

masked_image = arcnn.visualize.apply_mask(masked_image
se Morrison's logic era (dilate and erode to find
nel (lines), int
ked

onv the masked image to put black and white (1-bit)
ge_bw = cv.cvtColor(image_bw, cv.COLOR_BGR2GRAY)
esh, image_bw = cv.threshold(image_bw, 220, 255, cv.THRESH

to = sorted(contours, key=cv.contourArea, reverse=True)
find the bounding box of the largest contour
nding = [0, 0, 0, 0]

et the coords the bounding right corner of the bounding
nding_x = bounding_x + bounding_w
nding_y = bounding_y + bounding_h

Adam Geitgey

Second Edition

Python Machine Learning Pro Tips

Second Edition

Adam Geitgey

Second Edition, Version 2.0.0

Contents

1	Why Use Python for Machine Learning?	6
2	Choosing Between Python 2 and Python 3	9
3	NumPy and Vectorization	11
	Memory Efficiency	11
	Automatic Vectorization and Processor Efficiency	13
	Matrix Operations with NumPy	16
4	Pandas	17
5	Scraping Web Data with Pandas	23
6	Plotting data with Matplotlib	28
7	Speeding Up Data Processing with Parallelism	33
	The Normal Approach	34
	Parallelizing Code with Process Pools	36
	When to Use Process Pools	40
	Bypassing the Global Interpreter Lock	41
8	Working with Files Using Pathlib	42
	The Wrong Solution: Building File Paths by Hand	43

Contents

The Old Solution: Python's os.path Module	44
The Better Solution: Python 3's Pathlib	45
9 Images and Exif Orientation	48
How Digital Cameras Auto-Rotate Images	48
Why Does This Break so Many Python Computer Vision Applications?	51
Fixing the Problem	54
10 Reduce Bugs with Static Type Checking	56
Preventing Bugs with Type Annotations	57
How to Declare Types	59
Declaring Complex Types	61
Running a Type Checker	63
When to Use Type Declarations	64