. B. Marvasti, J. F. Aldana, M. d. M. Roldán García, General overview of ImageCLEF at the CLEF 2015 labs, in: Working Notes of CLEF 2015, Lecture Notes in Computer Science, Springer International Publishing, 2015.

[13] B. Caputo, H. Müller, B. Thomee, M. Villegas, R. Paredes, D. Zellhofer, H. Goeau, A. Joly, P. Bonnet, J. Martinez Gomez, I. Garcia Varea, C. Cazorla, ImageCLEF 2013: the vision, the data and the open challenges, in: Working Notes of CLEF 2013 (Cross Language Evaluation Forum), 2013.

[14] World Health Organization, et al., Global tuberculosis report 2019 (2019).

[15] Y. Dicente Cid, A. Kalinovsky, V. Liauchuk, V. Kovalev, , H. Müller, Overview of ImageCLEF-tuberculosis 2017 - predicting tuberculosis type and drug resistances, in: CLEF2017 Working Notes, CEUR Workshop Proceedings, CEUR-WS.org <http://ceur-ws.org>, Dublin, Ireland, 2017.

[16] Y. Dicente Cid, V. Liauchuk, V. Kovalev, , H. Müller, Overview of ImageCLEFtuberculosis 2018 - detecting multi-drug resistance, classifying tuberculosis type, and assessing severity score, in: CLEF2018 Working Notes, CEUR Workshop Proceedings, CEUR-WS.org <http://ceur-ws.org>, Avignon, France, 2018.

[17] Y. Dicente Cid, V. Liauchuk, D. Klimuk, A. Tarasau, V. Kovalev, H. Müller, 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, 2019.

[18] S. Kozlovski, V. Liauchuk, Y. Dicente Cid, A. Tarasau, V. Kovalev, H. Müller, Overview of ImageCLEFtuberculosis 2020 - automatic CT-based report generation, in: CLEF2020 Working Notes, CEUR Workshop Proceedings, CEUR-WS.org <http://ceur-ws.org>, Thessaloniki, Greece, 2020.

[19] Y. Dicente Cid, O. Jimenez-del-Toro, A. Depeursinge, H. Müller, Efficient and fully automatic segmentation of the lungs in CT volumes, in: O. Orcun Goksel, Jimenez-del-Toro, A. Foncubierta-Rodriguez, H. Müller (Eds.), Proceedings of the VISCERAL Challenge at ISBI, number 1390 in CEUR Workshop Proceedings, 2015, pp. 31–35.

[20] V. Liauchuk, V. Kovalev, 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, 2017.

[21] C. Moisii, R. Miron, M. E. Breaban, Identifying tuberculosis type in CTs, in: CLEF2021 Working Notes, CEUR Workshop Proceedings, CEUR-WS.org <http://ceur-ws.org>, Bucharest, Romania, 2021.

[22] H. Zunair, A. Rahman, N. Mohammed, ViPTT-Net: Video pretraining of spatio-temporal model for tuberculosis type classification from chest CT scans, in: CLEF2021 Working Notes, CEUR Workshop Proceedings, CEUR-WS.org <http://ceur-ws.org>, Bucharest, Romania, 2021.

[23] X. Lu, E. Y. Chang, C.-n. Hsu, J. Du, A. Gentili, Multi-Classification Study of the Tuberculosis with 3D CBAM-ResNet and EfficientNet, in: CLEF2021 Working Notes, CEUR Workshop Proceedings, CEUR-WS.org <http://ceur-ws.org>, Bucharest, Romania, 2021.

[24] E. Aghajanzadeh, B. Shomali, D. Aminshahidi, N. Ghassemi, Classification of Tuberculosis Type on CT Scans of Lungs using a fusion of 2D and 3D Deep Convolutional Neural Networks, in: CLEF2021 Working Notes, CEUR Workshop Proceedings, CEUR-WS.org <http://ceur-ws.org>, Bucharest, Romania, 2021.

[25] A. Hanganu, C. Simionescu, L.-G. Coca, A. Iftene, UAIC2021: Lung Analysis for Tuberculosis Classification, in: CLEF2021 Working Notes, CEUR Workshop Proceedings, CEUR-WS.org <http://ceur-ws.org>, Bucharest, Romania, 2021.

[26] J. M. Quintana, D. Florea, R. Deane, D. Parra, P. Pino, P. Messina, H. Lobel, PUC Chile team at TBT Task: Diagnosis of Tuberculosis Type using segmented CT scans, in: CLEF2021 Working Notes, CEUR Workshop Proceedings, CEUR-WS.org <http://ceur-ws.org>, Bucharest, Romania, 2021.

[27] T. Asakawa, R. Tsuneda, K. Shimizu, T. Komoda, M. Aono, ImageCLEF 2021: Deep categorizing tuberculosis cases using normalization and pseudo-color CT image, in: CLEF2021 Working Notes, CEUR Workshop Proceedings, CEUR-WS.org <http://ceur-ws.org>, Bucharest, Romania, 2021.

[28] U. Balwal, S. A. Yeragudipati, J. Bhuvana, T. T. Mirnalinee, Simple Neural Network based TB Classification, in: CLEF2021 Working Notes, CEUR Workshop Proceedings, CEUR-WS.org <http://ceur-ws.org>, Bucharest, Romania, 2021.

[29] J. Li, L. Yang, B. Yang, Lijie at ImageCLEFmed Tuberculosis 2021: EfficientNet Simplified Tuberculosis Case Classification, in: CLEF2021 Working Notes, CEUR Workshop Proceedings, CEUR-WS.org <http://ceur-ws.org>, Bucharest, Romania, 2021.
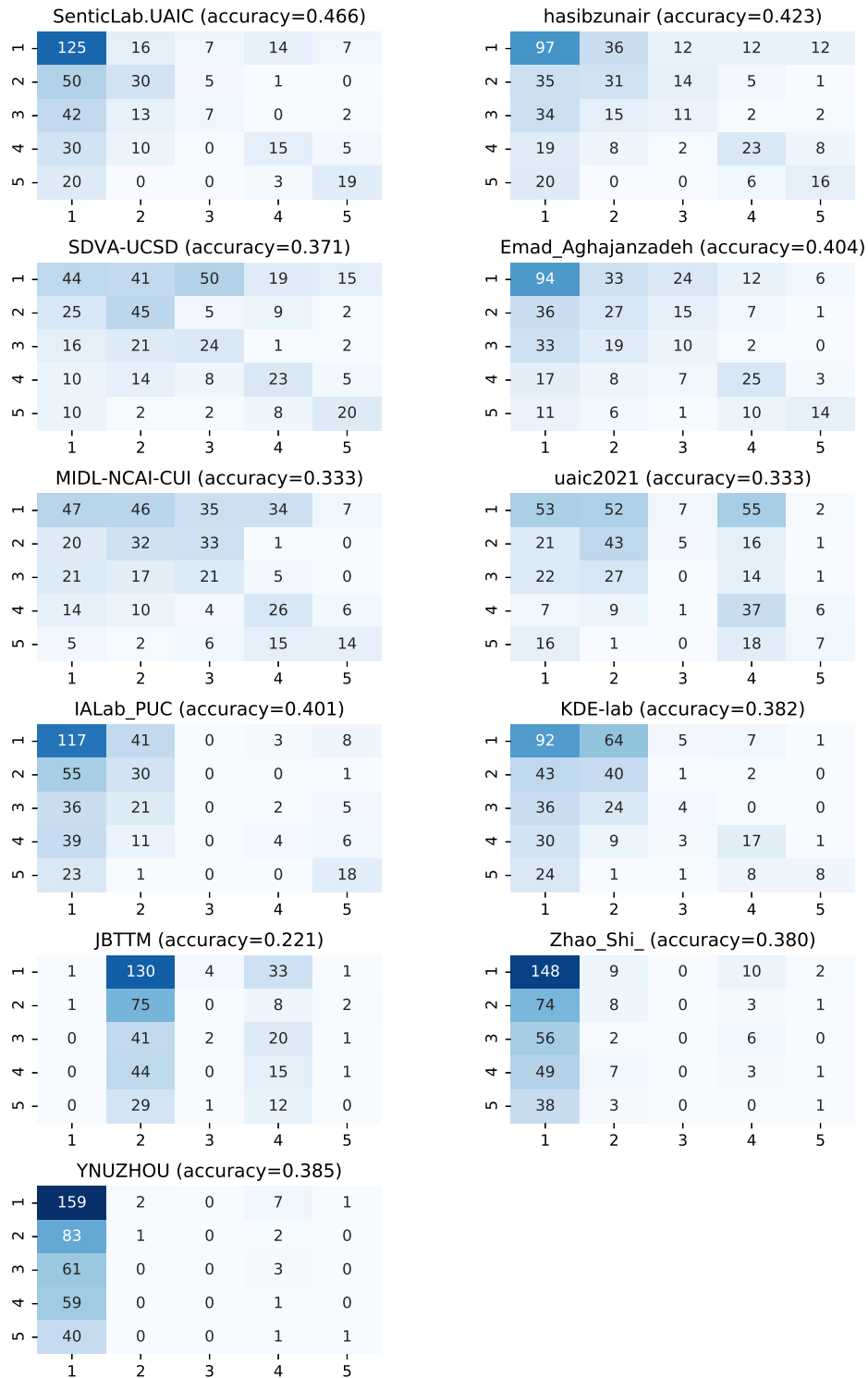
Figure 2: Confusion matrices obtained by the best run of each group. Vertical axes – true TB type, horizontal axes – predicted TB type.
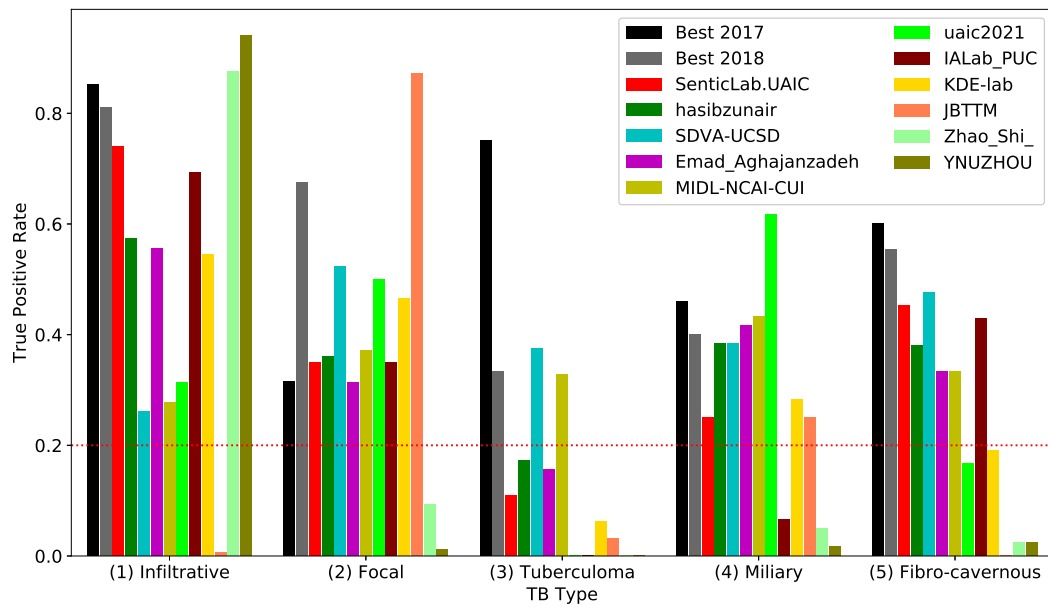
Figure 3: True positive rates obtained by the best run of each group.