

PyTorch Computer Vision Cookbook

Over 70 recipes to master the art of computer vision with deep learning and PyTorch 1.x

Michael Avendi



BIRMINGHAM - MUMBAI

PyTorch Computer Vision Cookbook

Copyright © 2020 Packt Publishing

All rights reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, without the prior written permission of the publisher, except in the case of brief quotations embedded in critical articles or reviews.

Every effort has been made in the preparation of this book to ensure the accuracy of the information presented. However, the information contained in this book is sold without warranty, either express or implied. Neither the author, nor Packt Publishing or its dealers and distributors, will be held liable for any damages caused or alleged to have been caused directly or indirectly by this book.

Packt Publishing has endeavored to provide trademark information about all of the companies and products mentioned in this book by the appropriate use of capitals. However, Packt Publishing cannot guarantee the accuracy of this information.

Commissioning Editor: Amey Varangaonkar

Acquisition Editor: Devika Battike

Content Development Editor: Nathanya Dias

Senior Editor: Ayaan Hoda

Technical Editor: Manikandan Kurup

Copy Editor: Safis Editing

Project Coordinator: Aishwarya Mohan

Proofreader: Safis Editing

Indexer: Manju Arasan

Production Designer: Jyoti Chauhan

First published: March 2020

Production reference: 1200320

Published by Packt Publishing Ltd.

Livery Place

35 Livery Street

Birmingham

B3 2PB, UK.

ISBN 978-1-83864-483-3

www.packt.com



Packt.com

Subscribe to our online digital library for full access to over 7,000 books and videos, as well as industry leading tools to help you plan your personal development and advance your career. For more information, please visit our website.

Why subscribe?

- Spend less time learning and more time coding with practical eBooks and Videos from over 4,000 industry professionals
- Improve your learning with Skill Plans built especially for you
- Get a free eBook or video every month
- Fully searchable for easy access to vital information
- Copy and paste, print, and bookmark content

Did you know that Packt offers eBook versions of every book published, with PDF and ePub files available? You can upgrade to the eBook version at www.packt.com and as a print book customer, you are entitled to a discount on the eBook copy. Get in touch with us at customercare@packtpub.com for more details.

At www.packt.com, you can also read a collection of free technical articles, sign up for a range of free newsletters, and receive exclusive discounts and offers on Packt books and eBooks.

About the author

Michael Avendi is a principal data scientist with vast experience in deep learning, computer vision, and medical imaging analysis. He works on the research and development of data-driven algorithms for various imaging problems, including medical imaging applications. His research papers have been published in major medical journals, including the *Medical Imaging Analysis* journal. Michael Avendi is an active Kaggle participant and was awarded a top prize in a Kaggle competition in 2017.

About the reviewers

Amin Ahmadi Tazehkandi is an Iranian author, software engineer and computer vision expert. He has worked in numerous software companies across the globe and has a long list of awards and achievements, including a countrywide hackathon win, and an award-winning paper as well. Amin is an avid blogger and long-time contributor to the open-source, cross-platform and computer vision developer communities. He is the proud author of *Computer Vision with OpenCV 3 and Qt5*, and *Hands-On Algorithms for Computer Vision*.

I would like to thank the love of my life and partner, Senem, for her patience and support. I would also like to thank little Naz, for helping her daddy type and work faster than ever.

Dr. Matthew Rever has a Ph.D. in Electrical Engineering from the University of Michigan, Ann Arbor. His career revolves around image processing, computer vision, and machine learning for scientific research purposes. He is well versed in C++, a language he has used over the last 20 years. Over the past few years, he has been using Matlab, Python, OpenCV, SciPy, scikit-learn, TensorFlow, and PyTorch. He believes it is important to stay with keep updated with the latest tools available in order to be most productive. He is also the author of Packt's *Computer Vision Projects with Python 3* and *Advanced Computer Vision Projects*.

Packt is searching for authors like you

If you're interested in becoming an author for Packt, please visit authors.packtpub.com and apply today. We have worked with thousands of developers and tech professionals, just like you, to help them share their insight with the global tech community. You can make a general application, apply for a specific hot topic that we are recruiting an author for, or submit your own idea.

Table of Contents

Preface	1
Chapter 1: Getting Started with PyTorch for Deep Learning	9
Technical requirements	10
Installing software tools and packages	11
How to do it...	11
Installing Anaconda	11
Installing PyTorch	12
Verifying the installation	13
Installing other packages	14
How it works...	14
Working with PyTorch tensors	15
How to do it...	16
Defining the tensor data type	16
Changing the tensor's data type	16
Converting tensors into NumPy arrays	17
Converting NumPy arrays into tensors	17
Moving tensors between devices	18
How it works...	19
See also	19
Loading and processing data	20
How to do it...	20
Loading a dataset	21
Data transformation	23
Wrapping tensors into a dataset	25
Creating data loaders	25
How it works...	26
Building models	27
How to do it...	27
Defining a linear layer	27
Defining models using nn.Sequential	28
Defining models using nn.Module	29
Moving the model to a CUDA device	30
Printing the model summary	31
How it works...	32
Defining the loss function and optimizer	33
How to do it...	33
Defining the loss function	33
Defining the optimizer	34
How it works...	34
See also	35
Training and evaluation	35

How to do it...	35
Storing and loading models	37
Deploying the model	38
How it works...	39
There's more...	40
Chapter 2: Binary Image Classification	43
Exploring the dataset	44
Getting ready	44
How to do it...	44
How it works...	47
Creating a custom dataset	48
How to do it...	48
How it works...	50
Splitting the dataset	51
How to do it...	51
How it works...	54
Transforming the data	54
How to do it...	54
How it works...	55
Creating dataloaders	55
How to do it...	56
How it works...	56
Building the classification model	57
How to do it...	58
How it works...	62
See also	64
Defining the loss function	64
How to do it...	65
How it works...	66
See also	66
Defining the optimizer	67
How to do it...	67
How it works...	68
See also	68
Training and evaluation of the model	69
How to do it...	69
How it works...	76
There's more...	78
Deploying the model	78
How to do it...	78
How it works...	81
Model inference on test data	81
Getting ready	82
How to do it...	82

How it works...	83
See also	84
Chapter 3: Multi-Class Image Classification	87
Loading and processing data	88
How to do it...	88
How it works...	95
There's more...	96
See also	96
Building the model	97
How to do it...	97
How it works...	102
There's more...	103
See also	103
Defining the loss function	103
How to do it...	104
How it works...	105
See also	105
Defining the optimizer	106
How to do it...	106
How it works...	107
See also	108
Training and transfer learning	108
How to do it...	109
How it works...	115
See also	117
Deploying the model	117
How to do it...	117
How it works...	121
Chapter 4: Single-Object Detection	123
Exploratory data analysis	125
Getting ready	125
How to do it...	126
How it works...	132
Data transformation for object detection	133
How to do it...	133
How it works...	139
There's more...	140
See also	143
Creating custom datasets	144
How to do it...	144
How it works...	150
Creating the model	152
How to do it...	152
How it works...	155

Defining the loss, optimizer, and IOU metric	156
How to do it...	157
How it works...	160
Training and evaluation of the model	161
How to do it...	161
How it works...	166
Deploying the model	167
How to do it...	167
How it works...	172
Chapter 5: Multi-Object Detection	173
Creating datasets	174
Getting ready	174
How to do it...	176
Creating a custom COCO dataset	176
Transforming data	183
Defining the Dataloaders	187
How it works...	188
Creating a YOLO-v3 model	192
How to do it...	193
Parsing the configuration file	193
Creating PyTorch modules	194
Defining the Darknet model	195
How it works...	196
Defining the loss function	202
How to do it...	203
How it works...	207
Training the model	211
How to do it...	211
How it works...	214
Deploying the model	216
How to do it...	216
How it works...	221
See also	223
Chapter 6: Single-Object Segmentation	225
Creating custom datasets	226
Getting ready	226
How to do it...	227
Data exploration	227
Data augmentation	229
Creating the datasets	230
How it works...	234
Defining the model	236
How to do it...	237
How it works...	240