Deep Learning: Modern Artificial Vision

Jaume Ivars Grimalt

March 2025

Acknowledgements

A special thanks to all those who contributed to this work...

Preface

This is the preface of the document...

Contents

Acknowledgements **1**

Preface **2**

I Introduction **1**

1 Introduction **2**

2 Deep Learnining **3**

3 Optimization **4**

4 Datasets **5**

5 Introduction to CNNs **6**
5.1 Image Classification 6

6 Model evaluation **7**

II State of the Art Implementations **8**

7 Classic architectures of CNNs **9**
7.1 FPNs 9
7.2 PANs 9
7.3 Deformable Convolutions 9

8 Classic architectures for Image Segmentation **10**
8.1 U-Net 10
8.2 U2-Net 10

9 Vision Transformers **11**
9.1 DETR 11
9.2 Deformable-DETR 11
9.3 CO-DETR 11

10 Bounding-box Object detection **12**

11 Semantic Segmentation **13**

12 Density Maps **14**

III Generative Models 15

13 Difussion Models **16**
13.1 Stable Difussion 16

14 Text to Image Difussion Models **17**

15 Layout to Image Difussion Models **18**

16 GANs **19**

17 Style Transfer **20**

Part I

Introduction

Chapter 1

Introduction

Chapter 2

Deep Learnining

Chapter 3

Optimization

Chapter 4

Datasets

Chapter 5

Introduction to CNNs

5.1 Image Classification

Chapter 6

Model evaluation

Part II

State of the Art Implementations

Chapter 7

Classic architectures of CNNs

7.1 FPNs

about FPNs

7.2 PANs

7.3 Deformable Convolutions

Chapter 8

Classic architectures for Image Segmentation

8.1 U-Net

8.2 U2-Net

Chapter 9

Vision Transformers

9.1 DETR

9.2 Deformable-DETR

9.3 CO-DETR

Chapter 10

Bounding-box Object detection

Chapter 11

Semantic Segmentation

Chapter 12

Density Maps

Part III

Generative Models

Chapter 13

Difussion Models

13.1 Stable Difussion

Chapter 14

Text to Image Difussion Models

Chapter 15

Layout to Image Difussion Models

Chapter 16

GANs

Chapter 17

Style Transfer