

Neural Network Computer Vision with OpenCV 5

*Build computer vision solutions
using Python and DNN module*

Gopi Krishna Nuti



www.bpbonline.com

Copyright © 2024 BPB Online

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 BPB Online 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.

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

First published: 2024

Published by BPB Online

WeWork

119 Marylebone Road

London NW1 5PU

UK | UAE | INDIA | SINGAPORE

ISBN 978-93-55516-961

www.bpbonline.com

Dedicated to

India, that is, Bharat

About the Author

Gopi Krishna Nuti is an experienced professional with 21 years of experience in IT industry. He has done his B. Tech in Computer Science from Andhra University, M.S in Business Analytics from State University of New York at Buffalo and an Executive MBA from Amrita University, Bengaluru. He has worked extensively in analytics and software development projects and has delivered award winning products and solutions. He has authored a book about Machine Learning and has multiple patents and research papers against his name. He is a faculty at various training events and a guest faculty at various Engineering Colleges in AP and Telangana. He is a member of the board of studies for Geetanjali Institute of Science and Technology. He is currently working as a Data Science Manager at Autodesk, Bengaluru. He also volunteers for MUST Research and is committed to democratizing AI for all. An incorrigible foodie, he is a passionate teacher and is obsessed with demystifying AI for the next generation of Software developers.

About the Reviewer

Charan is a product manager at Microsoft, where he works on developing innovative solutions for various domains. He has a keen interest in computer vision, natural language processing, and large language models, and how they can be applied to solve real-world problems.

Before joining Microsoft, Charan was a product manager at two different startups, where he led the development of products that dealt with intelligent document recognition and data extraction, and with using computer vision for grading agricultural produce. He has a rich experience in managing cross-functional teams, conducting user research, and launching products in different markets.

Charan enjoys astrophotography, bike riding and cooking in his spare time.

Acknowledgement

First and foremost, I express my heartfelt gratitude to mother Gnanaprasunamba.

Next, I express my sincere thanks to my wife Padma Latha and my son Dheeraj for sacrificing their share of my time and encouraging me to keep writing this book. I owe them a lot and hope to be worthy of their affection for me.

I also thank my extended family for their planned and inadvertent influence on my growth.

I would also like to acknowledge the valuable contributions of my colleagues and coworkers in these past two decades who have graciously taught me much.

I am also thankful to the team of BPB Publications for their guidance and patience in dealing with my eccentricities.

Finally, I would like to thank you, my readers, for your support and feedback.

Preface

Welcome to your essential guide to unraveling the complexities of image processing. Whether you are a seasoned developer or a beginner exploring the world of Computer Vision, this book offers a comprehensive journey from the roots of Computer Vision to practical implementation. It goes beyond theory, offering professionals a practical roadmap for integrating Computer Vision into their projects. With detailed discussions, hands-on code examples, and a focus on applications such as face detection and object recognition, this guide is tailored for those aiming to excel in the dynamic landscape of computer vision applications.

Whether you are in machine learning, automation, or image analysis, this book equips you with the skills to revolutionize your approach to visual data. Each chapter provides practical insights and examples, fostering innovation and excellence in your endeavors. Stay ahead of the curve with "Computer Vision using OpenCV DNN".

Chapter 1: Introduction to Computer Vision - traces the historical roots and the fundamental concepts that underpin Computer Vision.

Chapter 2: Basics of Imaging - dives into the essentials of imaging, laying the foundation for understanding image processing techniques.

Chapter 3: Challenges in Computer Vision - Explores the challenges and complexities encountered in real-world Computer Vision applications.

Chapter 4: Classical Solutions - delves into classical solutions, gaining insights into traditional approaches to image processing.

Chapter 5: Deep Learning and CNNs - Uncovers the power of deep learning and Convolutional Neural Networks (CNNs) in the context of Computer Vision.

Chapter 6: OpenCV DNN Module - Navigates the OpenCV DNN module, mastering its functionalities for efficient deep learning-based image processing.

Chapter 7: Modern Solutions for Image Classification - Elevates your skills by implementing modern solutions for image classification using Python and OpenCV.

Chapter 8: Modern Solutions for Object Detection - Discusses cutting-edge techniques for object detection, enhancing your ability to identify and locate objects in images.

Chapter 9: Faces and Text - Delves into the fascinating realms of face detection and recognition, along with optical character recognition.

Chapter 10: Running the Code – Gives detailed instructions on how to setup the runtime environments needed to run the code provided in the book.

Chapter 11: End-to-end Demo - Concludes your journey with an end-to-end demonstration, bringing together the concepts learned throughout the book.

Code Bundle and Coloured Images

Please follow the link to download the *Code Bundle* and the *Coloured Images* of the book:

<https://rebrand.ly/ehreg50>

The code bundle for the book is also hosted on GitHub at

<https://github.com/bpbpublications/Neural-Network-Computer-Vision-with-OpenCV-5>

In case there's an update to the code, it will be updated on the existing GitHub repository.

We have code bundles from our rich catalogue of books and videos available at

<https://github.com/bpbpublications>. Check them out!

Errata

We take immense pride in our work at BPB Publications and follow best practices to ensure the accuracy of our content to provide with an indulging reading experience to our subscribers. Our readers are our mirrors, and we use their inputs to reflect and improve upon human errors, if any, that may have occurred during the publishing processes involved. To let us maintain the quality and help us reach out to any readers who might be having difficulties due to any unforeseen errors, please write to us at :

errata@bpbonline.com

Your support, suggestions and feedbacks are highly appreciated by the BPB Publications' Family.

Did you know that BPB offers eBook versions of every book published, with PDF and ePub files available? You can upgrade to the eBook version at www.bpbonline.com and as a print book customer, you are entitled to a discount on the eBook copy. Get in touch with us at :

business@bpbonline.com for more details.

At **www.bpbonline.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 BPB books and eBooks.

Piracy

If you come across any illegal copies of our works in any form on the internet, we would be grateful if you would provide us with the location address or website name. Please contact us at **business@bpbonline.com** with a link to the material.

If you are interested in becoming an author

If there is a topic that you have expertise in, and you are interested in either writing or contributing to a book, please visit **www.bpbonline.com**. We have worked with thousands of developers and tech professionals, just like you, to help them share their insights 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.

Reviews

Please leave a review. Once you have read and used this book, why not leave a review on the site that you purchased it from? Potential readers can then see and use your unbiased opinion to make purchase decisions. We at BPB can understand what you think about our products, and our authors can see your feedback on their book. Thank you!

For more information about BPB, please visit **www.bpbonline.com**.

Join our book's Discord space

Join the book's Discord Workspace for Latest updates, Offers, Tech happenings around the world, New Release and Sessions with the Authors:

<https://discord.bpbonline.com>



Table of Contents

1. Introduction to Computer Vision.....	1
Introduction.....	1
Structure.....	2
Objectives	2
History of computer imaging	2
Retrieving information from images	3
Image processing.....	4
Representation	5
Manipulation.....	5
Flexibility	5
Reproducibility	5
Digital image processing.....	5
Conclusion.....	7
Exercises.....	7
2. Basics of Imaging.....	9
Introduction.....	9
Structure.....	10
Objectives	10
Pixels and image representation	11
<i>Pixels</i>	11
Color spaces	12
<i>Primary colors</i>	12
<i>Additive colors</i>	13
<i>Subtractive colors</i>	14
<i>Grayscale</i>	15
<i>Other color spaces</i>	16
<i>Pixels and color spaces</i>	17
<i>Examples</i>	18

Image filetypes.....	19
Video files	20
<i>Images and videos</i>	21
Programming for image data	22
<i>A brief history of computer image programming</i>	22
OpenCV: History and overview.....	23
Image processing code samples	25
<i>Opening, viewing and closing image files</i>	25
<i>CPP code</i>	25
<i>Python code</i>	26
<i>Videos and frames</i>	27
Programming with color spaces.....	30
<i>Grayscale</i>	30
<i>RGB image</i>	31
Conclusion.....	35
Exercises.....	35
3. Challenges in Computer Vision.....	37
Introduction.....	37
Structure.....	37
Objectives	38
Topics in computer vision	38
Complexity in image processing	40
Image classification	41
Object localization	42
Image segmentation.....	46
Character recognition	48
Conclusion.....	49
Exercises.....	49
Key terms.....	50
4. Classical Solutions.....	51
Introduction.....	51

Structure.....	51
Objectives	52
Solutions for challenges in computer vision	52
<i>Classical solutions</i>	53
<i>Modern solutions</i>	53
Algorithm families	53
<i>Morphological operations</i>	54
<i>Erosion and dilation of images</i>	55
<i>Closing and opening images</i>	58
<i>Thresholding</i>	62
<i>Detecting edges and corners</i>	65
<i>Image transformations</i>	68
<i>Region growing</i>	68
<i>Clustering</i>	69
<i>Template matching</i>	72
<i>Watershed algorithm</i>	75
<i>Foreground and background detection</i>	78
<i>Superpixels</i>	81
<i>Image pyramids</i>	82
<i>Convolution</i>	84
Conclusion.....	86
Exercises.....	87
Key terms.....	87
5. Deep Learning and CNNs	89
Introduction.....	89
Structure.....	89
<i>Objectives</i>	90
History of deep learning	90
Perceptron.....	91
Shallow learning networks	93
Deep learning networks	93
Weights, biases, and activation functions	94

<i>Weight</i>	94
<i>Bias</i>	95
<i>Activation function</i>	95
<i>Optimization function</i>	96
Convolutional neural networks	99
<i>CNNs versus fully connected networks</i>	101
Deep learning process.....	103
<i>Training</i>	103
<i>Techniques in training</i>	104
<i>Inference process</i>	104
<i>Techniques/tricks in inference</i>	105
Conclusion.....	106
Key terms.....	106
Exercises.....	106
6. OpenCV DNN Module	107
Introduction.....	107
Structure.....	107
Objectives	108
Deep learning frameworks	108
<i>TensorFlow</i>	108
<i>PyTorch</i>	109
<i>Keras</i>	109
Inference for computer vision	110
<i>Local inferencing</i>	110
<i>Local CPUs</i>	110
<i>Local GPUs</i>	111
<i>Cloud</i>	111
<i>Edge computing</i>	111
OpenCV DNN module.....	111
<i>History</i>	112
<i>Features and limitations</i>	112
<i>Capabilities</i>	112

<i>Limitations</i>	113
<i>Considerations</i>	114
<i>Supported layers</i>	114
<i>Unsupported layers and operations</i>	115
Important classes	116
Conclusion.....	117
Exercises.....	117
7. Modern Solutions for Image Classification	119
Introduction.....	119
Structure.....	119
Objectives	120
CNNs for classification	120
Inception-v3	121
<i>Keras</i>	125
<i>OpenCV DNN module</i>	127
ResNet.....	130
<i>Keras implementation</i>	132
<i>OpenCV DNN implementation</i>	134
MobileNetV2.....	136
<i>Keras implementation</i>	138
<i>OpenCV DNN implementation</i>	141
Comparison of models	143
Parameters for blobFromImage().....	144
Conclusion.....	145
Exercises.....	145
8. Modern Solutions for Object Detection	147
Introduction.....	147
Structure.....	147
Convolutional neural networks architecture for object detection.....	148
Faster region convolutional neural network	149
Single shot multibox detector	154

You only look once	160
YOLOv3.....	161
Overview of NMSBoxes() API.....	169
YOLOv5.....	169
Differences between YOLOv3 and v5.....	171
Obtaining v5 model ONNX file	178
Working with v6, v7 and v8.....	178
Conclusion.....	178
Exercises.....	179
9. Faces and Text	181
Introduction.....	181
Structure.....	181
Objectives	182
Face detection.....	182
Haar cascades	183
Deep learning approaches: YuNet	186
Face recognition.....	189
Face detection versus recognition.....	190
Face recognition using landmarks.....	190
Face recognizer module	191
Labeled Faces in the Wild dataset.....	193
FaceRecognizerSF class.....	194
Comparing faces.....	194
Text recognition	197
Text detection.....	197
Text recognition.....	200
OpenCV Model Zoo.....	204
Conclusion.....	205
Exercises.....	205
Key terms.....	205

10. Running the Code	207
Introduction.....	207
Structure.....	207
Objectives	208
Sequence of steps.....	208
Setting up Anaconda.....	208
<i>Installing Anaconda on Windows.....</i>	<i>209</i>
<i>Installing Anaconda on Ubuntu Linux</i>	<i>210</i>
Installing Git.....	211
<i>Installing Git on Windows.....</i>	<i>211</i>
<i>Installing Git on Ubuntu</i>	<i>212</i>
Setting up Python environment	213
Fetching the code.....	213
<i>Downloading the code</i>	<i>213</i>
<i>Fetch the weights.....</i>	<i>214</i>
Installing the libraries	214
Running the code	214
Conclusion.....	215
Exercises.....	215
11. End-to-end Demo	217
Introduction.....	217
Structure.....	217
Objectives	217
Code	218
<i>main_app.py.....</i>	<i>218</i>
<i>video_app_ui.py</i>	<i>219</i>
<i>image_processor.py.....</i>	<i>225</i>
<i>numberplate_recognizer.py</i>	<i>227</i>
<i>object_detector.py</i>	<i>232</i>
Running the code	238
Application design.....	238

Notes about codes	239
Conclusion.....	239
Exercises.....	240
Index	241-247

CHAPTER 1

Introduction to Computer Vision

Introduction

In a world where computers and cameras communicate seamlessly, the discipline of computer vision emerges as a profound domain. Envision a scenario where your computer assumes the role of an astute companion with the remarkable ability to comprehend visual data, much akin to your comprehension of textual content. Computer vision, in essence, imparts the capacity to perceive and comprehend the world through the lens of images and videos. It is akin to the endowment of sight and cognitive faculties to your computing machine.

Imagine presenting your computer with an image portraying an endearing feline creature. The computer, although lacking the faculty of perception akin to a human, possesses the competence to process the pixel-level data and decipher patterns and structures. It can discern, for instance, that the presence of pointed ears, fine whiskers, and a luxuriant tail coalesce to form the distinctive visage of a cat. The mechanism underpinning this comprehension is none other than image processing.

Image processing embodies the arsenal of tools with which the computer perfects and enhances the visual information at its disposal. It can effectuate alterations such as color correction, noise reduction, or the refinement of edges, endowing the depicted cat with even greater clarity and visual appeal.

Computer vision extends its capabilities beyond the identification of cats. It engenders awe-inspiring feats, including enabling autonomous vehicles to navigate roads and evade obstacles. It is proficient at tallying the human presence in a crowd and deciphering handwritten textual content. Furthermore, it furnishes invaluable assistance to medical practitioners in the identification of ailments from radiographic imagery, such as X-rays.

A noteworthy aspect of computer vision is its capacity for continuous learning and adaptation. Analogous to how human cognition improves with exposure and experience, computer vision is enhanced by accumulating additional data and knowledge. This dynamic field, steeped in innovation, imparts augmented intelligence and utility to technology across diverse domains, be it in the realms of healthcare, security, entertainment, or myriad other spheres. Computer vision, in its essence, bestows upon computers the precious gift of vision and comprehension, ushering in a realm brimming with possibilities.

Structure

The chapter will cover the following topics:

- History of computer imaging
- Retrieving information from images
- Image processing
- Representation
- Manipulation
- Flexibility
- Reproducibility
- Digital image processing

Objectives

The objective of this chapter is to introduce the contents discussed in later chapters. This chapter starts with a history of computer imaging and walks through image representation, processing, and manipulation. The chapter also introduces digital image processing and briefly explains the differences between digital and analog image processing.

History of computer imaging

The history of computer imaging is a fascinating journey that spans several decades. It has evolved from humble beginnings to become an integral part of our daily lives. Let us familiarize ourselves in detail with the history of computer imaging.

The roots of computer imaging can be traced back to the 1950s when computers were in their infancy. Researchers began exploring the idea of using computers to process and generate images. One of the earliest milestones was the development of the **Whirlwind** computer at **Massachusetts Institute of Technology (MIT)**, which could display simple graphics on a screen. In the 1960s, efforts to digitize images started to gain momentum. Researchers devised methods to convert photographs and other analog images into digital form. NASA played a significant role in advancing computer imaging technology by using digital images in space exploration and remote sensing. The 1970s saw the emergence of early computer graphics. The development of devices like the framebuffer allowed computers to display images directly on screens. Companies like Xerox PARC and Atari contributed to the growth of computer graphics, leading to the development of the first video games and interactive **graphical user interfaces (GUIs)**. In the medical field, computer imaging found applications in areas like **Computed Tomography (CT)** and **Magnetic Resonance Imaging (MRI)**, revolutionizing diagnostics. These technologies enabled doctors to visualize the human body's internal structures in previously impossible ways.

The advent of personal computers in the 1980s brought about desktop publishing. Applications like Adobe Photoshop and Adobe Illustrator revolutionized image editing and design. The field of computer vision gained momentum during this period. Researchers focused on teaching computers to interpret and understand images, laying the groundwork for facial recognition, object detection, and more. The 1990s saw the rise of digital photography with the introduction of consumer digital cameras. This technology made it easier for individuals to capture and share digital images. Advances in image sensors, image compression, and storage technologies played a pivotal role in the popularity of digital photography.

The entertainment industry embraced computer imaging for special effects in movies and the development of 3D animation in films like *Toy Story* by Pixar. Video games also evolved with increasingly realistic **computer-generated imagery (CGI)**.

In recent years, deep learning and artificial intelligence have fueled significant advancements in computer imaging. **convolutional neural networks (CNNs)** have revolutionized image recognition and processing. Applications include self-driving cars, facial recognition, medical image analysis, and more.

Today, computer imaging is an integral part of numerous industries, from healthcare to entertainment, and it continues to evolve rapidly. With the growing influence of AI and machine learning, we can expect even more exciting developments in computer imaging in the years to come.

Retrieving information from images

The notion of data being stored in and extracted from images is a significant aspect of computer vision and image processing. Images have been used as carriers of hidden