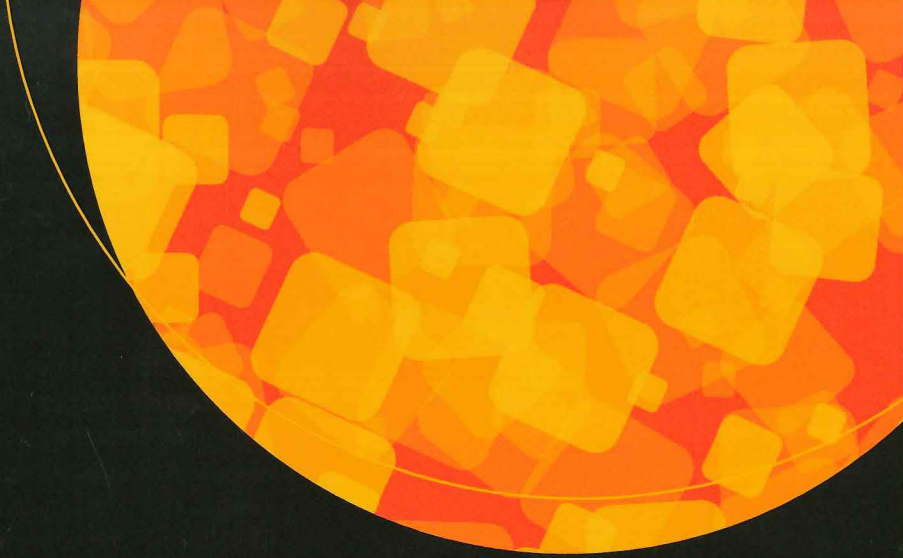




Computer Vision Using Deep Learning

Neural Network Architectures with
Python and Keras

Vaibhav Verdhhan

Apress®

สำนักหอสมุด มหาวิทยาลัยเชียงใหม่

p65
2201-

616580/4X
01250450X
1 22558628

Computer Vision Using Deep Learning

Neural Network Architectures
with Python and Keras



Vaibhav Verdhan

Apress®

Table of Contents

About the Author	xi
About the Technical Reviewer	xiii
Acknowledgments	xv
Introduction	xvii
Foreword	xix
 Chapter 1: Introduction to Computer Vision and Deep Learning	 1
1.1 Technical requirements.....	2
1.2 Image Processing using OpenCV	3
1.2.1 Color detection using OpenCV	4
1.3 Shape detection using OpenCV	6
1.3.1 Face detection using OpenCV.....	9
1.4 Fundamentals of Deep Learning	12
1.4.1 The motivation behind Neural Network	14
1.4.2 Layers in a Neural Network	15
1.4.3 Neuron	16
1.4.4 Hyperparameters	17
1.4.5 Connections and weight of ANN	18
1.4.6 Bias term	18
1.4.7 Activation functions	19
1.4.8 Learning rate	25
1.4.9 Backpropagation	26
1.4.10 Overfitting	28

TABLE OF CONTENTS

1.4.11 Gradient descent	29
1.4.12 Loss functions	31
1.5 How Deep Learning works?	32
1.5.1 Popular Deep Learning libraries	36
1.6 Summary.....	38
1.6.1 Further readings	39
Chapter 2: Nuts and Bolts of Deep Learning for Computer Vision	41
2.1 Technical requirements.....	42
2.2 Deep Learning using TensorFlow and Keras	42
2.3 What is a tensor?	43
2.3.1 What is a Convolutional Neural Network?	45
2.3.2 What is convolution?	46
2.3.3 What is a Pooling Layer?	51
2.3.4 What is a Fully Connected Layer?.....	52
2.4 Developing a DL solution using CNN	53
2.5 Summary.....	64
2.5.1 Further readings	66
Chapter 3: Image Classification Using LeNet.....	67
3.1 Technical requirements.....	68
3.2 Deep Learning architectures.....	68
3.3 LeNet architecture	69
3.4 LeNet-1 architecture	70
3.5 LeNet-4 architecture	71
3.6 LeNet-5 architecture	72
3.7 Boosted LeNet-4 architecture	75
3.8 Creating image classification models using LeNet.....	76
3.9 MNIST classification using LeNet	77

3.10 German traffic sign identification using LeNet	84
3.11 Summary.....	100
3.11.1 Further readings.....	101
Chapter 4: VGGNet and AlexNet Networks.....	103
4.1 Technical requirements.....	104
4.2 AlexNet and VGG Neural Networks.....	104
4.3 What is AlexNet Neural Network?	105
4.4 What is VGG Neural Network?	107
4.5 VGG16 architecture	107
4.6 Difference between VGG16 and VGG19	110
4.7 Developing solutions using AlexNet and VGG.....	111
4.8 Working on CIFAR-10 using AlexNet.....	113
4.9 Working on CIFAR-10 using VGG	128
4.10 Comparing AlexNet and VGG	136
4.11 Working with CIFAR-100	137
4.12 Summary.....	138
4.12.1 Further readings.....	139
Chapter 5: Object Detection Using Deep Learning	141
5.1 Technical requirements.....	142
5.2 Object Detection.....	142
5.2.1 Object classification vs. object localization vs. object detection	143
5.2.2 Use cases of Object Detection.....	144
5.3 Object Detection methods.....	146
5.4 Deep Learning frameworks for Object Detection.....	147
5.4.1 Sliding window approach for Object Detection	148
5.5 Bounding box approach	150

TABLE OF CONTENTS

5.6 Intersection over Union (IoU)	152
5.7 Non-max suppression	154
5.8 Anchor boxes	155
5.9 Deep Learning architectures	157
5.9.1 Region-based CNN (R-CNN)	157
5.10 Fast R-CNN	160
5.11 Faster R-CNN	162
5.12 You Only Look Once (YOLO)	165
5.12.1 Salient features of YOLO	166
5.12.2 Loss function in YOLO	167
5.12.3 YOLO architecture	169
5.13 Single Shot MultiBox Detector (SSD)	172
5.14 Transfer Learning	177
5.15 Python implementation	179
5.16 Summary	182
5.16.1 Further readings	184
Chapter 6: Face Recognition and Gesture Recognition	187
6.1 Technical toolkit	188
6.2 Face recognition	188
6.2.1 Applications of face recognition	190
6.2.2 Process of face recognition	192
6.2.3 DeepFace solution by Facebook	194
6.2.4 FaceNet for face recognition	199
6.2.5 Python implementation using FaceNet	206
6.2.6 Python solution for gesture recognition	208
6.3 Summary	217
6.3.1 Further readings	219

Chapter 7: Video Analytics Using Deep Learning.....	221
7.1 Technical toolkit	222
7.2 Video processing.....	222
7.3 Use cases of video analytics.....	223
7.4 Vanishing gradient and exploding gradient problem	225
7.5 ResNet architecture	230
7.5.1 ResNet and skip connection.....	230
7.5.2 Inception network.....	234
7.5.3 GoogLeNet architecture.....	237
7.5.4 Improvements in Inception v2	239
7.6 Video analytics	243
7.7 Python solution using ResNet and Inception v3.....	244
7.8 Summary.....	254
7.8.1 Further readings	255
Chapter 8: End-to-End Model Development.....	257
8.1 Technical requirements.....	258
8.2 Deep Learning project requirements	258
8.3 Deep Learning project process	262
8.4 Business problem definition	263
8.4.1 Face detection for surveillance	265
8.4.2 Source data or data discovery phase	268
8.5 Data ingestion or data management.....	270
8.6 Data preparation and augmentation	272
8.6.1 Image augmentation.....	274
8.7 Deep Learning modeling process	279
8.7.1 Transfer learning.....	282
8.7.2 Common mistakes/challenges and boosting performance	284

TABLE OF CONTENTS

8.8 Model deployment and maintenance.....289

8.9 Summary.....294

8.9.1 Further readings.....296

References.....297

Major activation functions and layers used in CNN297

Google Colab298

Index.....303