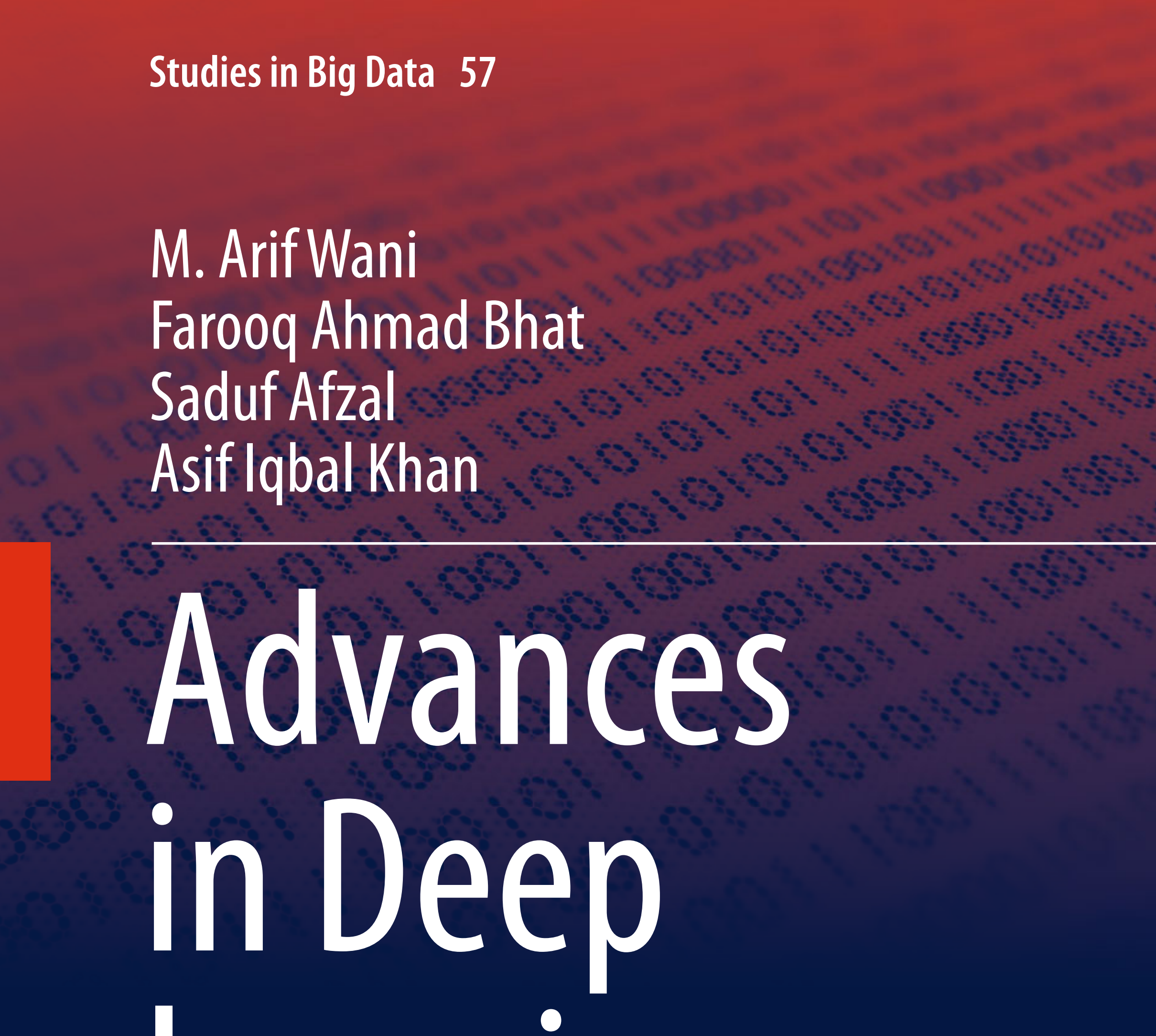
Studies in Big Data 57



M. Arif Wani
Farooq Ahmad Bhat
Saduf Afzal
Asif Iqbal Khan

Advances in Deep Learning

Springer

Studies in Big Data

Volume 57

Series editor

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland

The series “Studies in Big Data” (SBD) publishes new developments and advances in the various areas of Big Data—quickly and with a high quality. The intent is to cover the theory, research, development, and applications of Big Data, as embedded in the fields of engineering, computer science, physics, economics and life sciences. The books of the series refer to the analysis and understanding of large, complex, and/or distributed data sets generated from recent digital sources coming from sensors or other physical instruments as well as simulations, crowd sourcing, social networks or other internet transactions, such as emails or video click streams and other. The series contains monographs, lecture notes and edited volumes in Big Data spanning the areas of computational intelligence including neural networks, evolutionary computation, soft computing, fuzzy systems, as well as artificial intelligence, data mining, modern statistics and Operations research, as well as self-organizing systems. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution, which enable both wide and rapid dissemination of research output.

** Indexing: The books of this series are submitted to ISI Web of Science, DBLP, Ulrichs, MathSciNet, Current Mathematical Publications, Mathematical Reviews, Zentralblatt Math: MetaPress and Springerlink.

More information about this series at http://www.springer.com/series/11970

M. Arif Wani · Farooq Ahmad Bhat ·
Saduf Afzal · Asif Iqbal Khan

Advances in Deep Learning

Springer

M. Arif Wani
Department of Computer Sciences
University of Kashmir
Srinagar, Jammu and Kashmir, India

Farooq Ahmad Bhat
Education Department
Government of Jammu and Kashmir
Kashmir, Jammu and Kashmir, India

Saduf Afzal
Islamic University of Science
and Technology
Kashmir, Jammu and Kashmir, India

Asif Iqbal Khan
Department of Computer Sciences
University of Kashmir
Srinagar, Jammu and Kashmir, India

ISSN 2197-6503 ISSN 2197-6511 (electronic)
Studies in Big Data
ISBN 978-981-13-6793-9 ISBN 978-981-13-6794-6 (eBook)
https://doi.org/10.1007/978-981-13-6794-6

Library of Congress Control Number: 2019932671

© Springer Nature Singapore Pte Ltd. 2020
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 Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

This book discusses the state-of-the-art deep learning models used by researchers recently. Various deep architectures and their components are discussed in detail. Algorithms that are used to train deep architectures with fast convergence rate are illustrated with applications. Various fine-tuning algorithms are discussed for optimizing the deep models. These deep architectures not only are capable of learning complex tasks but can even outperform humans in some dedicated applications.

Despite the remarkable advances in this area, training deep architectures with a huge number of hyper-parameters is an intricate and ill-posed optimization problem. Various challenges are outlined at the end of each chapter. Another issue with deep architectures is that learning becomes computationally intensive when large volumes of data are used for training. The book describes a transfer learning approach for faster training of deep models. The use of this approach is demonstrated in fingerprint datasets.

The book is organized into eight chapters:

Chapter 1 starts with an introduction to machine learning followed by fundamental limitations of traditional machine learning methods. It introduces deep networks and then briefly discusses why to use deep learning and how deep learning works.

Chapter 2 of the book is dedicated to one of the most successful deep learning techniques known as convolutional neural networks (CNNs). The purpose of this chapter is to give its readers an in-depth but easy and uncomplicated explanation of various components of convolutional neural network architectures.

Chapter 3 discusses the training and learning process of deep networks. The aim of this chapter is to provide a simple and intuitive explanation of the backpropagation algorithm for a deep learning network. The training process has been explained step by step with easy and straightforward explanations.

Chapter 4 focuses on various deep learning architectures that are based on CNN. It introduces a reader to block diagrams of these architectures. It discusses how deep learning architectures have evolved while addressing the limitations of previous deep learning networks.

Chapter 5 presents various unsupervised deep learning architectures. The basics of architectures and associated algorithms falling under the unsupervised category are outlined.

Chapter 6 discusses the application of supervised deep learning architecture for face recognition problem. A comparison of the performance of supervised deep learning architecture with traditional face recognition methods is provided in this chapter.

Chapter 7 focuses on the application of convolutional neural networks (CNNs) for fingerprint recognition. This chapter extensively explains automatic fingerprint recognition with complete details of the CNN architecture and methods used to optimize and enhance the performance. In addition, a comparative analysis of deep learning and non-deep learning methods is presented to show the performance difference.

Chapter 8 explains how to apply the unsupervised deep networks to handwritten digit classification problem. It explains how to build a deep learning model in two steps, where unsupervised training is performed during the first step and supervised fine-tuning is carried out during the second step.

Srinagar, India

M. Arif Wani
Farooq Ahmad Bhat
Saduf Afzal
Asif Iqbal Khan

Contents

1 Introduction to Deep Learning 1
  1.1 Introduction 1
  1.2 Shallow Learning 2
  1.3 Deep Learning 2
  1.4 Why to Use Deep Learning 6
  1.5 How Deep Learning Works 8
  1.6 Deep Learning Challenges 9
  Bibliography 10

2 Basics of Supervised Deep Learning 13
  2.1 Introduction 13
  2.2 Convolutional Neural Network (ConvNet/CNN) 13
  2.3 Evolution of Convolutional Neural Network Models 14
  2.4 Convolution Operation 18
  2.5 Architecture of CNN 19
    2.5.1 Convolution Layer 20
    2.5.2 Activation Function (ReLU) 22
    2.5.3 Pooling Layer 25
    2.5.4 Fully Connected Layer 26
    2.5.5 Dropout 27
  2.6 Challenges and Future Research Direction 28
  Bibliography 28

3 Training Supervised Deep Learning Networks 31
  3.1 Introduction 31
  3.2 Training Convolution Neural Networks 31
  3.3 Loss Functions and Softmax Classifier 39
    3.3.1 Mean Squared Error (*L*2) Loss 39
    3.3.2 Cross-Entropy Loss 40
    3.3.3 Softmax Classifier 40

3.4 Gradient Descent-Based Optimization Techniques 41
  3.4.1 Gradient Descent Variants 41
  3.4.2 Improving Gradient Descent for Faster Convergence 43
3.5 Challenges in Training Deep Networks 45
  3.5.1 Vanishing Gradient 45
  3.5.2 Training Data Size 47
  3.5.3 Overfitting and Underfitting 47
  3.5.4 High-Performance Hardware 49
3.6 Weight Initialization Techniques 49
  3.6.1 Initialize All Weights to 0 50
  3.6.2 Random Initialization 50
  3.6.3 Random Weights from Probability Distribution 50
  3.6.4 Transfer Learning 51
3.7 Challenges and Future Research Direction 51
Bibliography 52

4 Supervised Deep Learning Architectures 53
4.1 Introduction 53
4.2 LeNet-5 53
4.3 AlexNet 55
4.4 ZFNet 58
4.5 VGGNet 60
4.6 GoogleNet 63
4.7 ResNet 66
4.8 Densely Connected Convolutional Network (DenseNet) 67
4.9 Capsule Network 71
4.10 Challenges and Future Research Direction 75
Bibliography 75

5 Unsupervised Deep Learning Architectures 77
5.1 Introduction 77
5.2 Restricted Boltzmann Machine (RBM) 77
  5.2.1 Variants of Restricted Boltzmann Machine 80
5.3 Deep Belief Network 81
  5.3.1 Variants of Deep Belief Network 83
5.4 Autoencoders 84
  5.4.1 Variations of Auto Encoders 86
5.5 Deep Autoencoders 88
5.6 Generative Adversarial Networks 91
5.7 Challenges and Future Research Direction 93
Bibliography 93

6 Supervised Deep Learning in Face Recognition 95
6.1 Introduction 95

6.2 Deep Learning Architectures for Face Recognition 95
6.2.1 VGG-Face Architecture 95
6.2.2 Modified VGG-Face Architecture 96
6.3 Performance Comparison of Deep Learning Models for Face Recognition 100
6.3.1 Performance Comparison with Variation in Facial Expression 103
6.3.2 Performance Comparison on Images with Variation in Illumination Conditions 104
6.3.3 Performance Comparison with Variation in Poses 106
6.4 Challenges and Future Research Direction 109
Bibliography 110

7 Supervised Deep Learning in Fingerprint Recognition 111
7.1 Introduction 111
7.2 Fingerprint Features 111
7.3 Automatic Fingerprint Identification System (AFIS) 114
7.3.1 Feature Extraction Stage 114
7.3.2 Minutia Matching Stage 115
7.4 Deep Learning Architectures for Fingerprint Recognition 116
7.4.1 Deep Learning for Fingerprint Segmentation 116
7.4.2 Deep Learning for Fingerprint Classification 123
7.4.3 Model Improvement Using Transfer Learning 127
7.5 Challenges and Future Research Direction 131
Bibliography 132

8 Unsupervised Deep Learning in Character Recognition 133
8.1 Introduction 133
8.2 Datasets of Handwritten Digits 133
8.3 Deep Learning Architectures for Character Recognition 136
8.3.1 Unsupervised Pretraining 136
8.3.2 Supervised Fine Tuning 138
8.4 Performance Comparison of Deep Learning Architectures 145
8.5 Challenges and Future Research Direction 149
Bibliography 149

About the Authors

Prof. M. Arif Wani completed his M.Tech. in Computer Technology at the Indian Institute of Technology, Delhi and his Ph.D. in Computer Vision at Cardiff University, UK. Currently, he is a Professor at the University of Kashmir, having previously served as a Professor at California State University Bakersfield. His main research interests are in gene expression datasets, face recognition techniques/algorithms, artificial neural networks and deep architectures. He has published many papers in reputed journals and conferences in these areas. He was honored with The International Technology Institute Award in 2002 by the International Technology Institute, California, USA. He is a member of many academic and professional bodies, e.g. the Indian Society for Technical Education, Computer Society of India, IEEE USA and Optical Society of America.

Dr. Farooq Ahmad Bhat completed his M.Phil. and Ph.D. in Computer Science at the University of Kashmir. His dissertation focused on ‘Efficient and robust convolutional neural network based models for face recognition’. Currently, his main interests are in artificial intelligence, machine learning and deep learning, areas in which he has published many articles.

Dr. Saduf Afzal teaches at the Islamic University of Science and Technology, Kashmir, India. She completed her BCA, MCA, M.Phil. and Ph.D. at the Department of Computer Science, University of Kashmir. She has also worked as an academic counselor for the MCA program at IGNOU University. Her main research interests are in machine learning, deep learning and neural networks. She has published many articles in high impact journals and conference proceedings.

Dr. Asif Iqbal Khan currently works as a Lecturer in the Higher Education Department, Kashmir, India. He completed his MCA, M.Phil. and Ph.D. at the Department of Computer Science, University of Kashmir. His main research interests are in machine learning, deep learning, and image processing. He is actively publishing in these areas.

Abbreviations

AE	Autoencoder
AI	Artificial intelligence
ANN	Artificial neural network
BN	Batch normalization
BP	Backpropagation
BPAG	Dropout-backpropagation with adaptive gain
BPGP	Dropout-backpropagation with pattern-based gain
CAE	Contractive autoencoder
CD	Contrastive divergence
CDBNs	Convolutional deep belief networks
CL	Convolutional layer
CNN	Convolutional neural network
CNN-AFC	CNN Architecture for Fingerprint Classification
ConvNet	Convolutional neural network
DAE	Denoising autoencoder
DBNs	Deep belief networks
DCT	Discrete cosine transform
DenseNet	Dense convolutional network
EBGM	Elastic bunch graph matching
FDR	False detection rate
GANs	Generative adversarial networks
GD	Gradient descent
GPUs	Graphics processing units
GWT	Gabor wavelet transform
ICA	Independent component analysis
IIIT-D	Indraprastha Institute of Information Technology, Delhi
ILSVRC	ImageNet Large-Scale Visual Recognition Challenge
ILSVRV	ImageNet Large-Scale Visual Recognition Competition
KL	Kullback–Leibler
LDA	Linear discriminant analysis

LRN	Local response normalization
M-DBNs	Modular deep belief networks
MDR	Missed detection rate
MLP	Multilayer perceptron
MrDBN	Multiresolution deep belief network
MSE	Mean squared error
NIST	National Institute of Standards and Technology
NIST-DB4	NIST Special Database 4
ORL	Olivetti Research Ltd face dataset
PCA	Principal component analysis
RBF	Radial basis function
RBM	Restricted Boltzmann machine
ReLU	Rectified Linear Unit
RMS	Root mean square
RoBMs	Robust restricted Boltzmann machines
RTRBMs	Recurrent temporal restricted Boltzmann machines
SGD	Stochastic gradient descent
SVM	Support vector machine
TRBM	Temperature-based restricted Boltzmann machine

Chapter 1
Introduction to Deep Learning

1.1 Introduction

Machine learning systems, with shallow or deep architectures, have ability to learn and improve with experience. The process of machine learning begins with the raw data which is used for extracting useful information that helps in decision-making. The primary aim is to allow a machine to learn useful information just like humans do. At abstract level, machine learning can be carried out using following approaches:

Supervised learning adapts a system such that for a given input data it produces a target output. The learning data is made up of tuples (*attributes*, *label*) where “attributes” represent the input data and “label” represents the target output. The goal here is to adapt the system so that for a new input the system can predict the target output. Supervised learning can use both continuous and discrete types of input data.

Unsupervised learning involves data that comprises of input vectors without any target output. There are different objectives in unsupervised learning, such as clustering, density estimation, and visualization. The goal of clustering is to discover groups of similar data items on the basis of measured or perceived similarities between the data items. The purpose of density estimation is to determine the distribution of the data within the input space. In visualization, the data is projected down from a high-dimensional space to two or three dimensions to view the similar data items.

Semi-supervised learning first uses unlabeled data to learn a feature representation of the input data and then uses the learned feature representation to solve the supervised task. The training dataset can be divided into two parts: the data samples with corresponding labels and the data samples where the labels are not known. Semi-supervised learning can involve not providing with an explicit form of error at each time but only a generalized reinforcement is received giving indication of how the system should change its behavior, and this is sometimes referred to as reinforcement

M. A. Wani et al., *Advances in Deep Learning*, Studies in Big Data 57,
https://doi.org/10.1007/978-981-13-6794-6_1

learning. Reinforcement learning has been successful in applications as diverse as autonomous helicopter flight, robot legged locomotion, cell-phone network routing, marketing strategy selection, factory control and efficient webpage indexing.

1.2 Shallow Learning

Shallow architectures are well understood and perform good on many common machine learning problems, and they are still used in a vast majority of today's machine learning applications. However, there has been an increased interest in deep architectures recently, in the hope to find means to solve more complex real-world problems (e.g., image analysis or natural language understanding) for which shallow architectures are unable to learn models adequately.

1.3 Deep Learning

Deep learning is a new area of machine learning which has gained popularity in recent past. Deep learning refers to the architectures which contain multiple hidden layers (deep networks) to learn different features with multiple levels of abstraction. Deep learning algorithms seek to exploit the unknown structure in the input distribution in order to discover good representations, often at multiple levels, with higher level learned features defined in terms of lower level features.

Conventional machine learning techniques are restricted in the way they process the natural data in its raw form. For decades, constructing a pattern recognition or machine learning system required considerable domain expertise and careful hand engineering to come up with a feature extractor that transformed the raw data (such as pixel values of an image) into suitable internal representation or feature vector from which the learning system, such as a classifier, could detect or classify patterns in the input. Deep learning allows inputting the raw data (pixels in case of image data) to the learning algorithm without first extracting features or defining a feature vector. Deep learning algorithms can learn the right set of features, and it does this in a much better way than extracting these features using hand-coding. Instead of handcrafting a set of rules and algorithms to extract features from raw data, deep learning involves learning these features automatically during the training process.

In deep learning, a problem is realized in terms of hierarchy of concepts, with each concept built on the top of the others. The lower layers of the model encode some basic representation of the problem, whereas higher level layers build upon these lower layers to form more complex concepts.

Given an image, the pixel intensity values are fed as inputs to the deep learning system. A number of hidden layers then extract features from the input image. These hidden layers are built upon each other in a hierarchal fashion. At first, the lower level layers of the network detect only edge-like regions. These edge regions are then used to define corners (where edges intersect) and contours (outlines of objects). The layers in the higher level combine corners and contours to lead to more abstract "object parts" in the next layer. The key aspect of deep learning is that these layers of features are not handcrafted and designed by human engineers; rather, they are learnt from data gradually using a general-purpose learning procedure.

Finally, the output layer classifies the image and obtains the output class label—the output obtained at the output layer is directly influenced by every other node available in the network. This process can be viewed as hierarchical learning as each layer in the network uses the output of previous layers as "building blocks" to construct increasingly more complex concepts at the higher layers. Figure 1.1 compares traditional machine learning approach based on handcrafted features to deep learning approach based on hierarchical representation learning.

Specifically, in deep learning meaningful representations from the input data are learnt by putting emphasis on buiding complicated mapping using a series of sim-

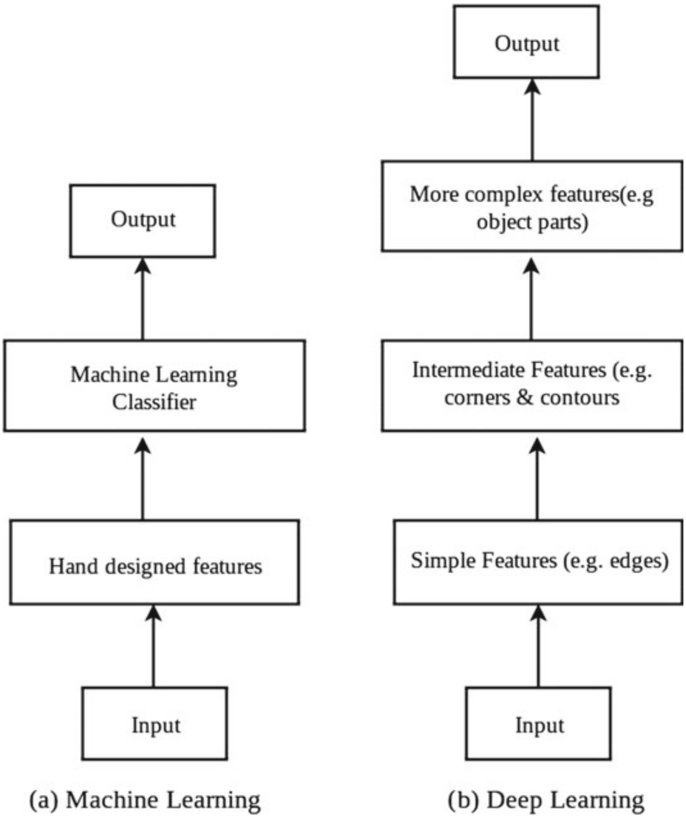


Fig. 1.1 **a** Conventional machine learning using hand-designed feature extraction algorithms **b** deep learning approach using hierarchy of representations that are learnt automatically