

Medical Dataset Collection for Artificial Intelligence-based Medical Image Analysis

– License: For Non-Commercial Use Only –

Yang WEN
School of Computer Science, UESTC
young.wen@foxmail.com

1. Introduction

Since finding proper dataset with sufficient well-annotated samples is vital for current AI-based research in the field of medical image analysis, I offer here a dataset collection to the community to help beginning researchers obtain the ideal dataset in a simple and straightforward way.

In total 32 public datasets was collected, of which 28 for medical imaging and 4 for natural image ones. The images of these datasets are captured by different cameras, thus vary from each other in modality, frame size and capacity. Detail information is demonstrated in Table. 1. For data accessibility, we also provide the websites of most datasets and hope this will help the readers reach the datasets.

For more information, please kindly visit my personal homepage: <https://wenyanger.github.io/>

1.1. Natural Images

ImageNet [21] Large Scale Visual Recognition Challenge 2012 (ILSVRC2012) is aiming to estimate the content of photographs for the purpose of retrieval and automatic annotation using a subset of the large hand-labeled ImageNet dataset (10,000,000 labeled images depicting 10,000+ object categories) as training. The general goal is to identify the main objects present in images. The training data, the subset of ImageNet containing the 1000 categories and 1.2 million images. A random subset of 50,000 of the images with labels will be released as validation data included in the development kit along with a list of the 1000 categories. The dataset is available at <http://image-net.org/challenges/LSVRC/2012/>

CIFAR100 [20] The CIFAR-100 is a dataset consists of 60000 32×32 colour images in 100 classes containing 600 images each. There are 500 training images and 100 testing images per class, which are randomly-selected images from each class. The 100 classes in the CIFAR-100 are grouped into 20 superclasses. In our experiments, we use the "fine" label, *i.e.*, 100 classes, for evaluation. The dataset is available at <http://www.cs.toronto.edu/~kriz/cifar.html>

ADE20K [35] We use both classification and segmentation labels. The dataset is available at <http://groups.csail.mit.edu/vision/datasets/ADE20K/>

PascalVOC [12] The dataset is available at <http://host.robots.ox.ac.uk/pascal/VOC/>

1.2. X-Ray Images

CheXpert [15] CheXpert is a large public dataset for chest radiograph interpretation, consisting of 224,316 chest radiographs of 65,240 patients. The data were collected from the chest radiographic examinations from Stanford Hospital, performed between October 2002 and July 2017 in both inpatient and outpatient centers, along with their associated radiology reports. The dataset is available at <https://stanfordmlgroup.github.io/competitions/chexpert/>

ChestXRay 2017 [19] The dataset was collected and labeled a total of 5,232 chest X-ray images from children, including 3,883 characterized as depicting pneumonia (2,538 bacterial and 1,345 viral) and 1,349 normal, from a total of 5,856 patients to train the AI system. The model was then tested with 234 normal images and 390 pneumonia images (242 bacterial and 148 viral) from 624 patients. The dataset is available at <https://data.mendeley.com/datasets/rscbjbr9sj/3>

LUNA [29] LUNg Nodule Analysis (LUNA) is a competition held to aid the development of the nodule detection algorithm. We use the lung segmentation of the challenge for our experiments. The dataset is available at <https://luna16.grand-challenge.org/Data/>

MURA [27] MURA (musculoskeletal radiographs) is a large dataset of bone X-rays. Algorithms are tasked with determining whether an X-ray study is normal or abnormal. The dataset is available at <https://stanfordmlgroup.github.io/competitions/mura/>

1.3. Ultrasound

BUS [34] The goal of the Breast Ultrasound Lesions Dataset (Dataset B) is to provide the images and ground

Table 1. Statistics of the datasets. Here CLS denotes classification, SEG denotes segmentation, Obj-D denotes object detection. For datasets with official division of training and testing sets, the number of images are shown in (number of training images) / (number of testing images), otherwise the total amounts of images.

Name	Modality	Target	Frame Size	# of Images	# of Classes	Task
CheXpert [15]	X-Ray	Lung	$\sim 390 \times 320$	224,316/624	2	CLS
ChestXRay 2017 [19]	X-Ray	Lung	$\sim 1000 \times 700$	5,232/624	2	CLS
LUNA [29]	CT	Lung	512×512	267	2	SEG
CHAOS-CT [18]	CT	Liver	512×512	2874	2	SEG
NIH-CT-82 [28]	CT	Pancreas	512×512	7,141	2	SEG
CHAOS-MRI [18]	MRI	Liver/Kidney/Spleen	320×320	992	5	SEG
CardiacMRI [2]	MRI	Heart	256×256	399	3	SEG
ACDC [4]	MRI	Heart	320×320	-	4	SEG
TCIA [24, 5]	MRI	Prostate	320×320	-	3	SEG
PROMISE12 [23]	MRI	Prostate	-	-	-	SEG
ISIC2019 [9]	Dermoscopy	Skin	1024×768	25,331	9	CLS
TNBC [25]	H/E Stained	Cell	512×512	50	2	SEG
GlaS [30]	H/E Stained	Cell	775×522	85/80	2	SEG
MoNu [22]	H/E Stained	Cell	1000×1000	30/14	2	SEG
BC-2015 [3]	H/E Stained	Breast cancer	2040×1536	269	2	CLS
DRISHTI-GS [31]	Fundoscopic	Optic Disc, Cup	2045×1752	50/51	2	SEG
DRIONS-DB [7]	Fundoscopic	Fundus	400×400	-	2	SEG
RIM-r3 [13]	Fundoscopic	Fundus	1072×1072	99/60	2	SEG/CLS
REFUGE [26]	Fundoscopic	Optic Disc, Cup/Glaucoma	1634×1634	400/400	2/2	SEG/CLS
HRF [6]	Fundoscopic	Vessel/Glaucoma	3504×2336	45	3	SEG/CLS
DRIVE [32]	Fundoscopic	Vessel	565×584	20/20	2	SEG
MESSIDOR [11]	Fundoscopic	Diabetic Retinopathy	2240×1488	1,200	4/4	CLS
STARE [14]	Fundoscopic	Diabetic Retinopathy/Vessel	700×650	397	15	SEG/CLS
EyePACS [16]	Fundoscopic	Diabetic Retinopathy	4000×4000	35126	5	CLS
BUS [34]	Ultrasound	Breast lesion	500×500	163	2	SEG
BUSI [1]	Ultrasound	Breast lesion	500×500	780	3	SEG/CLS
OCT2017 [19]	OCT	Fundus	$\sim 512 \times 512$	108,309/1,000	4	CLS
KIMCCS [17]	Cervoscope	Cervical	$\sim 3000 \times 3000$	1,463	3	CLS
CIFAR100 [20]	Natural Image	-	32×32	50,000/10,000	100	CLS
ImageNet [21]	Natural Image	-	$\sim 500 \times 375$	1.3M/60000	1000	CLS
ADE20K [35]	Natural Image	-	$\sim 1000 \times 1000$	20,210/2000	2603/826	SEG/CLS
PascalVOC [12]	Natural Image	-	$\sim 1000 \times 1000$	20,210/2000	2603/826	Obj-D

truth for the automatic interpretation and analysis of breast ultrasound images. UDIAT-Centre Diagnostic, Corporaci Parc Taul, Sabadell (Spain) has copyright on the data and is the principal distributor of Dataset B. University of Girona and Manchester Metropolitan University are involved in an ongoing effort to develop this dataset to aid research efforts in the general area of developing, testing and evaluating algorithms for breast ultrasound lesions analysis. The dataset is available at www2.docm.mmu.ac.uk/STAFF/M.Yap/dataset.php

BUSI [1] The data collected at baseline include breast ultrasound images among women in ages between 25 and 75 years old. This data was collected in 2018. The number of patients is 600 female patients. The dataset consists of 780 images with an average image size of 500×500 pixels. The images are in PNG format. The ground truth images are presented with original images. The images are categorized into three classes, which are normal, benign, and malignant. The dataset is available at <https://scholar.cu.edu.sg/?q=afahmy/pages/dataset>

1.4. MRI & CT Images

CHAOS [18] CHAOS challenge aims the segmentation of abdominal organs (liver, kidneys and spleen) from CT and MRI data. Segmentation of liver from computed tomography (CT) data sets, which are acquired at portal phase after contrast agent injection for pre-evaluation of living donated liver transplantation donors. Segmentation of four abdominal organs (*i.e.* liver, spleen, right and left kidneys) from magnetic resonance imaging (MRI) data sets acquired with two different sequences (T1-DUAL and T2-SPIR). The dataset is available at <https://chaos.grand-challenge.org/Data/>

NIH-CT-82 [28] The National Institutes of Health Clinical Center performed 82 abdominal contrast enhanced 3D CT scans from 53 male and 27 female subjects. Seventeen of the subjects are healthy kidney donors scanned prior to nephrectomy. The remaining 65 patients were selected by a radiologist from patients who neither had major abdominal pathologies nor pancreatic cancer lesions. A medical student manually

performed slice-by-slice segmentations of the pancreas as ground-truth and these were verified/modified by an experienced radiologist. The dataset is available at <https://wiki.cancerimagingarchive.net/display/Public/Pancreas-CT>

CardiacMRI [2] The dataset was comprised of short axis cardiac MR image sequences acquired from 33 subjects, for a total of 7,980 2D images. Each patient's image sequence consisted of exactly 20 frames and the number of slices acquired along the long axis of the subjects ranged between 8 and 15. Spacing-between slices ranged between 6 and 13 mm. Each image where both the endocardial and epicardial contours of the left ventricle were visible, was manually segmented, to provide the ground truth. The manual segmentation was performed by the first author and took approximately 3 weeks of full time work. This resulted in 5011 manually segmented MR images, with a total of 10,022 endocardial and epicardial contours.

ACDC [4] The ACDC dataset [4] was acquired from two MRI scanners (1.5T and 3.0T) and used in the STACOM 2017 ACDC challenge. It comprises 150 patients divided into the same 5 subgroups and annotated with three classes, namely, the right ventricles (RV), left ventricles (LV), and myocardium (Myo). The MRI data with associated annotations for 100 patients is publicly released, and the MRI data for the other 50 patients is released with their annotations held by the challenge organizer. We leverage the MRI data for the 100 released patients with annotations for cross-dataset validation. The dataset is available at <https://www.creatis.insa-lyon.fr/Challenge/acdc/index.html>

Prostate segmentation The Prostate dataset contains 40 patients from the PROSTATE-DIAGNOSIS collection [24] scanned with a 1.5T Philips Achieva MRI scanner. It is split into 30 patients for training, 5 for testing, and 5 for the competition (not used in our experiments). The labels are provided by Cancer Imaging Archive (TCIA) site [8]. The image size is 400×400 or 320×320 . This dataset has two labeled categories, peripheral zone (PZ) and central gland (CG). The dataset is available at <https://wiki.cancerimagingarchive.net/display/DOI/NCI-ISBI+2013+Challenge%3A+Automated+Segmentation+of+Prostate+Structures>

PROMISE12 Challenge [23] The MICCAI Prostate MR Image Segmentation-challenge 2012, to segment the prostate in transversal T2-weighted MR images. The data includes both patients with benign disease (e.g. benign prostatic hyperplasia) and prostate cancer. Additionally, to test the robustness and generalizability of the algorithms, data will be from multiple centers and multiple MRI device vendors. Differences in scanning protocols will also be present in the data, e.g. patient with and without an endorectal coil. The dataset is available at <https://promise12.grand-challenge.org/Details/>

There are 50 training cases available for download.

These cases include a transversal T2-weighted MR image of the prostate. The training set is a representative set of the types of MR images acquired in a clinical setting. The data is multi-center and multi-vendor and has different acquisition protocols (e.g. differences in slice thickness, with/without endorectal coil). The set is selected such that there is a spread in prostate sizes and appearance. For each of the cases in the training set, a reference segmentation is also included.

Why do we want to segment the prostate on MR images? Determination of prostate volume (PV) facilitates an assessment of prostate disorders and, for prostate cancer, in conjunction with other parameters, can help predict the pathologic stage of disease, offers insights into the prognosis, and helps predict treatment response. Prostate-specific antigen (PSA) levels have been modified to derive the PSA density by incorporating PV calculations to help guide clinical decisions. The clinical value of prostate-specific antigen density, however, is dependent on the quality of the PV estimate. The accuracy and variability of PV determinations pose limitations to its usefulness in clinical practice. Information on the size/PV, shape, and location of the prostate relative to adjacent organs is also an essential part of surgical planning for prostatectomy, radiation therapy, and emerging minimally invasive therapies, such as cryotherapy and high-intensity focused ultrasound (HIFU). Recently, the high spatial resolution and soft-tissue contrast offered by MRI makes it the most accurate method available for obtaining this kind of information. This, combined with the potential of MRI to localize and grade prostate cancer, has led to a rapid increase in its adoption and increasing research interest in its use for this application. Furthermore, and of particular relevance to the MICCAI community, is the fact that accurate prostate MRI segmentation is an essential pre-processing task for computer-aided detection and diagnostic algorithms, as well as a number of multi-modality image registration algorithms, which aim to enable MRI-derived information on anatomy and tumor location and extent to aid therapy planning and guidance.

1.5. Hematoxylin & Eosin Stained Images

Pathology Datasets Collection from Andrew Janowczyk This page is a collection of some of my open-sourced deep learning work's supplemental materials (i.e., tutorials / code / datasets from papers) <http://www.andrewjanowczyk.com/deep-learning/>

A Great Review of Pathology Datasets: Machine learning methods for histopathological image analysis <https://arxiv.org/pdf/1709.00786.pdf>

GlaS [30] Glands are important histological structures, which has been shown that malignant tumours arising from

glandular epithelium, also known as adenocarcinomas, are the most prevalent form of cancer. Accurate segmentation of glands is often a crucial step to obtain reliable morphological statistics. Nonetheless, the task by nature is very challenging due to the great variation of glandular morphology in different histologic grades. Up until now, the majority of studies focus on gland segmentation in healthy or benign samples, but rarely on intermediate or high grade cancer, and quite often, they are optimised to specific datasets. In GlaS challenge, participants are encouraged to run their gland segmentation algorithms on images of Hematoxylin and Eosin (H&E) stained slides. The dataset is provided together with ground truth annotations by expert pathologists. The dataset is available at <https://warwick.ac.uk/fac/sci/dcs/research/tia/glascontest/about/>

1.6. Fundoscopic & OCT Images

EyePACS [16] This datasets is released along a competition on Kaggle, of which sponsored by EyePACS platform, aiming for comprehensive and automated method of diabetic retinopathy screening using image classification, pattern recognition, and machine learning. The dataset is available at <https://www.kaggle.com/c/diabetic-retinopathy-detection/>

DRISHTI-GS [31] Drishti-GS is a dataset meant for validation of segmenting OD, cup and detecting notching. The images in the Drishti-GS dataset have been collected and annotated by Aravind Eye hospital, Madurai, India. The dataset is divided into two: a training set and a testing set of images. Training images (50) are provided with groundtruths for OD and Cup segmentation and notching information. The dataset is available at <https://cvit.iit.ac.in/projects/mip/drishti-gs/mip-dataset2/Home.php>

REFUGE [26] REFUGE challenge is partnering with OMIA to widen the opportunities to present your work at MICCAI. In addition to the traditional oral and poster presentations of OMIA, REFUGE offers the chance to try your software on a real challenge this year with fundus images. The goal of the challenge is to evaluate and compare automated algorithms for glaucoma detection and optic disc/cup segmentation on a standard dataset of retinal fundus images. The dataset is available at <https://refuge.grand-challenge.org/Home2020/>

DRIVE [32] The Digital Retinal Images for Vessel Extraction (DRIVE) database has been established to enable comparative studies on segmentation of blood vessels in retinal images. The photographs for the DRIVE database were obtained from a diabetic retinopathy screening program in The Netherlands. The set of 40 images has been divided into a training and a test set, both containing 20 images. All human observers that manually segmented the vasculature were instructed and trained by an experi-

enced ophthalmologist. They were asked to mark all pixels for which they were for at least 70% certain that they were vessel. The dataset is available at <https://drive.grand-challenge.org/>

HRF [6] High-Resolution Fundus (HRF) Image Database has been established by a collaborative research group to support comparative studies on automatic segmentation algorithms on retinal fundus images. The public database contains at the moment 15 images of healthy patients, 15 images of patients with diabetic retinopathy and 15 images of glaucomatous patients, along with the segmentation annotations of vessels. The dataset is available at <https://www5.cs.fau.de/research/data/fundus-images/>

MESSIDOR [11] MESSIDOR stands for Methods to Evaluate Segmentation and Indexing Techniques in the field of Retinal Ophthalmology (in French). The Messidor database has been established to facilitate studies on computer-assisted diagnoses of diabetic retinopathy. The dataset is available at <http://www.adcis.net/en/third-party/messidor/>

STARE [14] The STARE (STructured Analysis of the Retina) Project was conceived and initiated in 1975 by Michael Goldbaum, M.D., at the University of California, San Diego. It contains 400 images attached with corresponding diagnosis, including (0) Normal; (1) Hollenhorst Emboli; (2) Branch Retinal Artery Occlusion; (3) Cilio-Retinal Artery Occlusion; (4) Branch Retinal Vein Occlusion; (5) Central Retinal Vein Occlusion; (6) Hemi-Central Retinal Vein Occlusion; (7) Background Diabetic Retinopathy; (8) Proliferative Diabetic Retinopathy; (9) Arteriosclerotic Retinopathy; (10) Hypertensive Retinopathy; (11) Coat's; (12) Macroaneurism; (13) Choroidal Neovascularization; (14) Others. The STARE also provides some segmentation annotations of vessels. The dataset is available at <http://cecas.clemson.edu/ahoover/stare/>

OCT 2017 [19] The dataset obtained 108,312 images (among these, there are 37,206 with choroidal neovascularization, 11,349 with diabetic macular edema, 8,617 with drusen, and 51,140 normal) from 4,686 patients. The dataset is available at <https://data.mendeley.com/datasets/rsbjbr9sj/3>

1.7. Dermoscopy Images

ISIC 2019 [9, 33, 10] The International Skin Imaging Collaboration (ISIC) has developed the ISIC Archive, an international repository of dermoscopic images, for both the purposes of clinical training, and for supporting technical research toward automated algorithmic analysis for skin cancer. The goal for ISIC 2019 is classify dermoscopic images among nine categories: 1. Melanoma; 2. Melanocytic nevus; 3. Basal cell carcinoma; 4. Actinic keratosis; 5. Benign keratosis (solar lentigo / seborrheic keratosis / lichen planus-like keratosis); 6. Dermatofibroma;

7. Vascular lesion; 8. Squamous cell carcinoma; 9. None of the others. There are 25,331 images available for training across 8 different categories. The dataset is available at <https://challenge2019.isic-archive.com/>

References

- [1] Walid Al-Dhabyani, Mohammed Gomaa, Hussien Khaled, and Aly Fahmy. Dataset of breast ultrasound images. *Data in brief*, 28:104863, 2020. 2
- [2] Alexander Andreopoulos and John K Tsotsos. Efficient and generalizable statistical models of shape and appearance for analysis of cardiac mri. *Medical Image Analysis*, 12(3):335–357, 2008. 2, 3
- [3] Teresa Araujo, Guilherme Aresta, Eduardo Castro, Jose Rouco, Paulo Aguiar, Catarina Eloy, Antonio Polonia, and Aurelio Campilho. Classification of breast cancer histology images using convolutional neural networks. *PLOS ONE*, 12(6), 2017. 2
- [4] Olivier Bernard, Alain Lalande, Clement Zotti, Frederick Cervenansky, Xin Yang, Pheng-Ann Heng, Irem Cetin, Karim Lekadir, Oscar Camara, Miguel Angel Gonzalez Ballester, et al. Deep learning techniques for automatic mri cardiac multi-structures segmentation and diagnosis: is the problem solved? *IEEE transactions on medical imaging*, 37(11):2514–2525, 2018. 2, 3
- [5] Nicolas Bloch, Anant Madabhushi, Henkjan Huisman, John Freymann, Justin Kirby, Andinet Enquobahrie, Carl Jaffe, Larry Clarke, and Keyvan Farahani. NCI-ISBI 2013 challenge: automated segmentation of prostate structures. *The Cancer Imaging Archive*, 370, 2015. 2
- [6] Attila Budai, Rüdiger Bock, Andreas Maier, Joachim Hornegger, and Georg Michelson. Robust vessel segmentation in fundus images. *International journal of biomedical imaging*, 2013, 2013. 2, 4
- [7] Enrique J. Carmona, Mariano Rincón, Julián García-Feijó, and José M. Martínez-De-La-Casa. Identification of the optic nerve head with genetic algorithms. *Artificial Intelligence in Medicine*, 43(3):243–259, 2008. 2
- [8] Kenneth Clark, Bruce Vendt, Kirk Smith, John Freymann, Justin Kirby, Paul Koppel, Stephen Moore, Stanley Phillips, David Maffitt, Michael Pringle, et al. The cancer imaging archive (tcia): maintaining and operating a public information repository. *Journal of digital imaging*, 26(6):1045–1057, 2013. 3
- [9] Noel C F Codella, David A Gutman, M Emre Celebi, and Brian Helba, et al. Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging, hosted by the international skin imaging collaboration (isic). *arXiv: Computer Vision and Pattern Recognition*, 2016. 2, 4
- [10] Marc Combalia, Noel C F Codella, Veronica Rotemberg, Brian Helba, Veronica Vilaplana, Ofer Reiter, Allan C Halpern, Susana Puig, and Josep Malvehy. Bcn20000: dermoscopic lesions in the wild. *arXiv: Image and Video Processing*, 2019. 4
- [11] Etienne Decencière, Xiwei Zhang, Guy Cazuguel, Bruno Lay, Béatrice Cochener, Caroline Trone, Philippe Gain, Richard Ordonez, Pascale Massin, Ali Erginay, et al. Feedback on a publicly distributed image database: the messidor database. *Image Analysis & Stereology*, 33(3):231–234, 2014. 2, 4
- [12] Mark Everingham, SM Ali Eslami, Luc Van Gool, Christopher KI Williams, John Winn, and Andrew Zisserman. The pascal visual object classes challenge: A retrospective. *International journal of computer vision*, 111(1):98–136, 2015. 1, 2
- [13] Francisco Fumero, Silvia Alayón, José L Sanchez, Jose Sigut, and M Gonzalez-Hernandez. Rim-one: An open retinal image database for optic nerve evaluation. In *2011 24th international symposium on computer-based medical systems (CBMS)*, pages 1–6. IEEE, 2011. 2
- [14] A. D. Hoover, Valentina Kouznetsova, and Michael Goldbaum. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *IEEE Transactions on Medical Imaging*, 19(3):203–210, 2000. 2, 4
- [15] Jeremy Irvin, Pranav Rajpurkar, Michael Ko, and Yifan Yu, et al. Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison. 1, 2
- [16] Kaggle EyePACS Competition. Identify signs of diabetic retinopathy in eye images. Available at: <https://www.kaggle.com/c/diabetic-retinopathy-detection/> [Accessed 4 May. 2020], 2015. 2, 4
- [17] Kaggle IMCCS Competition. Intel & mobileodt cervical cancer screening-which cancer treatment will be most effective? Available at: <https://www.kaggle.com/c/intel-mobileodt-cervical-cancer-screening/overview/description> [Accessed 4 May. 2020], 2017. 2
- [18] Ali Emre Kavur, M. Alper Selver, Oğuz Dicle, Mustafa Barış, and N. Sinem Gezer. CHAOS - Combined (CT-MR) Healthy Abdominal Organ Segmentation Challenge Data, Apr. 2019. 2
- [19] Daniel S Kermany, Michael Goldbaum, Wenjia Cai, Carolina CS Valentim, Huiying Liang, Sally L Baxter, Alex McKeown, Ge Yang, Xiaokang Wu, Fangbing Yan, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*, 172(5):1122–1131, 2018. 1, 2, 4
- [20] Alex Krizhevsky, Geoffrey Hinton, et al. Learning multiple layers of features from tiny images. 2009. 1, 2
- [21] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012. 1, 2
- [22] Neeraj Kumar, Ruchika Verma, Sanuj Sharma, Surabhi Bhargava, Abhishek Vahadane, and Amit Sethi. A dataset and a technique for generalized nuclear segmentation for computational pathology. *IEEE Transactions on Medical Imaging*, 36(7):1550–1560, 2017. 2
- [23] Geert Litjens, Robert Toth, Wendy van de Ven, Caroline Hoeks, Sjoerd Kerkstra, Bram van Ginneken, Graham Vincent, Gwenael Guillard, Neil Birbeck, Jindang Zhang, et al. Evaluation of prostate segmentation algorithms for mri: the promise12 challenge. *Medical image analysis*, 18(2):359–373, 2014. 2, 3

- [24] Bloch N, Madabhushi A, Huisman H, Freymann J, Kirby J, Grauer M, Enquobahrie A, Jaffe C, Clarke L, and Farahani K. NCI-ISBI 2013 Challenge: Automated Segmentation of Prostate Structures. *The Cancer Imaging Archive*, Available online: <http://doi.org/10.7937/K9/TCIA.2015.zF0vIOPv>, 9, 2015. 2, 3
- [25] Peter Naylor, Marick Laé, Fabien Reyat, and Thomas Walter. Segmentation of nuclei in histopathology images by deep regression of the distance map. *IEEE transactions on medical imaging*, 38(2):448–459, 2018. 2
- [26] José Ignacio Orlando, Huazhu Fu, João Barbosa Breda, and Karel van Keer, et al. Refuge challenge: A unified framework for evaluating automated methods for glaucoma assessment from fundus photographs. *Medical image analysis*, 59:101570, 2020. 2, 4
- [27] Pranav Rajpurkar, Jeremy Irvin, Aarti Bagul, Daisy Ding, Tony Duan, Hershel Mehta, Brandon Yang, Kaylie Zhu, Dillon Laird, and Robyn L Ball. Mura: Large dataset for abnormality detection in musculoskeletal radiographs. 2017. 1
- [28] Holger R Roth, Le Lu, Amal Farag, and Hoo-Chang Shin, et al. Deeporgan: Multi-level deep convolutional networks for automated pancreas segmentation. In *International conference on medical image computing and computer-assisted intervention*, pages 556–564. Springer, 2015. 2
- [29] Arnaud Arindra Adiyoso Setio, Alberto Traverso, Thomas De Bel, Moira SN Berens, Cas van den Bogaard, Piergiorgio Cerello, Hao Chen, Qi Dou, Maria Evelina Fantacci, Bram Geurts, et al. Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the luna16 challenge. *Medical image analysis*, 42:1–13, 2017. 1, 2
- [30] Korsuk Sirinukunwattana, David RJ Snead, and Nasir M Rajpoot. A stochastic polygons model for glandular structures in colon histology images. *IEEE transactions on medical imaging*, 34(11):2366–2378, 2015. 2, 3
- [31] Jayanthi Sivaswamy, S. R. Krishnadas, Gopal Datt Joshi, Madhulika Jain, and A. Ujjwaft Syed Tabish. Drishti-gs: Retinal image dataset for optic nerve head(ONH) segmentation. In *IEEE International Symposium on Biomedical Imaging*, 2014. 2, 4
- [32] Joes Staal, Michael D Abramoff, Meindert Niemeijer, Max A Viergever, and Bram Van Ginneken. Ridge-based vessel segmentation in color images of the retina. *IEEE transactions on medical imaging*, 23(4):501–509, 2004. 2, 4
- [33] Philipp Tschandl, Cliff Rosendahl, and Harald Kittler. The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific Data*, 5(1):180161, 2018. 4
- [34] Moi Hoon Yap, Gerard Pons, Joan Martí, Sergi Ganau, Melcior Sentís, Reyer Zwiggelaar, Adrian K Davison, and Robert Martí. Automated breast ultrasound lesions detection using convolutional neural networks. *IEEE journal of biomedical and health informatics*, 22(4):1218–1226, 2017. 1, 2
- [35] Bolei Zhou, Hang Zhao, Xavier Puig, Sanja Fidler, Adela Barriuso, and Antonio Torralba. Semantic understanding of scenes through the ade20k dataset. *arXiv preprint arXiv:1608.05442*, 2016. 1, 2