KDE-med-lab at ImageCLEF 2024: Identify data and detect generative models using CNN by lung segmentation based on U-net.

Notebook for the KDE-med-lab at CLEF 2024

Tetsuya Asakawa[1,*], Kazuki Shimizu[2], Kei Nomura[2] and Masaki Aono[1]

[1]Toyohashi University of Technology, 1-1 Hibarigaoka, Tempaku-cho, Toyohashi, Aichi, Japan, 441-8580
[2]Toyohashi Heart Center, 21-1 Gobutori, Oyama-cho, Toyohashi, Aichi, Japan, 441-8530

Abstract

CLEF 2024 ImageCLEF Gans Task is an example of the challenging research problem in the field of CT image analysis. The purpose of this research is to detect the synthetic biomedical image data to determine which real images were used in training to produce the generated images in the first subtask, and detect generated model in the second subtask. We propose lung images using segmentation based on U-net. And, we employ fine-tuning deep neural network model. In addition, the first subtask is using transfer learning at ImageCLEFmed GANS the training dataset for Task 1 and Task 2. Our submissions (KDE-lab team) on the task test dataset reached accuracy value of about 50.6% in the first subtask and reached ARI (Adjusted Rand Index) of about 0.226 in the second subtask.

Keywords

Tuberculosis, Deep Learning, Lung segmentation, U-net

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 seven last years[1], [2], [3], [4], [5], [6], [7], [8]. In ImageCLEF 2024 the main task [8], "ImageCLEFmed GANS," is treated as a computed tomography (CT) report. The goal of task is to detect the synthetic biomedical image data to determine which real images were used in training to produce the generated images in the first subtask, and detect generated model in the second subtask.

In this paper, we propose lung images using segmentation based on U-net. And, we employ fine-tuning deep neural network model by convolutional neural network (CNN) models or Vision Transformer (ViT). In addition, the first subtask is using transfer learning at ImageCLEF Gans the training dataset for Task 1 and Task 2.

The new contributions of this paper are the proposition of novel feature building techniques by lung segmentation based on U-net. In Section 2, we describe the conducted task and the ImageCLEFmed GANS 2024 dataset. In Section 3, we introduce the image experimental settings, and features used in this study. In Section 4, we describe the experiments we performed. In Section 5, we provide our conclusions.

CLEF 2024: Conference and Labs of the Evaluation Forum, September 09–12, 2024, Grenoble, France
*Corresponding author.
asakawa.tetsuya.um@tut.jp (T. Asakawa); shimizu@heart-center.or.jp (K. Shimizu); k.nomura@heart-center.or.jp (K. Nomura); masaki.aono.ss@tut.jp (M. Aono)
0000-0002-8345-7094 (T. Asakawa); 0009-0000-3448-7986 (K. Shimizu); 0000-0003-2838-7844 (K. Nomura); 0000-0003-1383-1076 (M. Aono)

2. ImageCLEFmed GANS 2024 Dataset

The Gans task of the ImageCLEF 2024 Challenge included partial 2D gray-scale chest CT images[9]. There are two subtasks. We describe as below.

2.1. First subtask-: dentify training data ”fingerprints”

The development dataset comprises data for two different generative models organized as follows: Model 1 (representing the ground truth for the test dataset of the previous edition) consists of 10k generated images and 200 images annotated as used/not used for training to generate the images. Specifically, 100 images were utilized for training, while the remaining 100 were not.

Model 2 consists of 10k generated images and 6k annotated images marked as used/not used for training to generate the images. Specifically, 3k images were utilized for training, while the remaining3k were not.

The test dataset has been structured the two subsets of real images have been mixed, with no disclosed proportion between unused and used ones. This edition, two generative models will be evaluated to study the similarity between the real and synthetic data, so two sets of generated and real images are provided as shown in Fig 1(a).

2.2. Second subtask: Detect generative models ”fingerprints”

The training dataset consists of 600 images generated using three different generative models. Each model is represented by 200 images of size 256x256 and are organized in annotated folders.

This test task involves working with a dataset comprising 3000 computed tomography (CT) slices, each sized at 256x256 pixels and grayscale. These slices were generated using four distinct generative models as shown in Fig 1(b).

3. Proposed Method

We propose lung images using segmentation based on U-net. And, we employ fine-tuning deep neural network model. In addition, the first subtask is using transfer learning at ImageCLEFmed GANS the training dataset for Task 1 and Task 2. We will detail our proposed system in the following section.

3.1. Lungs image applying mask based on U-net

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. We decompressed the files and extracted the mask only lung based on U-net[10], [11], as shown in Fig 2. In addition, we extracted Lungs image applying mask.

3.2. Proposed Method the first subtask

3.2.1. Two-stage transfer learning in the first subtask at Baseline

We propose fine-tuning deep neural network model that uses two-stage transfer learning. The first stage transfer learning uses ImageCLEF Gans training dataset1. The second stage transfer learning uses ImageCLEF Gans training dataset2. It shows Fig 3. We don’t use Lungs image applying mask in Baseline. And, We used to a deep neural network models of Densenet 121 as Baseline. In addition, CNN feature is passed out on ”K-means” processing to predict the two classes of features.

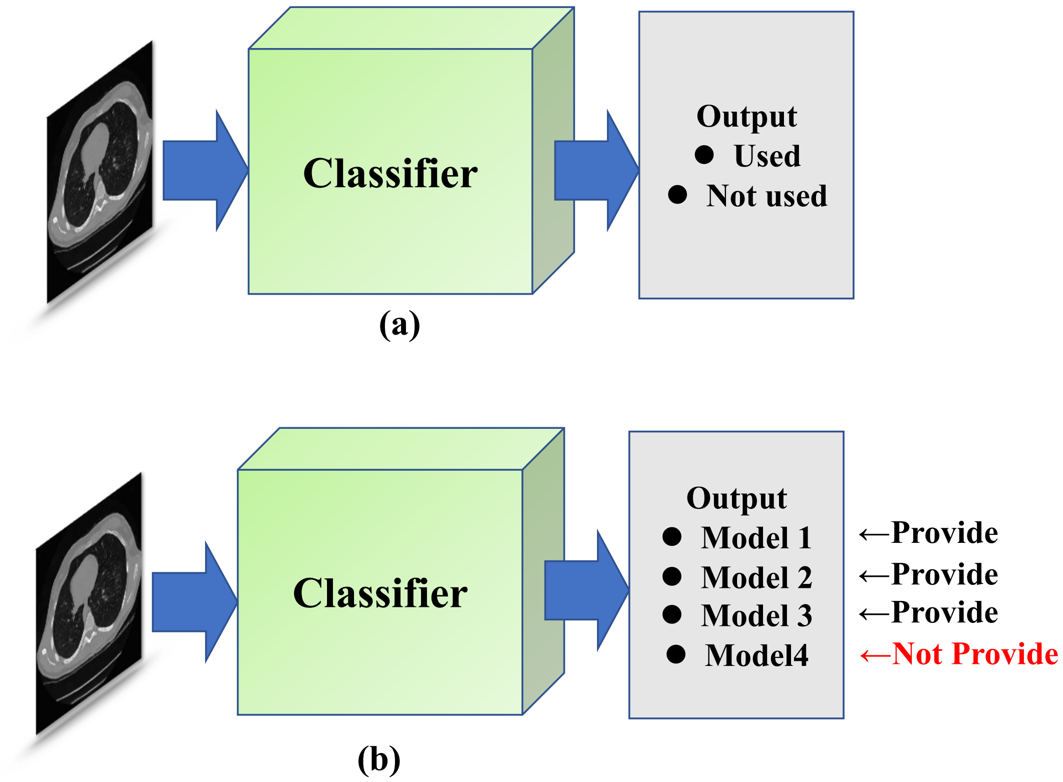

Figure 1: Overview of ImageCLEFmed GANS 2024 Tasks.

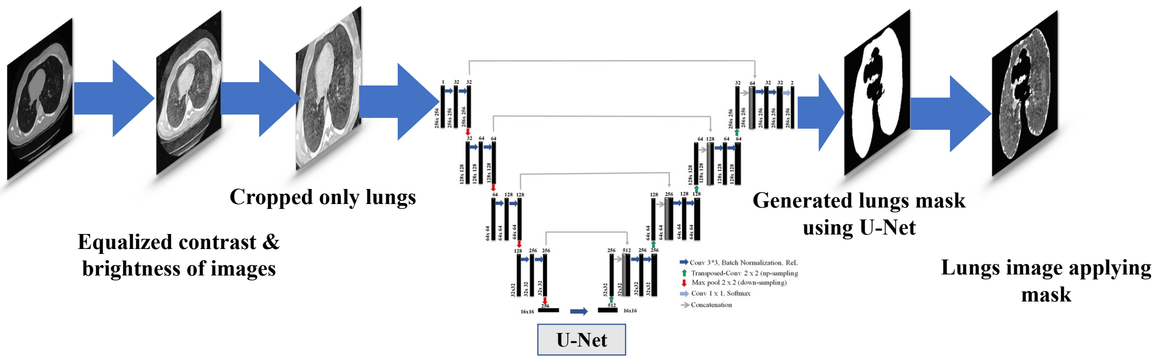

Figure 2: Lung images using segmentation based on U-net.

3.2.2. Two-stage transfer learning in the first subtask at Propose model

We propose fine-tuning deep neural network model that uses two-stage transfer learning. The first stage transfer learning uses ImageCLEF Gans training dataset1. The second stage transfer learning uses ImageCLEF Gans training dataset2. We use Lungs image applying mask using U-net in the Fig 4. We used to five deep neural network models: Swin Transformer, Densenet 121, Inception-Resnet V2, EfficientNetB03, and Inception V3. In addition, CNN feature is passed out on "K-means" processing to predict the two classes of features.

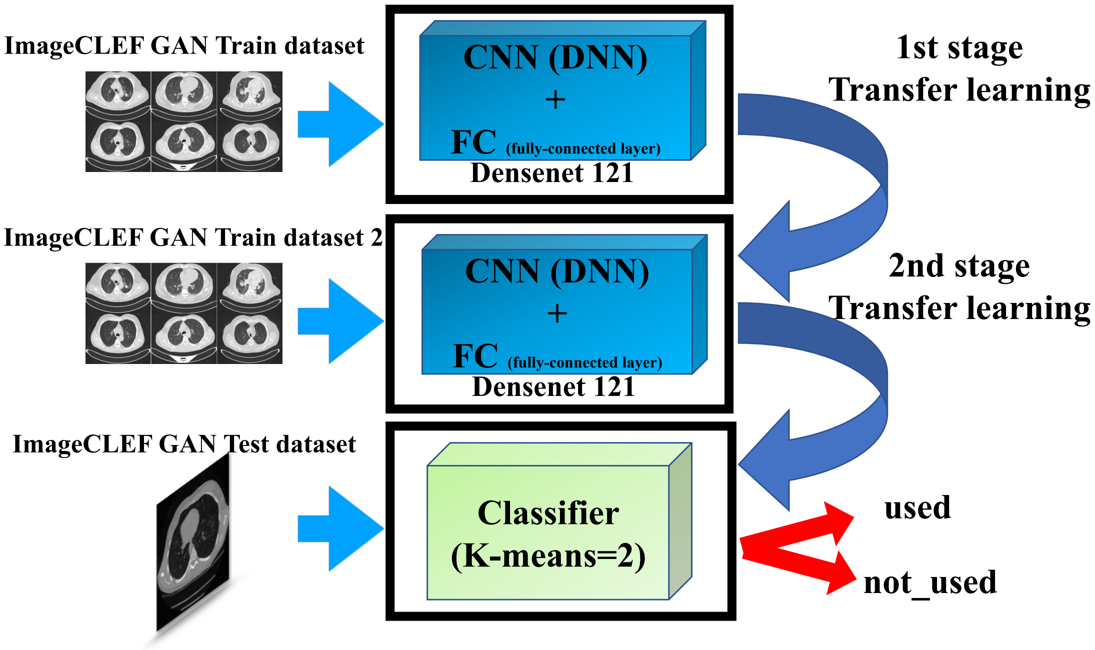

Figure 3: Overview of two-stage transfer learning at Baseline.

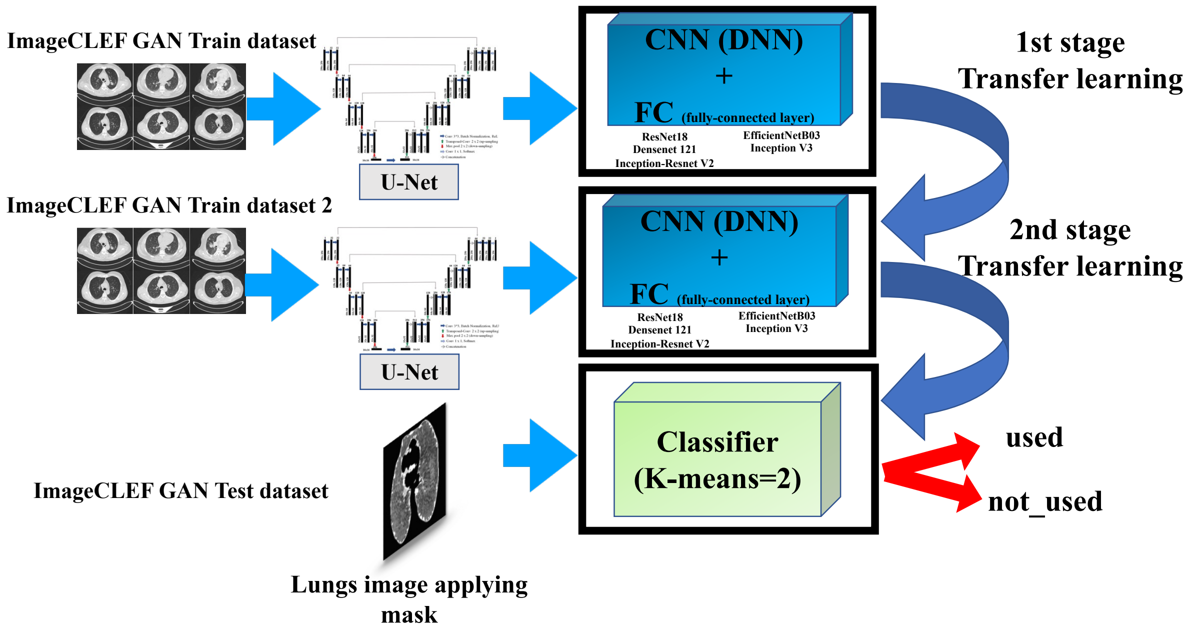

Figure 4: Overview of two-stage transfer learning at Proposed model.

3.3. Baseline and Proposed Method the second subtask

We propose fine-tuning deep neural network model. The second subtask is unsupervised learning. Therefore, we employed K-Means clustering that is the most popular unsupervised learning algorithm [12]. K-Means clustering is used to find intrinsic groups within the unlabelled dataset and draw inferences from them.

In Fig 5, it's shown that Baseline and Proposed model. Fig 5 (a) is Baseline, and Fig 5 (b) is Proposed

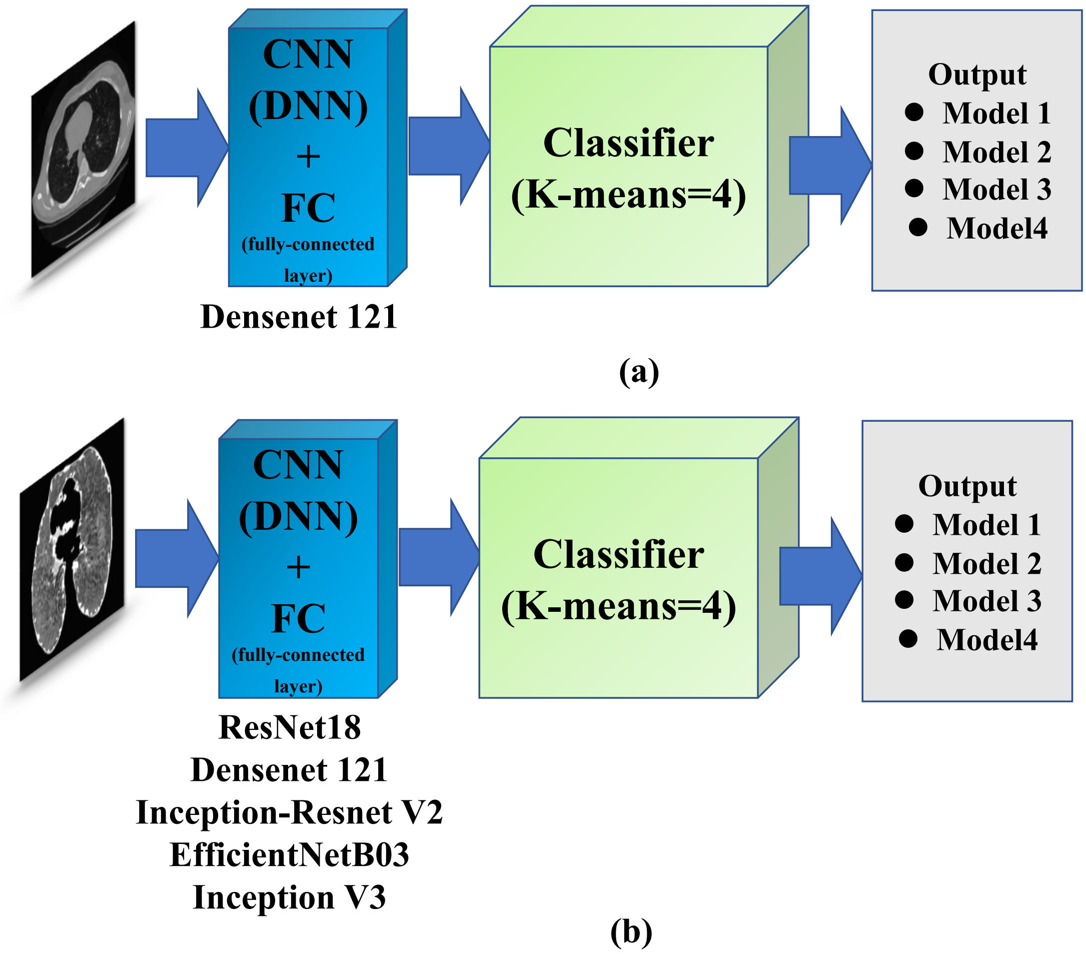

Figure 5: Overview of Baseline and Proposed Method the second subtask.

model. We noted that use Lungs image applying mask using U-net at Fig 5 (b) as Proposed model . We used to five deep neural network models: Swin Transformer, Densenet 121, Inception-Resnet V2, EfficientNetB03, and Inception V3, and used to only Densenet 121 as Baseline. In addition, CNN feature is passed out on ”K-means clustering” processing to predict the four models.

4. Experiments

4.1. Experimental parameters

Here, we have 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 50 using early-stopping. For the implementation, we employed Tensorflow[13] as our deep learning framework. For the implementation, we employed Tensorflow as our deep learning framework. These experiments were performed using PyTorch on Ubuntu 20.04. The workstation has an Intel Xeon 6242RXeon(20core/3.10GHz/TDP:205W) CPU with 16GB of 6 RAM and an NVIDIA RTX A6000 GPU.

4.2. Submission in the first subtask

Here, Table 1 shows the results of submission. Finally, we employed Densenet 121 as Baseline. And we employed Swin Transformer, Densenet 121, Inception-Resnet V2, EfficientNetB03, and Inception V3

Table 1
Submission results for the first subtask-Identify training data "fingerprints"

Model name	F1	Test Dataset 1 Accuracy	Precision	Recall	F1	Test Dataset 2 Accuracy	Precision	Recall	F1
Densenet 121 (Baseline)	0.455	**0.506**	**0.509**	0.331	0.401	0.497	0.497	**0.521**	0.509
Swin Transformer+U-net	**0.51**	0.493	0.495	**0.652**	**0.562**	0.502	0.502	0.42	0.457
Inception-Resnet V2+U-net	0.488	0.504	0.504	0.422	0.459	**0.527**	**0.528**	0.505	**0.516**
Inception V3+U-net	0.46	0.468	0.47	0.505	0.487	0.479	0.475	0.398	0.433
Densenet 121+U-net	0.443	0.502	0.503	0.314	0.386	0.49	0.49	0.511	0.5
EfficientNetB03+U-net	0.019	0.5	0.545	0.003	0.005	0.497	0.443	0.017	0.033

Table 2
Submission results for the second subtask-Detect generative models'

Model name	Submission Score (ARI)
Densenet 121 (Baseline)	0.0382425658
ResNet18+U-net	0.0138561082
Densenet 121+U-net	**0.2263399455**
Inception-Resnet V2+U-net	0.1234268768
EfficientNetB03+U-net	0.0452833430
Inception V3+U-net	0.0918183962

applying U-net mask images.

Here, in terms of the score at the total, Swin Transformer+U-net is best score.

We explain the submission score in more detail. In Test Dataset 1 of Accuracy and Precision, and Test Dataset 2 of Recall, Densenet 121 (Baseline) is best score. In Test Dataset 2 of Accuracy, Precision, and Recall, Inception-Resnet V2+U-net is best score. Finally, Test Dataset 1 of Precision and Recall, Swin Transformer+U-net is best score. This is due to the effect of the shifted window-based self-attention in the Swin transformer, which calculates features of different sizes of regions, from the features of the entire image to the features of the details.

4.3. Submission in the second subtask

Here, Table 2 shows the results of submission. The evaluation is Score as ARI (Adjusted Rand Index). Standard clustering methods used Rand Index Finally, we employed Densenet 121 as Baseline. And we employed ResNet18, Densenet 121, Inception-Resnet V2, EfficientNetB03, and Inception V3 applying U-net mask images. Here, in terms of the score, Densenet 121+U-net is best score.

5. Conclusions

In this study, we proposed lung images using segmentation based on U-net by real and generated image from chest CT images.

In addition, we could perform fine-tuning deep neural network model. In addition, the first subtask is using transfer learning at ImageCLEFmed GANS the training dataset for Task 1 and Task 2.

The experimental results demonstrate that our proposed models out-perform some models in terms of the accuracy in the first subtask and ARI in the second subtask. Therefore, we believe that using U-net to pre-process an image is fefective.

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.

Acknowledgments

Part of this research was carried out with the support of the Grant for Toyohashi Heart Center Smart Hospital Joint Research Course and the Grant-in-Aid for Scientific Research (C) (issue numbers 22K12149 and 22K12040).

References

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

[2] 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.

[3] 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.

[4] 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.

[5] S. Kozlovski, V. Liauchuk, Y. Dicente Cid, V. Kovalev, H. Müller, Overview of imagecleftuberculosis 2021 - ct-based tuberculosis type classification, in: CLEF2021 Working Notes, CEUR Workshop Proceedings, CEUR-WS.org <http://ceur-ws.org>, Bucharest, Romania, 2021.

[6] B. Ionescu, H. Müller, R. Péteri, J. Rückert, A. Ben Abacha, A. G. S. de Herrera, C. M. Friedrich, L. Bloch, R. Brüngel, A. Idrissi-Yaghir, H. Schäfer, S. Kozlovski, Y. D. Cid, V. Kovalev, L.-D. Ştefan, M. G. Constantin, M. Dogariu, A. Popescu, J. Deshayes-Chossart, H. Schindler, J. Chamberlain, A. Campello, A. Clark, Overview of the ImageCLEF 2022: Multimedia Retrieval in Medical, Social Media and Nature Applications, in: Experimental IR Meets Multilinguality, Multimodality, and Interaction, Proceedings of the 13th International Conference of the CLEF Association (CLEF 2022), LNCS Lecture Notes in Computer Science, Springer, Bologna, Italy, 2022.

[7] B. Ionescu, H. Müller, A. Drăgulinescu, W. Yim, A. Ben Abacha, N. Snider, G. Adams, M. Yetisgen, J. Rückert, A. García Seco de Herrera, C. M. Friedrich, L. Bloch, R. Brüngel, A. Idrissi-Yaghir, H. Schäfer, S. A. Hicks, M. A. Riegler, V. Thambawita, A. Storås, P. Halvorsen, D. J. A. A. A. R. I. C. V. K. A. S. G. I. Nikolaos Papachrysos, Johanna Schöler, H. Manguinhas, L. Ştefan, M. G. Constantin, M. Dogariu, J. Deshayes, A. Popescu, Overview of ImageCLEF 2023: Multimedia retrieval in medical, socialmedia and recommender systems applications, in: Experimental IR Meets Multilinguality, Multimodality, and Interaction, Proceedings of the 14th International Conference of the CLEF Association (CLEF 2023), Springer Lecture Notes in Computer Science LNCS, Thessaloniki, Greece, 2023.

[8] B. Ionescu, H. Müller, A. Drăgulinescu, W. Yim, A. Ben Abacha, N. Snider, G. Adams, M. Yetisgen, J. Rückert, A. García Seco de Herrera, C. M. Friedrich, L. Bloch, R. Brüngel, A. Idrissi-Yaghir, H. Schäfer, S. A. Hicks, M. A. Riegler, V. Thambawita, A. Storås, P. Halvorsen, D. J. A. A. A. R. I. C. V. K. A. S. G. I. Nikolaos Papachrysos, Johanna Schöler, H. Manguinhas, L. Ştefan, M. G. Constantin, M. Dogariu, J. Deshayes, A. Popescu, Overview of ImageCLEF 2024: Multimedia retrieval in medical applications, in: Experimental IR Meets Multilinguality, Multimodality, and Interaction,

Proceedings of the 15th International Conference of the CLEF Association (CLEF 2024), Springer Lecture Notes in Computer Science LNCS, Grenoble, France, 2024.

[9] A. Andrei, A. Radzhabov, D. Karpenka, Y. Prokopchuk, V. Kovalev, B. Ionescu, H. Müller, Overview of 2024 ImageCLEFmedical GANs Task – Investigating Generative Models' Impact on Biomedical Synthetic Images, in: CLEF2024 Working Notes, CEUR Workshop Proceedings, CEUR-WS.org, Grenoble, France, 2024.

[10] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation (2015). URL: https://arxiv.org/pdf/1505.04597.pdf. arXiv:1505.04597.

[11] C. Liu, M. Pang, Lung ct image segmentation via dilated u-net model and multi-scale gray correlation-based approach, Circuits, Systems, and Signal Processing 43 (2024) 1697–1714. URL: https://link.springer.com/content/pdf/10.1007/s00034-023-02532-x.pdf. doi:10.1007/s00034-023-02532-x.

[12] E. Ahn, A. Kumar, D. Feng, M. Fulham, J. Kim, Unsupervised feature learning with k-means and an ensemble of deep convolutional neural networks for medical image classification (2019). URL: https://arxiv.org/pdf/1906.03359.pdf. arXiv:1906.03359.

[13] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, R. Jozefowicz, Y. Jia, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, M. Schuster, R. Monga, S. Moore, D. Murray, C. Olah, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, X. Zheng, 2015, TensorFlow, Large-scale machine learning on heterogeneous systems, URL: https://github.com/tensorflow.