Real and Generated Image Classification using Multi-stage Transfer Learning

Notebook for the ImageCLEFmedical GANs 2023 Lab at CLEF 2023

Tetsuya Asakawa[1]*, Hiroki Shinoda[2]†, Takuya Togawa[3]†, Kazuki Shimizu[3]† and Masaki Aono[4] †

[1] *Toyohashi University of Technology, 1-1 Hibarigaoka Tenpaku, Toyohashi, Aichi, Japan*
[2] *Toyohashi University of Technology, 1-1 C3-504 Hibarigaoka Tenpaku, Toyohashi, Aichi, Japan*
[3] *Toyohashi Heart center, 21-1 Gobutori Tenpaku, Oyama, Toyohashi, Aichi, Japan*
[4] *Toyohashi University of Technology, 1-1 C-510 Hibarigaoka Tenpaku, Toyohashi, Aichi, Japan*

Abstract
CLEF 2023 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. We propose fine-tuning deep neural network model that uses multi-stage transfer learning. The first stage transfer learning uses caisia dataset. The second stage transfer learning uses COVID-19 dataset. The third stage transfer learning uses ImageCLEF Gans training dataset. The fourth stage transfer learning uses ImageCLEF Gans test dataset. Our submissions (KDE-lab team) on the task test dataset reached accuracy value of about 52% and reached a precision value of about 51.8%.

Keywords
Tuberculosis, Deep Learning, Multi-stage transfer learning, Classification.

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]. In ImageCLEF 2023 the main task [7], "ImageCLEFmed Tuberculosis," is treated as a computed tomography (CT) report. The goal of the task is to determine which real images were used in training to produce the generated images.

In this paper, we employ a new fine-tuning neural network model that uses features extracted by pre-trained convolutional neural network (CNN) models or Vision Transformer (ViT). In addition, the several datasets are used to transfer learning.

The new contributions of this paper are the proposition of novel feature building techniques, the multi-stage transfer learning from the CNN models and ViT. In Section 2, we describe the conducted

CLEF 2023: Conference and Labs of the Evaluation Forum, September 18–21, 2023, Thessaloniki, Greece
*Corresponding author.
†These authors contributed equally.

asakawa.tetsuya.um@tut.jp (T. Asakawa); shinoda.hiroki@tut.jp (H. Shinoda); t.togawa0316@gmail.com (T. Togawa) shimizu@heart-center.or.jp (K. Shimizu) masaki.aono.ss@tut.jp (M. Aono)
https://kde.cs.tut.ac.jp (M. Aono)
0000-0002-8345-7094 (T. Asakawa); 0009-0008-2850-5015 (H. Shinoda); 0009-0006-6822-5427 (T. Togawa); 0009-0000-3448-7986 (K. Shimizu); 0000-0003-1383-1076 (M. Aono)

task and the ImageCLEF2023 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.

2. ImageCLEF 2023 Dataset

The Gans task of the ImageCLEF 2023 Challenge included partial 3D patient chest CT images [8]. The training dataset for task includes 500 artificial images, 80 real images which were not used for training generative neural networks as well as 80 real images taken from the image set which has been used for training corresponding generative model. The test dataset was created in similar way. The only difference is no proportion of non-used and used ones has been disclosed. Thus, a total of 10,000 generated and 200 real images are provided. Thus, each image among the 200 images must be classified as used or unused to generate the 10,000 artificial images.

3. Proposed Method

We propose to predict real or generated Tuberculosis from CT scan images. This paper proposes fine-tuning deep neural network model that uses multi-stage transfer learning. In addition, we use a CT scan movie not CT scan images. We will detail our proposed system in the following section.

3.1. Dataset Using Multi-stage transfer learning

We used caisia dataset [9]. caisia dataset is natural images and generated images using GAN based on natural images. There is nothing to medical images. We used COVID-19 dataset at next steps [10]. COVID-19 dataset is X-ray lung images of COVID-19 and generated images using GAN based on that. The next is using at ImageCLEF Gans training dataset [11] and uses ImageCLEF Gans test dataset.

3.2. Multi-stage transfer learning

We propose fine-tuning deep neural network model that uses multi-stage transfer learning. The first stage transfer learning uses caisia dataset. The second stage transfer learning uses COVID-19 dataset. The third stage transfer learning uses ImageCLEF Gans training dataset. The fourth stage transfer learning uses ImageCLEF Gans test dataset. We used to four deep neural network models: Conv model, ResNet18, VGG11, ViT B/32.

3.2.1. Conv model

We employ from Conv model using multi-stage transfer learning. As illustrated in Fig. 1. Conv Model is 5 block convolution neural network. In addition, we extract a 512-dimensional vector from fc layer. This convolution neural network repeated 3 times.

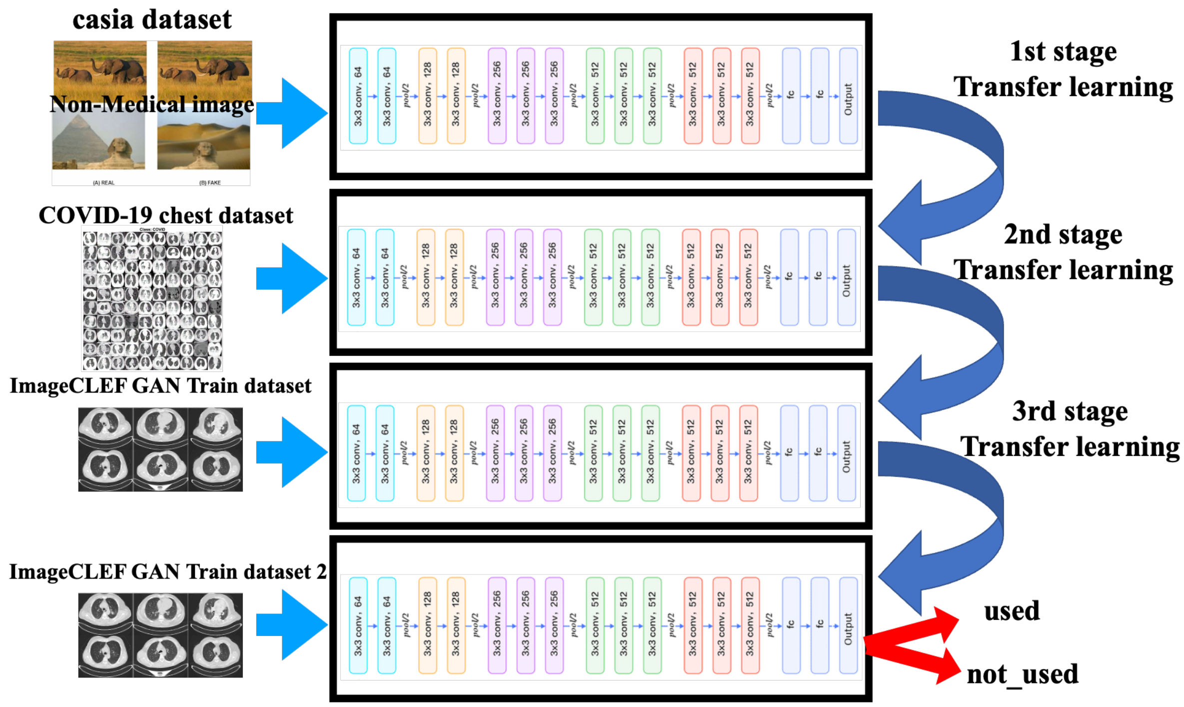

Figure 1: Overview of multi-stage transfer learning using Conv model

3.2.2. ResNet18

We employ from ResNet18 using multi-stage transfer learning. As illustrated in Fig. 2. We used ResNet18 to apply small dataset.

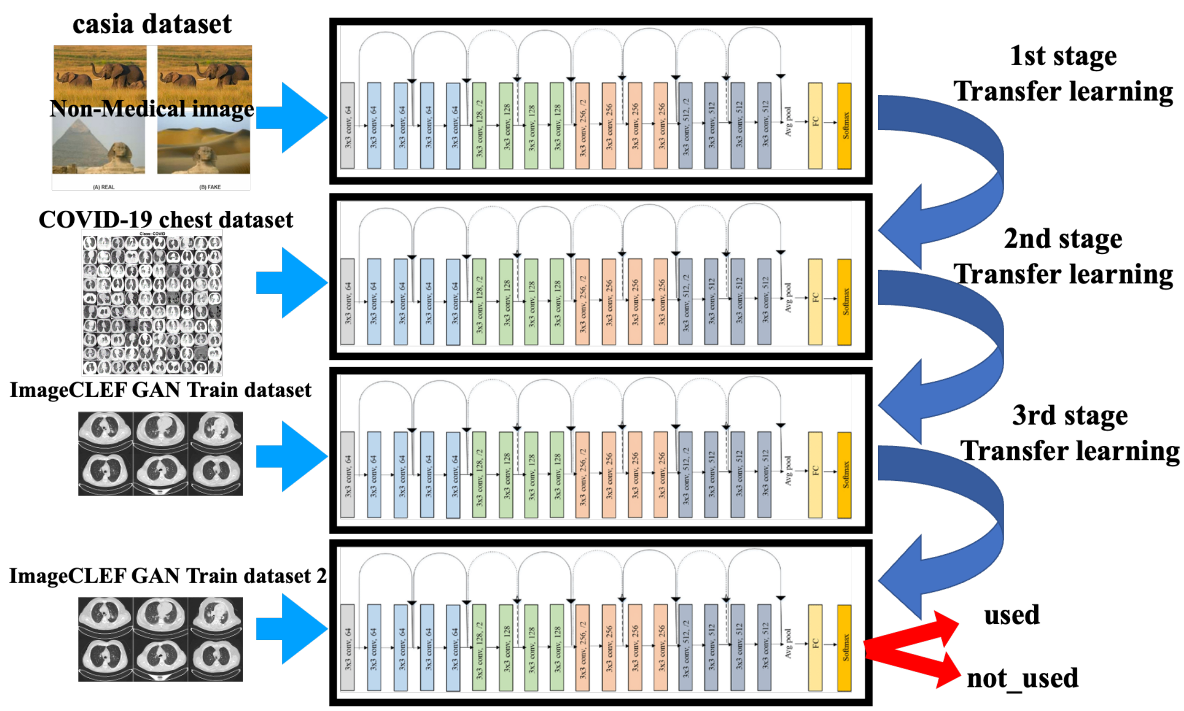

Figure 2: Overview of multi-stage transfer learning using ResNet18

3.2.3. VGG11

We employ from VGG11 using multi-stage transfer learning. As illustrated in Fig. 3. We extract a 1000-dimensional vector from fc layer.

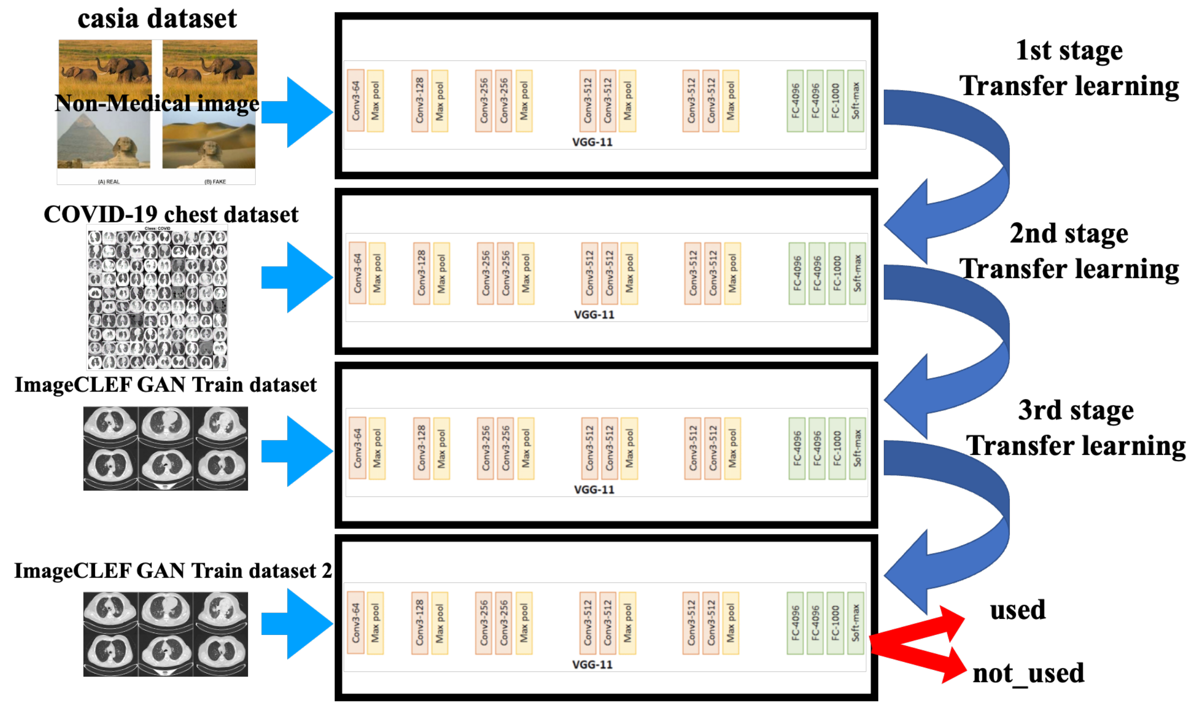

Figure 3: Overview of multi-stage transfer learning using VGG11

3.2.4. ViT B/32

We employ from ViT B/32 using multi-stage transfer learning. As illustrated in Fig. 4.

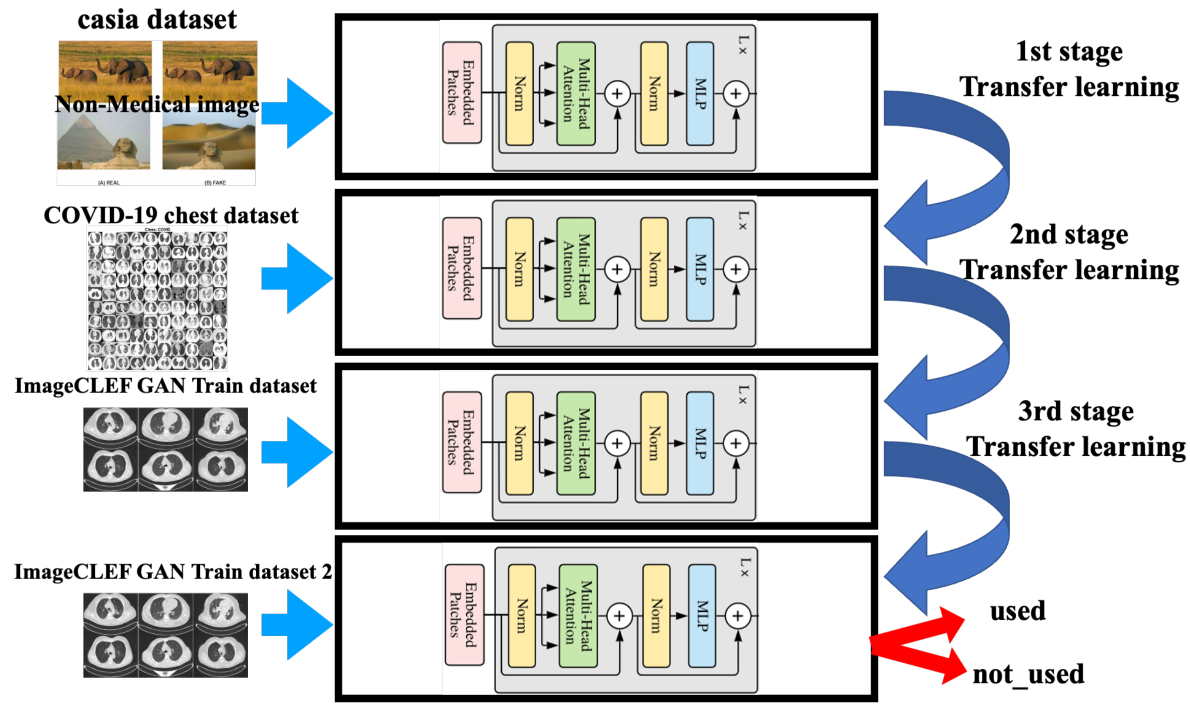

Figure 4: Overview of multi-stage transfer learning using ViT B/32

4. Experiments

4.1. Experimental parameters

Here, we have divided the filtering data into training and validation datasets with a ratio of 8:2 in multi-transfer learning. 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 using early-stopping. For the implementation, we employed Tensorflow[12] as our deep learning framework.

4.2. Submission at several models

For the evaluation of the multi-label and multi-class classification, we employed the accuracy, precision, specificity, recall, and f1_score. Table 1 shows the results. finally, we employed Conv model, ResNet18, VGG11, ViT B/32 for the training and validation datasets and the test data. Here, in terms of the accuracy (Table1), precision(Table1), specificity(Table2), recall(Table2), and f1_score(Table2), ViT B/32 is best score.

Table 1

Our submitted of accuracy and precision for the proposed model.

	Model	accuracy	precision
Submission1	Conv model	0.495	0.494
Submission3	ResNet18	0.505	0.507
Submission4	VGG11	0.49	0.488
Submission5	ViT B/32	0.52	0.518

Table 2

Our submitted of specificity, recall and f1_score for the proposed model.

	Model	specificity	recall	f1_score
Submission1	Conv model	0.55	0.44	0.465
Submission3	ResNet18	0.67	0.34	0.505
Submission4	VGG11	0.55	0.43	0.49
Submission5	ViT B/32	0.47	0.57	0.52

5. Conclusions

In this study, we proposed several CNN models or ViT for predicting real and generated image from chest CT images. In addition, we could perform multi-stage transfer learning. Specifically, The first stage transfer learning uses caisia dataset. The second stage transfer learning uses COVID-19 dataset. The third stage transfer learning uses ImageCLEF Gans training dataset. The fourth stage transfer learning uses ImageCLEF Gans test dataset.

The experimental results demonstrate that our proposed models out-perform some models in terms of the accuracy and precision. Therefore, we believe that using multi-stage transfer learnig 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.

Acknowledgements

A 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 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.

[2] Y. Dicente Cid, V. Liauchuk, V. Kovalev, , H. Müller, Overview of ImageCLEFtuberculosis 2018 - detecting multi-drug resistance, classifying tuberculosis type, and assessing sever-ity 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 imageclef tuberculosis 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. Ko-valev, 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 Interna-tional 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 sys-tems applications, in: Experimental IR Meets Multilinguality, Multimodality, and Inter-action, Proceedings of the 14th International Conference of the CLEF Association (CLEF 2023), Springer Lecture Notes in Computer Science LNCS, Thessaloniki, Greece, 2023.

[8] A. Andrei, A. Radzhabov, I. Coman, V. Kovalev, B. Ionescu, H. Müller, Overview of ImageCLEFmedical GANs 2023 task – Identifying Training Data "Fingerprints" in Synthetic Biomedical Images Generated by GANs for Medical Image Security, in: CLEF2023 Working Notes, CEUR Workshop Proceedings, CEUR-WS.org, Thessaloniki, Greece, 2023.
[9] S. Vathana, casia dataset, URL: https://www.kaggle.com/datasets/sophatvathana/casia-dataset.
[10] M. Rahimzadeh, A. Attar, S. M. Sakhaei, A fully automated deep learning-based network for detecting covid-19 from a new and large lung ct scan dataset, Biomedical Signal Processing and Control 68 (2021) 102588. URL: https://www.sciencedirect.com/science/article/pii/S1746809421001853. doi:https://doi.org/10.1016/j.bspc.2021.102588.
[11] B. I. Alexandra Andrei, H. Müller, Imageclef2023 gans dataset, ???? URL: https://github.com/ImageCLEF/2023_ImageCLEFmed_GANs.
[12] 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.