

Multi-faceted Deep Learning

Jenny Benois-Pineau • Akka Zemmari
Editors

Multi-faceted Deep Learning

Models and Data

Editors

Jenny Benois-Pineau
LaBRI UMR 5800
University of Bordeaux
Talence Cedex, France

Akka Zemmari
LaBRI UMR 5800
University of Bordeaux
Talence Cedex, France

ISBN 978-3-030-74477-9 ISBN 978-3-030-74478-6 (eBook)
<https://doi.org/10.1007/978-3-030-74478-6>

© Springer Nature Switzerland AG 2021

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

*“And now, if you will set us to our task,
We will serve you four and twenty hours a
day . . .”*

Rudyard Kipling, The secret of machines

*We dedicate this book to our students.
Be curious, be inventive, persevere and
serve the Dame Science!*

Preface

Today, artificial intelligence approaches penetrate all areas of societal activity. One of its main branches, artificial neural networks, got a new life with the drastic augmentation of computational capacities due to graphical processing units and cloud computing. Neural networks have become deep. Deep learning is now a winner in all supervised machine learning approaches which have been ever used for data mining and decision-making.

These tools are specifically interesting in the field which has been traditionally called “multimedia.” Indeed, this field supplies a huge amount of heterogeneous data: images, video, audio and music, text, and multimodal signals. Furthermore, these data have a spatio-temporal grid structure which is convenient for one of the varieties of deep learning networks, such as convolutional neural networks.

Hence, in this book, we tried to provide a snapshot of methods, models, and data which are being developed or used in this research community. This book is a collective work of selected researchers at the French National Network GDR-ISIS and ACM Special Interest Group on Multimedia. We hope this book will be interesting for young researchers, student, and professionals who are employing existing models and designing new ones in the framework of deep learning.

Bordeaux, France
December 2020

Jenny Benois-Pineau
Akka Zemmari

Acknowledgments

The editors of this book acknowledge the French National Research Network CNRS GDR-ISIS and also ACM-SIGMM which enabled the authors of this book to collaborate together.

Contents

1	Introduction	1
	Jenny Benois-Pineau	
2	Deep Neural Networks: Models and Methods	5
	Akka Zemmari and Jenny Benois-Pineau	
3	Deep Learning for Semantic Segmentation	39
	Alexandre Benoit, Badih Ghattas, Emna Amri, Joris Fournel, and Patrick Lambert	
4	Beyond Full Supervision in Deep Learning	73
	Nicolas Thome	
5	Similarity Metric Learning	103
	Stefan Duffner, Christophe Garcia, Khalid Idrissi, and Atilla Baskurt	
6	Zero-Shot Learning with Deep Neural Networks for Object Recognition	127
	Yannick Le Cacheux, Hervé Le Borgne, and Michel Crucianu	
7	Image and Video Captioning Using Deep Architectures	151
	Danny Francis and Benoit Huet	
8	Deep Learning in Video Compression Algorithms	175
	Ofer Hadar and Raz Birman	
9	3D Convolutional Networks for Action Recognition: Application to Sport Gesture Recognition	199
	Pierre-Etienne Martin, Jenny Benois-Pineau, Renaud Péteri, Akka Zemmari, and Julien Morlier	
10	Deep Learning for Audio and Music	231
	Geoffroy Peeters and Gaël Richard	

11	Explainable AI for Medical Imaging: Knowledge Matters	267
	Pascal Bourdon, Olfa Ben Ahmed, Thierry Urruty, Khalifa Djemal, and Christine Fernandez-Maloigne	
12	Improving Video Quality with Generative Adversarial Networks	293
	Leonardo Galteri, Lorenzo Seidenari, Tiberio Uricchio, Marco Bertini, and Alberto del Bimbo	
13	Conclusion	315
	Jenny Benois-Pineau and Akka Zemmari	