ImageCLEF 2021: Deep categorizing tuberculosis cases using normalization and pseudo-color CT image

Tetsuya Asakawa[1], Riku Tsuneda[2], Kazuki Shimizu[3], Takuyuki Komoda[4] and Masaki Aono[5]

[1,2,5] *Toyohashi University of Technology, 1-1 Hibarigaoka Tenpaku, Toyohashi, Aichi, Japan*
[3,4] *Toyohashi Heart center, 21-1 Gobutori Tenpaku, Oyama, Toyohashi, Aichi, Japan*

Abstract

The ImageCLEF 2021 Tuberculosis task is an example of a challenging research problem in the field of computed tomography (CT) image analysis. The purpose of this study is to make accurate estimates for five labels (infiltrative, focal, tuberculoma, miliary, and fibrocavernous) based on lung images. We describe the tuberculosis task and approach for chest CT image analysis and then perform a single-label CT image analysis using the task dataset. We propose an image processing and fine-tuning deep neural network model that uses inputs from convolutional neural network features. This paper presents several approaches for applying normalization and pseudo-color to the extracted 2D images, for applying mask data to the extracted 2D image data, and for extracting a set of 2D projection images based on the 3D chest CT data. Our submissions for the task test dataset achieved an unweighted Cohen's kappa of 0.117 and an accuracy of 0.382.

Keywords

Tuberculosis, Deep Learning, Normalization, Pseudo-color

1. Introduction

With the spread of various diseases (e.g., tuberculosis (TB), COVID-19, and influenza), medical research has been performed to develop and implement the necessary treatments for viruses. However, there is no method currently available to identify such diseases early. An early diagnosis method is needed to provide the necessary treatment, develop specific medicines, and prevent the deaths of patients.

Accordingly, a significant amount of effort has been invested in medical image analysis research in recent years. In fact, a task dedicated to TB has been adopted as part of the ImageCLEF evaluation campaign for the five last years [1][2][3][4][5]. In ImageCLEF 2021 the main task [6], "ImageCLEFmed Tuberculosis," is treated as a computed tomography (CT) report. The goal of this subtask is to automatically categorize each TB case into one of the following five types: infiltrative, focal, tuberculoma, miliary, or fibrocavernous. Accordingly, the goal of this study is to automatically categorize the TB type from 3D CT images of TB patients.

In this paper, we employ a new fine-tuning neural network model that uses features extracted by pre-trained convolutional neural network (CNN) models as input. The existing CNN model had weak classifications; therefore, we propose a new fully connected two layers. The new contributions of this paper are the proposition of novel feature building techniques, the incorporation of features from the proposed CNN model, and the use of several forms of pre-processing to predict TB from the images. In Section 2, we describe the conducted task and the ImageCLEF2021 dataset. In Section 3, we

introduce the image pre-processing, experimental settings, and features used in this study. In Section 4, we describe the experiments we performed. In Section 5, we provide our conclusions.

2. ImageCLEF 2021 Dataset

The TB task of the ImageCLEF 2021 Challenge included partial 3D patient chest CT images [7]. The dataset contained the chest CT scan imaging data, including 917 images for the training (development) dataset and 421 images for the test dataset. Some of the scans include additional meta-information, which may vary depending on data availability for different cases. Each CT image corresponds to only one TB type. In this edition, each CT scan corresponds to one patient. Using the CT image data, our goal is to automatically extract and categorize each TB case into one of the following five types: (1) Infiltrative, (2) Focal, (3) Tuberculoma, (4) Miliary, (5) fibrocavernous Table 1 lists the labels for the chest CT scan in the training dataset.

Table 1
Presence of labels for the chest CT scan in the training dataset.

Label	In Training set (number of patients)
Infiltrative	420
Focal	226
Tuberculoma	101
Miliary	100
fibrocavernous	70
Total	917

3. Proposed Method

We propose a single-label analysis system to predict the TB type from CT scan images. The first step is input data pre-processing. After introducing our pre-processing of the input data, we describe our deep neural network model, which enables single-label outputs given the CT scan images. In addition, optionally in the first step, we can use a CT scan movie instead of CT scan images. We detail our proposed system in the following subsections.

3.1. Input data pre-processing

The 3D CT scans in the training and test datasets are provided in compressed Nifti format. We decompressed the files and extracted the slices along the z-axis of the 3D image, as shown in fig. 1. For each Nifti image, we obtained a number of slices, according to the dimensions, ranging from 110 to 250 images for the z-dimension. After extracting the slices along the z-axis, we filtered the slices of each patient using mask1 and mask2 data [8][9]. The mask1 data provide more accurate masks but tend to miss large abnormal regions of the lungs in the most severe TB cases. The mask2 data provide more rough bounds but behave more stably in terms of including lesion areas. We extracted the filtered CT scan images. We noticed that all slices contain relevant information, including bone, space, fat, and skin, in addition to the lungs that could help classify the samples. This is why we added a step to the filter and selected a number of slices per patient. We call this data the **Applying mask CT data**.

In addition, as shown in fig. 2, we implemented pseudo-color on the normalization mask CT data. We call this data the **normalization mask CT data**.

In addition, as shown in fig. 3, we perform pseudo color for normalization mask CT data. We call this data **pseudo color CT data.**
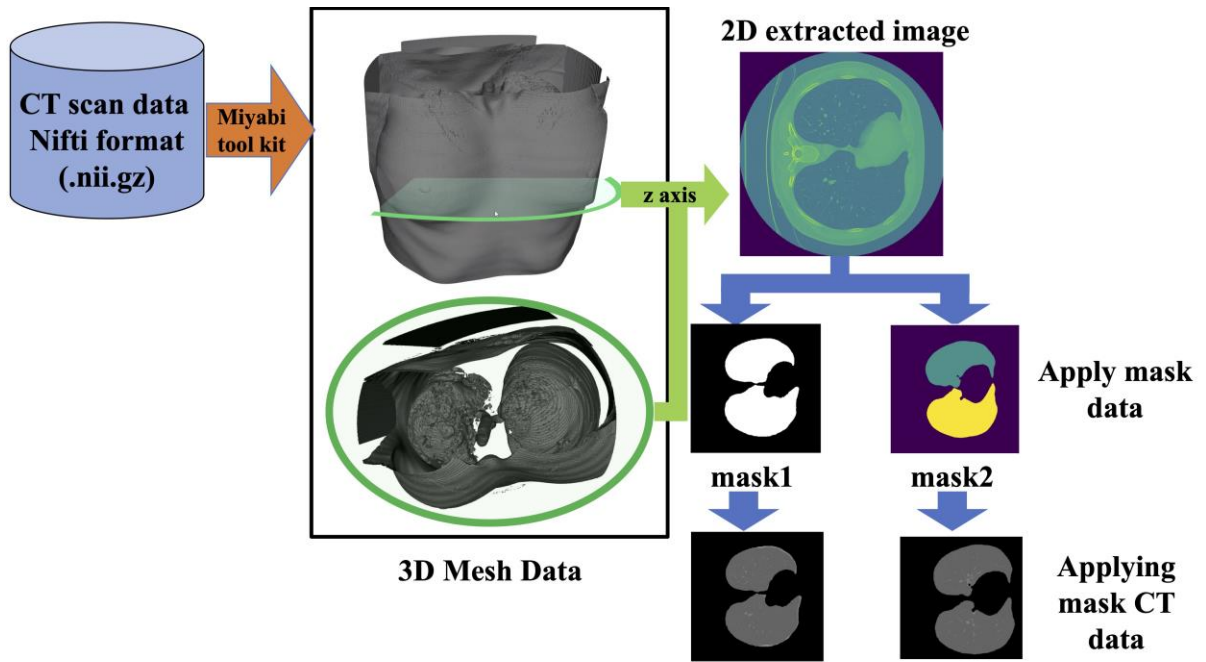
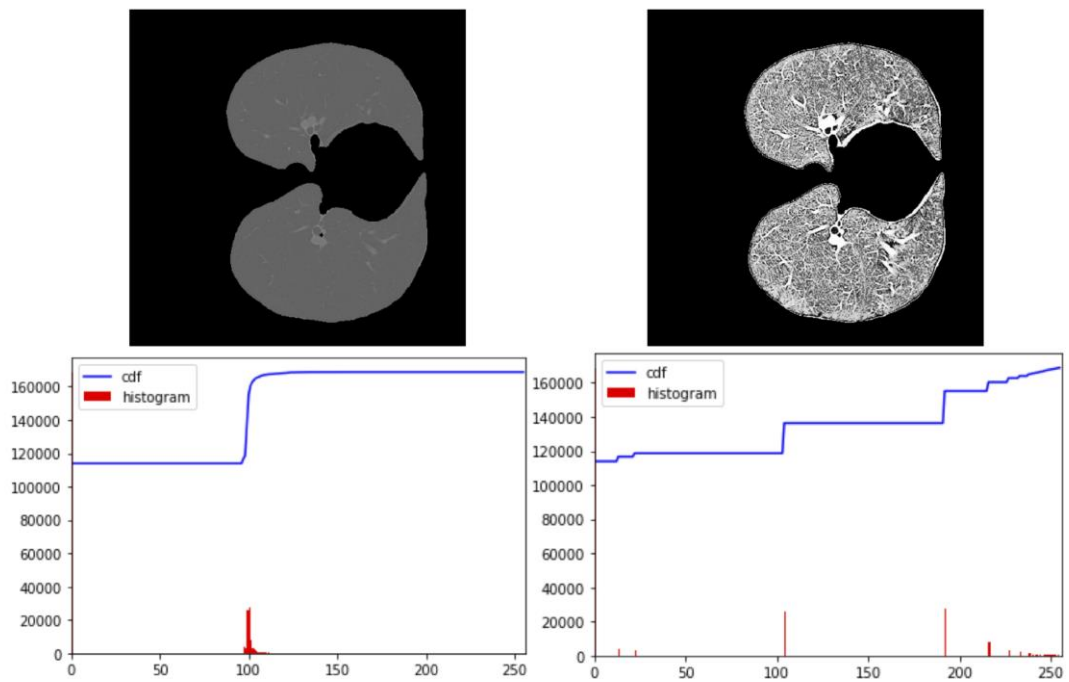**Figure 1**: Pre-processing of the input data applying mask data.

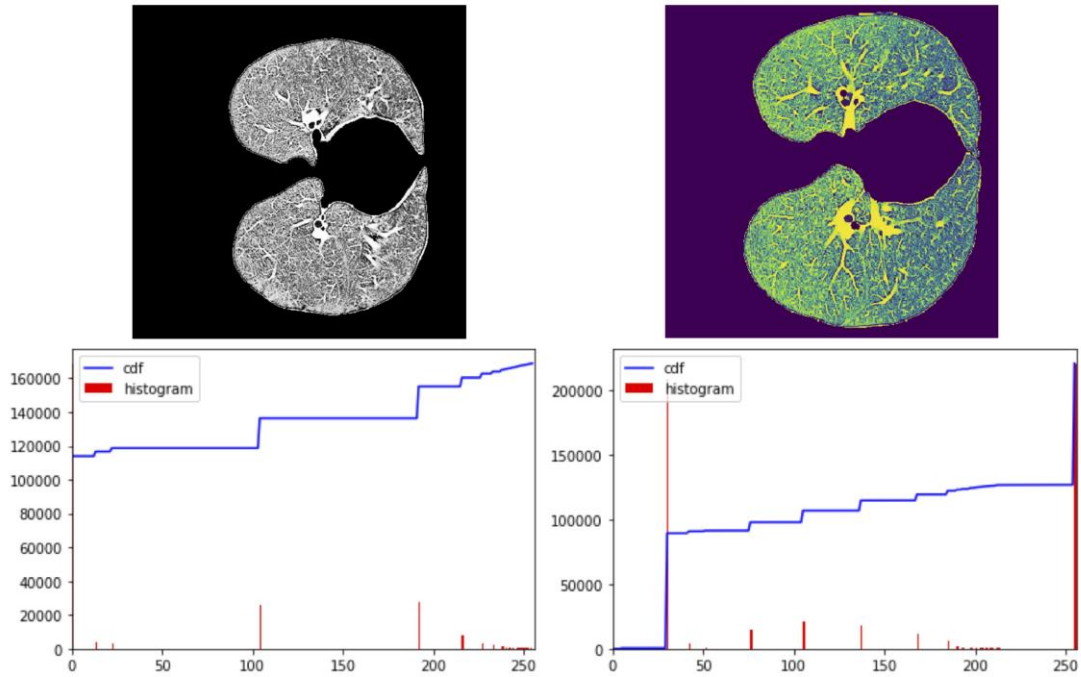**Figure 2**: Pre-processing of the input data using normalization.

Figure 3: Pre-processing of the input data using pseudo color.

3.2. Proposed deep neural network model

To solve this single-label problem, we propose fine-tuning neural network models that allow inputs coming from end-to-end CNN features.
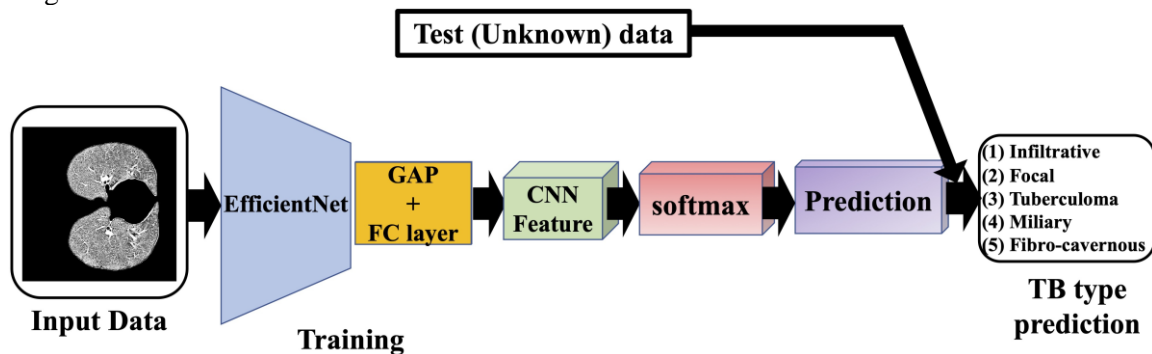

Figure 4: Our proposed method for feature extraction.

3.2.1. Training and Validation sets

The training dataset consists of 107,955 and 105,494 images extracted from the applying mask1 and mask2 CT datasets, respectively, for the z-axis.

We divided the training dataset at random into training and validation datasets with a ratio of 8:2. The CNN features were extracted using pre-trained CNN-based neural networks, including EfficientNet B05. To deal with the above features, we propose a deep neural network architecture.

Our system incorporates CNN features, which can be extracted using deep CNNs pre-trained on ImageNet [10] such as EffcientNet B05[11]. Because of the lack of datasets in visual sentiment analysis, we adopted transfer learning for the feature extraction to prevent overfitting. We decreased the dimensions of the fully connected layers used in the CNN models. In addition, we extracted the vector to 2048 dimensions.

3.2.2. Training and Validation sets and Test data

We employed the unweighted Cohen's kappa and accuracy to fine-tune the above CNN model.

As illustrated in fig. 4, the CNN features are combined and represented by an integrated feature as a linearly weighted average, where the weights are w3 for the CNN features. The CNN features are passed through "Fusion" processing to generate the integrated features, followed by a "softmax" activation function.

3.3.   Single-label probability

We propose the method illustrated in Algorithm 1. The input is a collection of features extracted from each image with K types of diseases, while the output is a K-dimensional hot vector.

In Algorithm 1, we assume that the extracted CNN features are represented by their probabilities. For each TB case, we sum the features, followed by the median of the result, which is denoted as T_i^k in Algorithm 1. In short, the vector S_i represents the output of each hot vector. We repeat this computation until all the test (unknown) images are processed.

4. Experiments
4.1.   Unweighted Cohen's Kappa and Accuracy of training and validation sets

The training dataset consists in Applying mask1 and mask2 CT data, and the normalization mask1 and mask2 CT data. The training dataset consists of 105 494,107 955 images extracted for the mask1 and mask2 CT data respectively.

Here, we have divided the filtering data into training and validation datasets with a ratio of 8:2. We determined the following hyper-parameters: the batch size is 256, the optimization function is stochastic gradient descent with a learning rate of 0.001 and a momentum of 0.9, and the number of epochs is 200. For the implementation, we employed Tensorflow[12] as our deep learning framework.

For the evaluation of the single-label classification, we employed the un-weighted Cohen's kappa and the accuracy. Table 2 shows the results. finally, we employed EfficientNet B05 for the training and validation datasets and the test data. The results are given in Section (4.2).

Table 2
Unweighted Cohen's Kappa and Accuracy of training and validation sets for fine-tuning EfficientNet B05.

Mask	Pre-processing	Unweighted Cohen's Kappa	Accuracy
	applying mask	0.213	0.443
mask1	applying mask and normalization	0.199	0.443
	applying mask, normalization, pseudo color	0.215	0.475
	applying mask	0.215	0.495
mask2	applying mask and normalization	0.244	0.489
	applying mask, normalization, pseudo color	0.183	0.448

Table 3
Numbers of images with the five labels for the chest CT scans in the training dataset.

Mask	Pre-processing	In Training set (number of images)
mask1	Infiltrative	49058
	Focal	25722
	Tuberculoma	11293
	Miliary	11692
	fibrocavernous	7729
mask2	Infiltrative	50035
	Focal	26203
	Tuberculoma	11552
	Miliary	12030
	fibrocavernous	8135

4.2. Results for the training and validation datasets and the test data using our proposed model

The test dataset consisted of 59 835 and 60 758 images extracted from the applying mask1 and mask2 CT data, respectively, as show in Table 3.

It is likely that our proposed models will give better results after more advanced data pre-processing including the use of several types of CT images and data augmentation. Here as described above, we employed fine-tuning CNN models in EfficientNet B05 based on several pre-processing methods.

Table 4 shows the results. Here, we compare the results in terms of the unweighted Cohen's kappa and the accuracy. For mask1 and normalization on fine-tuning EfficientNet B05, our proposed CNN model has good values of un-weighted Cohen's kappa and accuracy.

In addition, results of the other participants' submissions with their un-weighted Cohen's kappa and accuracy are shown in Table 5. Here, we compare the results in terms of the unweighted Cohen's kappa and the accuracy.

For our team, KDE-lab, our proposed CNN model has the best unweighted Cohen's kappa and accuracy.

The results achieved by our submissions are well ranked compared to those at the top of the list given in Table 5. Note that several runs in the table belong to the same teams and likely do not differ significantly. In terms of the unweighted Cohen's kappa, our model ranks 8th. In terms of the accuracy, our model ranks 7th.

Table 4
Results of experiments for single-label classification.

Mask	Pre-processing	Unweighted Cohen's Kappa	Accuracy
mask1	applying mask	0.016	0. 382
	applying mask and normalization	**0.117**	**0.382**
	applying mask, normalization, pseudo color	0.069	0.371
mask2	applying mask	0.015	0.372
	applying mask and normalization	0.085	0.375
	applying mask, normalization, pseudo color	0.081	0.373

Table 5
The best participants' runs submitted for the CTR subtask.

Group name	Rank	Unweighted Cohen's Kappa	Accuracy
SenticLab.UAIC	1	0.221	0.446
hasibzunair	2	0.200	0.423
SDVA-UCSD	3	0.190	0.371
Emad-Aghajanzadeh	4	0.181	0.333
MIDL-NCAI-CUI	5	0.140	0.333
uaic2021	6	0.129	0.401
IALab PUC	7	0.120	0.401
KDE-Lab	**8**	**0.117**	**0.381**
JBTTM	9	0.038	0.221
Zhao_Shi_	10	0.015	0.380
YNUZHOU	11	-0.08	0.385

5. Conclusions

In this study, we proposed image pre-processing and a CNN model for predicting five labels (infiltrative, focal, tuberculoma, miliary, and fibrocavernous) from chest CT images. We performed a lung CT image analysis in which we proposed a deep neural network model that enabled the inputs to be derived from the CNN features. To predict the five labels, we introduced a threshold-based single-label prediction algorithm.

Specifically, after training our deep neural network using the pre-processed images, we were able to predict the categories of the five types of TB cases from unknown CT scan images. The experimental results demonstrate that our proposed models out-perform some models in terms of the unweighted Cohen's kappa and the accuracy. For the unweighted Cohen's kappa, our model achieved a good value. As a consequence, we believe that using normalization to pre-process an image is effective.

In the future, given an arbitrary X-ray, CT, echo, or magnetic resonance imaging image might be included the optimal weights for the neural networks. Moreover, we hope our proposed model will encourage further research into the early detection of diseases (such as TB, COVID-19, and influenza) or unknown diseases.

6. Acknowledgment

7. References

[1]   Yashin Dicente Cid, Alexander Kalinovsky, Vitali Liauchuk, Vassili Kovalev, , and Henning M¨uller. Overview of ImageCLEFtuberculosis 2017 - predicting tubercu-losis type and drug resistances. In CLEF2017 Working Notes, CEUR Workshop Proceedings, Dublin, Ireland, September 11-14 2017. CEUR-WS.org <http://ceur-ws.org>.

[2]   Bogdan Ionescu, Henning M¨uller, Mauricio Villegas, Alba Garc´ıa Seco de Herrera, Carsten Eickhoff, Vincent Andrearczyk, Yashin Dicente Cid, Vitali Liauchuk, Vas-sili Kovalev, Sadid A. Hasan, Yuan Ling, Oladimeji Farri, Joey Liu, Matthew Lun-gren, Duc-Tien Dang-Nguyen, Luca Piras, Michael Riegler, Liting Zhou, Mathias Lux, and Cathal Gurrin. Overview of ImageCLEF 2018: Challenges, datasets and evaluation. In Experimental IR Meets

Multilinguality, Multimodality, and Interac-tion, Proceedings of the Ninth International Conference of the CLEF Association (CLEF 2018), Avignon, France, September 10-14 2018. LNCS Lecture Notes in Computer Science, Springer.

[3]   Bogdan Ionescu, Henning M¨uller, Renaud P´eteri, Yashin Dicente Cid, Vitali Li-auchuk, Vassili Kovalev, Dzmitri Klimuk, Aleh Tarasau, Asma Ben Abacha, Sa-did A. Hasan, Vivek Datla, Joey Liu, Dina Demner-Fushman, Duc-Tien Dang-Nguyen, Luca Piras, Michael Riegler, Minh-Triet Tran, Mathias Lux, Cathal Gur-rin, Obioma Pelka, Christoph M. Friedrich, Alba Garc´ıa Seco de Herrera, Narciso Garcia, Ergina Kavallieratou, Carlos Roberto del Blanco, Carlos Cuevas Rodr´ıguez, Nikos Vasillopoulos, Konstantinos Karampidis, Jon Chamberlain, Adrian Clark, and Antonio Campello. ImageCLEF 2019: Multimedia Retrieval in Medicine, Lifel-ogging, Security and Nature: Multimedia Retrieval in Medicine, Lifelogging, Secu-rity 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, September 9-12 2019. LNCS Lecture Notes in Computer Science, Springer.

[4]   Obioma Pelka, Christoph M Friedrich, Alba Garc´ıa Seco de Herrera, and Hen-ning M¨uller. Medical image understanding: Overview of the ImageCLEFmed 2020 concept prediction task. In CLEF2020 Working Notes, Workshop Proceedings, Thessaloniki, Greece, September 22-25 2020. CEUR-WS.org.

[5]   Serge Kozlovski, Vitali Liauchuk, Yashin Dicente Cid, Aleh Tarasau, Vassili Ko-valev, and Henning M¨uller. Overview of ImageCLEFtuberculosis 2021 - automatic CT-based report generation. In Overview of ImageCLEF tuberculosis 2021 - CT-based Tuberculosis Type Classification, CEUR Workshop Proceedings, Bucharest, Romania, September 21-24 2021. CEUR-WS.org <http://ceur-ws.org>.

[6]   Bogdan Ionescu, Henning M¨uller, Renaud P´eteri, Asma Ben Abacha, Vivek Datla, Sadid A. Hasan, Dina Demner-Fushman, Serge Kozlovski, Vitali Liauchuk, Yashin Dicente Cid, Vassili Kovalev, Obioma Pelka, Christoph M. Friedrich, Alba Garc´ıa Seco de Herrera, Van-Tu Ninh, Tu-Khiem Le, Liting Zhou, Luca Piras, Michael Riegler, P˚al Halvorsen, Minh-Triet Tran, Mathias Lux, Cathal Gurrin, Duc-Tien Dang-Nguyen, Jon Chamberlain, Adrian Clark, Antonio Campello, Dim-itri Fichou, Raul Berari, Paul Brie, Mihai Dogariu, Liviu Daniel S¸tefan, and Mi-hai Gabriel 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), Thessaloniki, Greece, September 22-25 2020. LNCS Lecture Notes in Computer Science, Springer.

[7]   Serge Kozlovski, Vitali Liauchuk, Yashin Dicente Cid, Aleh Tarasau, Vassili Ko-valev, and Henning M¨uller. Overview of ImageCLEFtuberculosis 2020 - auto-matic CT-based report generation. In CLEF2020 Working Notes, CEUR Work-shop Proceedings, Thessaloniki, Greece, September 22-25 2020. CEUR-WS.org <http://ceur-ws.org>.

[8]   Yashin Dicente Cid, Oscar Alfonso Jim´enez del Toro, Adrien Depeursinge, and Henning M¨uller. Efficient and fully automatic segmentation of the lungs in ct volumes. In Orcun Goksel, Oscar Alfonso Jim´enez del Toro, Antonio Foncubierta-Rodr´ıguez, and Henning M¨uller, editors, Proceedings of the VISCERAL Anatomy Grand Challenge at the 2015 IEEE ISBI, CEUR Workshop Proceedings, pages 31–35. CEUR-WS, May 2015.

[9]   Vitali Liauchuk and Vassili Kovalev. Imageclef 2017: Supervoxels and co-occurrence for tuberculosis CT image classification. In Linda Cappellato, Nicola Ferro, Lorraine Goeuriot, and Thomas Mandl, editors, Working Notes of CLEF 2017 - Conference and Labs of the Evaluation Forum, Dublin, Ireland, September11-14, 2017, volume 1866 of CEUR Workshop Proceedings. CEUR-WS.org, 2017.10.

[10]   Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexan-der C. Berg, and Li Fei-Fei. ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision (IJCV), 115(3):211–252, 2015.

[11]   Mingxing Tan and Quoc V. Le. Efficientnet: Rethinking model scaling for convo-lutional neural networks. ICML 2019, 05 2019.

[12]   Google. Tensorflow. https://github.com/tensorflow.