RETHINKING COMPUTER-AIDED PELVIS SEGMENTATION

Siming Yuan[1], Qing Liu[1], Shenghui Liao[1, †], Fuchang Han[1], Haitao Wei[2], Yingqi Zhang[3]

[1]School of Computer Science and Engineering, Central South University, Changsha 410083, China
[2] Central South E3D Digital Medical and Virtual Reality Research Center, Changsha 410083, China
[3] Tongji Hospital, School of Medicine, Tongji University, Shanghai 200000, China

ABSTRACT

As an important structure connecting the spine and lower limbs, the abnormal pelvis is one of the threats to human health worldwide, leading to millions of deaths every year. Although early diagnosis and treatment can greatly improve the chances of survival, it remains a major challenge, especially for pelvis fractures. Computer-aided pelvis segmentation (CPS) is a promising choice for the abnormal pelvis due to the great success of deep learning. However, when it comes to the abnormal pelvis diagnosis, the lack of training data has hampered the progress of CPS. To solve this problem, we have published a large-scale pelvis dataset, namely the Pelvis Computed tomography (CT) image (PCT14K) dataset. This dataset contains 14487 CT slices with the corresponding label for pelvis areas, while the existing largest public pelvis dataset part comes from existing data sets related to other organs with a lot of redundant information, and another part comes from the orthopedic hospital without a corresponding label. The proposed dataset enables the training of sophisticated segmentation networks for high-quality CPS. Some mainstream segmentation algorithms are trained and evaluated on the proposed PCT14K dataset and served as the baselines for future research. The published dataset will be available at https://github.com/YUAN-SIMING/PCT14K.

Index Terms— Pelvis, deep learning, CT images, segmentation algorithms

1. INTRODUCTION

The pelvis is an important structure connecting the spine and lower limbs. It not only plays a crucial role in maintaining the stability of the body and protecting the internal organs of the abdomen, but it is also closely related to female childbirth. Pelvis abnormalities, such as hip dysplasia [1] and pelvis fractures, can seriously affect human health. Accurate pelvis bone segmentation is essential for assessing the severity of the pelvic injury, helping surgeons to make correct judgments, and choosing the appropriate surgical method.

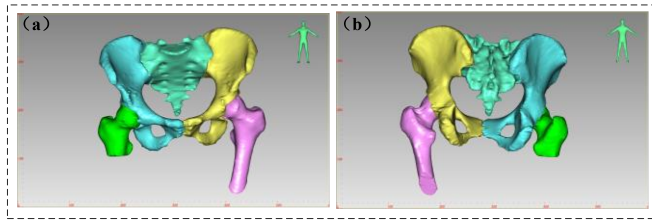

Fig. 1. CT image of the pelvis.

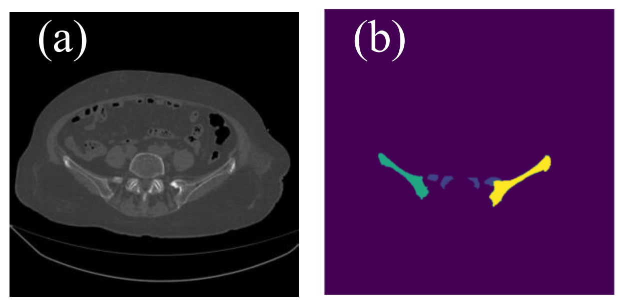

Fig. 2. Two-dimensional slices. (a) Original image. (b) Ground truth.

Computed tomography (CT) has high resolution and can clearly show the characteristics of human anatomy and diseased tissue regions. Therefore, CT technology has also become a major issue when dealing with related diseases in the pelvis region in clinical practice auxiliary means. However, the current traditional segmentation method relies on the doctor's clinical experience to manually delineate and segment layer-by-layer CT images. This method takes a long time, not only greatly increases the work intensity of the doctor, but is also prone to errors, and the accuracy can not be guaranteed. In addition, the segmentation results are also subjective, and different doctors will produce different segmentation results [2]. Thanks to the powerful representation ability of deep learning, especially convolutional neural networks (CNNs) [3, 4], machines have outperformed human in many fields, such as face recognition, image classification [5], object detection [6], and edge detection [6]. Deep learning can capture detailed information [6] and never feel

† Corresponding author: lsh@csu.edu.cn

This work was supported by the National Natural Science Foundation of China (No.61772556), National Key R&D Program of China (No.2018 YFB1107100, No.2016 YFC1100600).

tired like humans. A natural idea is to use deep learning for computer-aided pelvic bone segmentation (CPS) of CT images. However, deep learning is always data-hungry, and it is difficult to collect large-scale pelvis data because much data related to patient privacy cannot be made public, and medical image annotation is also very complicated. The lack of publicly available CT images has prevented computer-aided pelvis bone segmentation (CPS) from successfully applying deep learning for improving performance. Although there are some data sets related to pelvis bone [7, 8, 9], only a few of them are open source, and the number is small (less than 5 people or 200 slices), which is much lower than that of other organs quantity. For example, the existing largest public CT image dataset for pelvis bone segmentation is the CTPelvic1K dataset proposed. The CTPelvic1K dataset has seven sources, five of which are from existing data sets [10], these are data sets related to other organs, so some redundant processing is required to get the required pelvic data. This is not only time-consuming and laborious but also may introduce some low-quality data or do not contain redundant information in the pelvis area. The remaining two are from orthopedic hospitals, one of which is postoperative images with metal artifacts, so there is no label and can only be used for unsupervised learning.

In order to actually deploy the CPS system to help patients with pelvis abnormalities around the world, we must first solve the problem of insufficient data. In this paper, we contribute to the community with a large-scale pelvis CT image (PCT14K) dataset, through long-term cooperation with Shanghai Tongji Hospital.

Moreover, by comparing many different classic image segmentation algorithms, we found that BiseNet [12] and LedNet [11] can be used as valuable segmentation tools for computer-assisted pelvic segmentation. These methods can be used as benchmarks for future CPS research. Compared with the previous CPS dataset, the main advantages of this new data set are the following: (1) Unlike the previous dataset [17,18,19,20] only contains less than 5 people or 200 slices, PCT14K has 14487 CT images, so that PCT14K makes it possible to train very deep CNNs. (2) PCT14K includes five categories: sacrum, left ilium, right ilium, left femur, and right femur. Their decision is based on the consensus reached by experts in many orthopedic fields, unlike the previous data set that simply divides the entire pelvis. This enables future CPS systems to adapt to more complex real worlds and provide people with more detailed disease analyses. The label of PCT14K is jointly marked by an associate professor with 5 years of experience and an associate professor with 23 years of experience. The PCT14K data set has been desensitized by the data provider and approved by the ethics committee of Shanghai Tongji Hospital. It complies with the principles of the Helsinki Clinical Research Declaration. In addition, all participants in this study signed a written consent form before undergoing the examination.

By comparing many different classic image segmentation algorithms, we found that BiseNet [12] and LedNet [11] can be used as valuable segmentation tools for computer-assisted pelvic segmentation. These methods can be used as benchmarks for future CPS research.

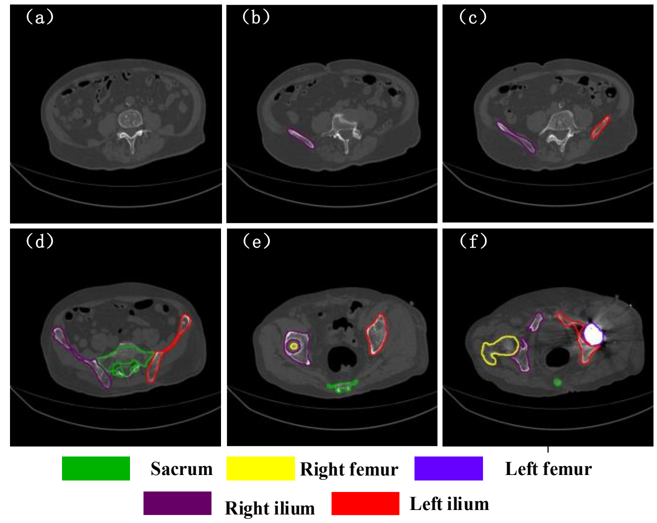

Fig. 3. Different slice positions of the same patient with different numbers of ROIs (the above picture is patient 1). (a) Slice position 17 does not include the pelvic area. (b) Slice position 24, which includes one type in the pelvic area. (c) Slice position 34, which includes two types in the pelvic area. (d) No. 76 slice position, which includes three types in the pelvic area. (e) No. 166 slice position, which includes four types in the pelvic area. (f) No. 213 Slice location, including five categories in the pelvic area.

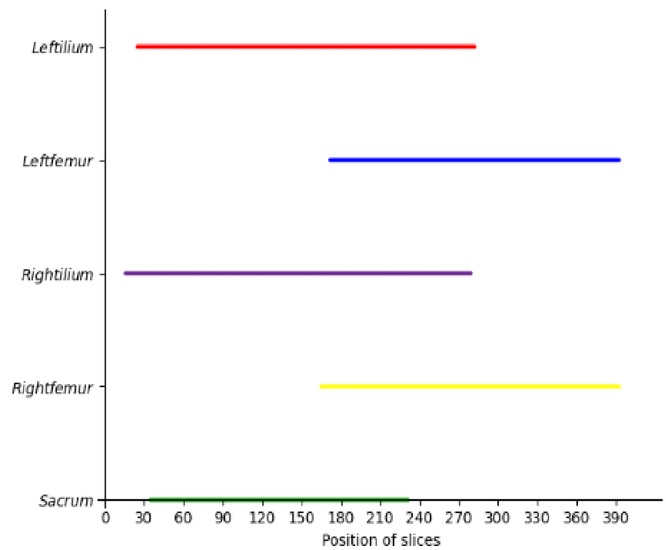

Fig. 4. ROI statistics of the same patient at different slice positions (the above picture is patient No. 1). Among them, the ordinate is the five class labels, and the abscissa is the location of the slice. If the category i exist at position x, then draw a point at position (x, i). Different types are represented by different colors.

In summary, our contributions are twofold:

- We build a large-scale CPS dataset that is much larger, better annotated, and more realistic than the existing Pelvic dataset, enabling the training of deep CNNs.

- We establish the CPS benchmark by using the existing image segmentation network for pelvic segmentation, which is expected to set a good start for future CPS research.

2. THE PELVIC CT (PCT14K) DATASET

In this section, we will introduce our dataset in detail from four aspects: CT image collection, mask annotation, data description, and data statistics.

CT Image Collection. The main difficulty in collecting CT images of the pelvis is that CT images of pelvic abnormalities have a high degree of privacy because these data contain the patient's name, age, gender, reproductive organs, and other information, leaking these data will expose people to the risk of illegality, so it is almost impossible for individuals to get the original data. In order to overcome this difficulty, we cooperated with Shanghai Tongji Hospital to collect CT images of the pelvis. As shown in Figure 1, there were 42 patients who underwent hip replacement due to fractures or bone tumors. Among them, there were males and females.

Mask Annotation. The acquired pelvic data is segmented on the two-dimensional image by experienced orthopedic surgeons in top hospitals through E3D (http://www.e3d-med.com/) software, and the required areas are retained to form a two-dimensional mask. The software automatically converts a two-dimensional mask into a three-dimensional mask, performs a secondary segmentation of each area, and separates each area of the pelvis to obtain and export the two-dimensional mask of each area. Our segmentation is performed on a two-dimensional image, as shown in Figure 2.

Data Description. Our data contains a total of 42 patients, each patient contains a different number of two-dimensional slices, a total of 14487 CT slices. The slice spacing is 1mm or 2mm, and the resolution of all slices is 512×512. Some of our pictures have inconsistent ROIs, as shown in Figure 3 and Figure 4.

Data Statistics. Finally, we analyzed the attributes of the PCT14K dataset, as shown in Figure 5. The abscissa indicates the area of category *i*/image area, and the ordinate indicates cumulative image percentage. Each point (x, y) on the curve means: in y% of the images, the area of category *i*/image area is less than or equal to x%. According to Figure 5, we analyzed the average size of the objects in the data set. Generally speaking, smaller objects are more difficult to segment. As shown in Figure 5, the average size of the sacrum is smaller.

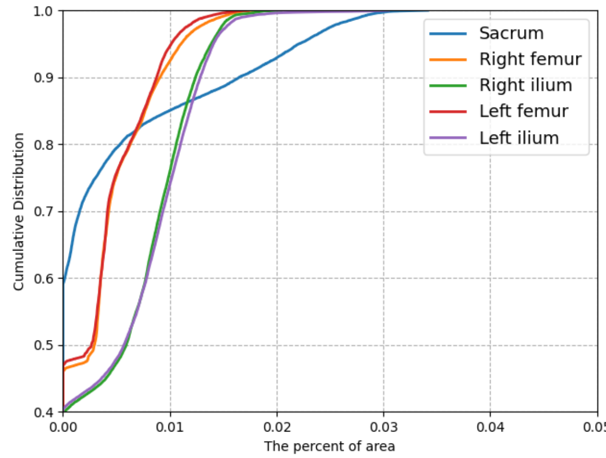

Fig.5. The distribution of instance sizes for the sacrum, right femur, right ilium, left femur and left ilium.

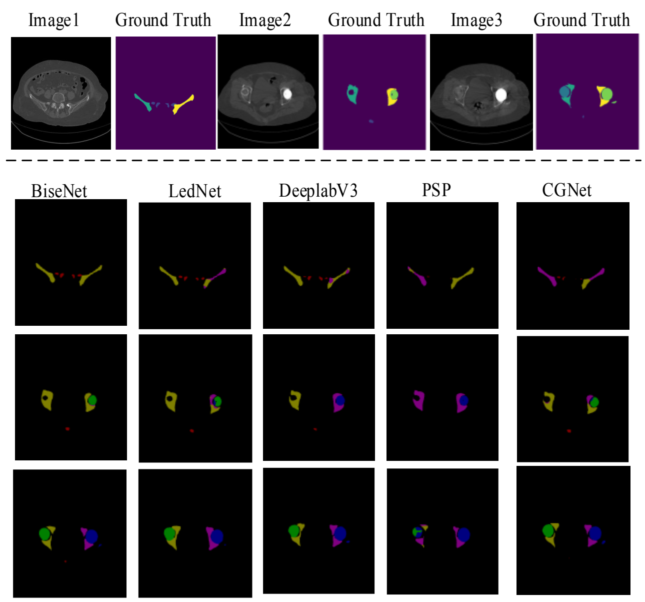

Fig.6. Visual comparison of our segmentation result on the PCT14K test set.

3. EXPERIMENTAL RESULTS

In this section, we establish some baselines for our pelvic segmentation by using some mainstream segmentation algorithms. PCT14K dataset contains 14487 images size of 512×512 and five predefined categories. They are the sacrum, left ilium, right ilium, left femur, and right femur. All categories are labelled by two experts. PCT14K dataset is divided into Two sets: 35 groups of data for train,7 groups of data for test.

Experimental Setting. Our experiment is based on the existing common image segmentation methods, as shown in Table 1. Different methods correspond to different backbone networks. The parameters in the backbone network are initi-

Table 1. Performance comparison with the state-of-the-arts on PCT14K.

Model	Resolution	Backbone	mIoU (%)	ACC (%)
LedNet [11]	512*512	Resnet50 [23]	74.258	99.410
BiseNet [12]	512*512	Resnet18 [23]	75.979	99.468
PSPNet [13]	512*512	Resnet50	66.036	99.263
DeeplabV3 [14]	512*512	Resnet50	72.954	99.394
OCNet [15]	512*512	Resnet50	65.557	99.254
CCNet [16]	512*512	Resnet50	63.436	99.200
DANet [17]	512*512	Resnet50	53.947	98.940
DUNet [18]	512*512	Resnet50	63.091	99.204
FCN [19]	512*512	VGG16 [24]	51.289	98.649
Cgnet [20]	512*512	Resnet50	67.102	99.271
ESPNet [21]	512*512	Resnet50	54.132	98.845
DenseASPP [22]	512*512	Densenet121 [25]	50.691	98.909

alized using the pre-trained model on ImageNet [26]. Parameters are optimized by SGD on one GPU. Hyper-parameters includes: base learning rate (0.001), weight decay (0.0001), momentum (0.9), batch size (2) and iteration epoch (60). Each experiment uses 2 GPUs. During the training phase, random mirrors, random scale 0.5~2, and gaussian blur are randomly used to augment the training set.

Comparison with State-of-the-arts. We compared LedNet [11], BiseNet [12], PSPNet [13], DeeplabV3 [14], OCNet [15], CCNet [16],DANet [17], DUNet [18], FCN[19], Cgnet [20], ESPNet [21], DenseASPP [22]. All of the results are fine-tuned on PCT14K. We repeat each experiment three times and take the best result of the three times. Table 1 shows the results. Obviously, BiseNet [12] has achieved 75.979% mean intersection over union (mIoU) and 99.468% accuracy (ACC), and LedNet [11] has achieved 74.258% mIoU and 99.41% ACC, which shows that BiseNet [12] and LedNet [11] can be used as valuable pelvic segmentation tools.

Visualization. We visualized the segmentation result map on the PCT14K test set. Results are reported in Figure 6. It can be seen from the segmentation result graph that PSPNet [13] cannot accurately predict smaller areas, such as the small red dots in the first and second rows of the figure. In contrast, BiseNet [12] and LedNet [11] have better prediction results.

4. CONCLUSION

Early diagnosis and treatment of pelvis abnormalities can greatly improve the survival rate. Unfortunately, the diagnosis and treatment of pelvis abnormalities remain a major challenge. Inspired by the success of deep learning, deep learning-based CPS is a promising research direction. However, the lack of data has prevented deep learning from bringing progress for CPS. In this paper, we build a large-scale pelvis dataset PCT14K with Mask annotations, enabling the training of deep CNNs for pelvis abnormalities diagnosis and treatment. We build an initial benchmark for CPS by further proposing some baselines. This new PCT14K dataset and benchmark are expected to promote the research in CPS, and better CPS systems are expected to be designed with new powerful deep networks.

5. REFERENCES

[1] Pavel Kotlarsky, Reuben Haber, Victor Bialik, and Mark Eidelman, “Developmental dysplasia of the hip: What has changed in the last 20 years?,” *World Journal of Orthopedics*, vol. 6, no. 11, pp. 886-901, 2015.

[2] Andreas Schicho, Stefan A. Schmidt, Kevin Seeber, Alain Olivier, Peter H. Richter, and Florian Gebhard, “Pelvic X-ray misses out on detecting sacral fractures in the elderly–Importance of CT imaging in blunt pelvic trauma,” *Injury international Journal of the Care of the Injured*, vol. 47, pp. 707-710, 2016.

[3] Shang-Hua Gao, Ming-Ming Cheng, Kai Zhao, Xin-Yu Zhang, Ming-Hsuan Yang, and Philip Torr, “Res2Net: A new multi-scale backbone architecture,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, pp. 652-662, 2020.

[4] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770-778.

[5] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, “Delving deep into rectifiers: Surpassing human-level performance on imagenet classification,” in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 1026-1034.

[6] Kaiming He, Georgia Gkioxari, Piotr Dollar, and Ross Girshick, “Mask R-CNN,” in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 2961-2969.

[7] K. Punnam Chandar, and T. Satyasavithri, "Segmentation and 3d visualization of pelvic bone from CT scan images," In *IACC*. IEEE, Feb 2016, pp. pp. 430–433.

[8] Robert Hemke, Colleen G. Buckless, Andrew Tsao, Benjamin Wang, and Martin Torriani, "Deep learning for automated segmentation of pelvic muscles, fat, and bone from CT studies for body composition assessment," *Skeletal Radiology*, vol. 49, no. 3, pp. 387-395, 2020.

[9] Pei-Yuan Lee, Jiing-Yih Lai, Yu-Sheng Hu, Chung-Yi Huang, Yao-Chen Tsai, and Wen-Der Ueng, "Virtual 3D planning of pelvic fracture reduction and implant placement," *Biomedical Engineering: Applications, Basis and Communications*, vol. 24, no. 3, pp. 245-262, 2012.

[10] C. Daniel Johnson, M.D, M.M.M, Mei-Hsiu Chen, Ph.D, Alicia Y. Toledano, and et al, "Accuracy of CT colonography for detection of large adenomas and cancers," *New England Journal of Medicine*, vol. 359, no. 12, pp. 1207–1217.

[11] Yu Wang, Quan Zhou, Jia Liu, Jian Xiong, Guangwei Gao, Xiaofu Wu, and Longin Jan Latecki, "Lednet: A Lightweight Encoder-Decoder Network for Real-Time Semantic Segmentation," in *2019 IEEE International Conference on Image Processing (ICIP)*, 2019, pp. 1860-1864.

[12] Changqian Yu, Jingbo Wang, Chao Peng, Changxin Gao, Gang Yu, and Nong Sang, "BiSeNet: Bilateral Segmentation Network for Real-time Semantic Segmentation," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 325-341.

[13] Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, and Jiaya Jia, "Pyramid Scene Parsing Network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2881-2890.

[14] Liang-Chieh Chen, George Papandreou, Florian Schroff, and Hartwig Adam, "Rethinking atrous convolution for semantic image segmentation," *arXiv preprint*, 2017.

[15] Yuhui Yuan, Lang Huang, Jianyuan Guo, Chao Zhang, Xilin Chen, and Jingdong Wang, "Ocnet: Object context network for scene parsing," *arXiv preprint*, 2018.

[16] Zilong Huang, Xinggang Wang, Lichao Huang, Chang Huang, Yunchao Wei, and Wenyu Liu, " CCNet: Criss-Cross Attention for Semantic Segmentation," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019, pp. 603-612.

[17] Jun Fu, Jing Liu, Haijie Tian, Yong Li, Yongjun Bao, Zhiwei Fang, and Hanqing Lu, "Dual Attention Network for Scene Segmentation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019*, pp. 3146-3154.

[18] Qiangguo jin, Zhaopeng Meng, Tuan D.Pham, Qi Chen, Leyi Wei, and Ran Su, "DUNet: A deformable network for retinal vessel segmentation," *Knowledge-Based Systems*, vol. 178, pp. 149-162, 2019.

[19] Jonathan Long, Evan Shelhamer, and Trevor Darrell, "Fully Convolutional Networks for Semantic Segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 3431-3440.

[20] Tianyi Wu, Sheng Tang, Rui Zhang, Juan Cao, and Yongdong Zhang, "Cgnet: A light-weight context guided network for semantic segmentation," *IEEE Transactions on Image Processing*, vol. 30, pp. 1169-1179, 2020.

[21] Sachin Mehta, Mohammad Rastegari, Anat Caspi, Linda Shapiro, and Hannaneh Hajishirzi, "ESPNet: Efficient Spatial Pyramid of Dilated Convolutions for Semantic Segmentation," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 552-568.

[22] Maoke Yang, Kun Yu, Chi Zhang, Zhiwei Li, and Kuiyuan Yang, "DenseASPP for Semantic Segmentation in Street Scenes," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 3684-3692.

[23]Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep Residual Learning for Image Recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770-778.

[24] Simonyan, Karen, and Andrew Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint*, 2014.

[25] Gao Huang, Zhuang Liu, and Laurens Van Der Maaten, "Densely connected convolutional networks," *IEEE Computer Society*, pp. 1063-6919, 2017.

[26] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, and et al., "Imagenet large scale visual recognition challenge," *IJCV*, vol. 115, no. 3, pp. 211–252, 2015.