Deep Learning for Speech Classification and Speaker Recognition

A strong reference on the problem of signal and speech enhancement, describing the newest developments in this exciting field. The general emphasis is on noise reduction, because of the large number of applications that can benefit from this technology.

Speech Enhancement

This volume is an initiative undertaken by the IEEE Computational Intelligence Society’s Task Force on Security, Surveillance and Defense to consolidate and disseminate the role of CI techniques in the design, development and deployment of security and defense solutions. Applications range from the detection of buried explosive hazards in a battlefield to the control of unmanned underwater vehicles, the delivery of superior video analytics for protecting critical infrastructures or the development of stronger intrusion detection systems and the design of military surveillance networks. Defense scientists, industry experts, academicians and practitioners alike will all benefit from the wide spectrum of successful applications compiled in this volume. Senior undergraduate or graduate students may also discover uncharted territory for their own research endeavors.

2000 Sentences (South-East England) (the)

This book will help you to understand the realms of Deep Learning from A-Z. In today's modern world, Deep Learning has taken over the reins of Machine Learning and Artificial Intelligence. We are all familiar with the term "Deep Learning." But have you ever wondered what it really is? This book will cover all the major features of Deep Learning and how you can use it practically with the help of Python language. It has been designed as a guide to master Deep Learning. You will learn about the basics of Deep Learning, its models, supervised and unsupervised learning, how to perform GUI testing with the help of Deep Learning, and installing Jupyter. It also covers topics like neural networks, pattern recognition, and ANN. You will learn about predictive analysis in Python, Random forest model in Python, and common algorithms used in automation. We also cover the Speech recognition mechanism using Deep Learning with practical examples like Markov models, MFCC Algorithm, and DTW based recognition. Don't know how to use Robot Framework with Python? Don't worry. We will help you out. What about Python key points for Deep Learning? We got you covered. If you are looking for a guide to implement Principal Component Analysis in Python, then this book is for...
you. You will also learn about bulk image compression and front end GUI testing. It has been designed to prepare you to understand the word embedding in Python and linear regression modeling. So what are you waiting for? It is your chance to become a master of Deep Learning. Click the BUY NOW button to get started.

Dynamic Speech Models

Deep Learning

This textbook explains Deep Learning Architecture, with applications to various NLP Tasks, including Document Classification, Machine Translation, Language Modeling, and Speech Recognition. With the widespread adoption of deep learning, natural language processing (NLP), and speech applications in many areas (including Finance, Healthcare, and Government) there is a growing need for one comprehensive resource that maps deep learning techniques to NLP and speech and provides insights into using the tools and libraries for real-world applications. Deep Learning for NLP and Speech Recognition explains recent deep learning methods applicable to NLP and speech, provides state-of-the-art approaches, and offers real-world case studies with code to provide hands-on experience. Many books focus on deep learning theory or deep learning for NLP-specific tasks while others are cookbooks for tools and libraries, but the constant flux of new algorithms, tools, frameworks, and libraries in a rapidly evolving landscape means that there are few available texts that offer the material in this book. The book is organized into three parts, aligning to different groups of readers and their expertise. The three parts are: Machine Learning, NLP, and Speech Introduction. The first part has three chapters that introduce readers to the fields of NLP, speech recognition, deep learning, and machine learning with basic theory and hands-on case studies using Python-based tools and libraries. Deep Learning Basics. The five chapters in the second part introduce deep learning and various topics that are crucial for speech and text processing, including word embeddings, convolutional neural networks, recurrent neural networks, and speech recognition basics. Theory, practical tips, state-of-the-art methods, experimentations and analysis in using the methods discussed in theory on real-world tasks. Advanced Deep Learning Techniques for Text and Speech. The third part has five chapters that discuss the latest and cutting-edge research in the areas of deep learning that intersect with NLP and speech. Topics including attention mechanisms, memory augmented networks, transfer learning, multi-task learning, domain adaptation, reinforcement learning, and end-to-end deep learning for speech recognition are covered using case studies.

Deep Learning

"While deep learning algorithms have made significant progress in automatic speech recognition and natural language processing, they require a significant amount of labelled training data to perform effectively. As such, these applications have not been extended to languages that have only limited amount of data available, such as extinct or endangered languages. Another problem caused by the rise of deep learning is that individuals with malicious intents have been able to leverage these algorithms to create fake contents that can pose serious harm to security and public safety. In this work, we explore the solutions to both of these problems. First, we investigate different data augmentation methods and acoustic architecture designs to improve automatic speech recognition performance on low-resource languages. Data augmentation for audio often involves changing the characteristic of the audio without modifying the ground truth. For example, different background noise can be added to an utterance while maintaining the content of the speech. We also explored how different acoustic model paradigms and complexity affect performance on low-resource languages. These methods are evaluated on Seneca, an endangered language spoken by a Native American tribe, and Iban, a low-resource language spoken in Malaysia and Brunei. Secondly, we explore methods to determine speaker identification and audio spoofing detection. A spoofing attack involves using either a text-to-speech voice conversion application to generate audio that mimic the identity of a target speaker. These methods are evaluated on the ASVSpoof 2019 Logical Access dataset containing audio generated using various methods of voice conversion and text-to-speech synthesis."--Abstract.

Proceedings of the International Conference on Data Engineering and Communication Technology
This two-volume book contains research work presented at the First International Conference on Data Engineering and Communication Technology (ICDECT) held during March 10–11, 2016 at Lavasa, Pune, Maharashtra, India. The book discusses recent research technologies and applications in the field of Computer Science, Electrical and Electronics Engineering. The aim of the Proceedings is to provide cutting-edge developments taking place in the field data engineering and communication technologies which will assist the researchers and practitioners from both academia as well as industry to advance their field of study.

**Python Deep Learning Cookbook**

Deep learning methods are achieving state-of-the-art results on challenging machine learning problems such as describing photos and translating text from one language to another. In this new laser-focused Ebook, finally cut through the math, research papers and patchwork descriptions about natural language processing. Using clear explanations, standard Python libraries and step-by-step tutorial lessons you will discover what natural language processing is, the promise of deep learning in the field, how to clean and prepare text data for modeling, and how to develop deep learning models for your own natural language processing projects.

**New Era for Robust Speech Recognition**

An accessible introduction to the artificial intelligence technology that enables computer vision, speech recognition, machine translation, and driverless cars. Deep learning is an artificial intelligence technology that enables computer vision, speech recognition in mobile phones, machine translation, AI games, driverless cars, and other applications. When we use consumer products from Google, Microsoft, Facebook, Apple, or Baidu, we are often interacting with a deep learning system. In this volume in the MIT Press Essential Knowledge series, computer scientist John Kelleher offers an accessible and concise but comprehensive introduction to the fundamental technology at the heart of the artificial intelligence revolution. Kelleher explains that deep learning enables data-driven decisions by identifying and extracting patterns from large datasets; its ability to learn from complex data makes deep learning ideally suited to take advantage of the rapid growth in big data and computational power. Kelleher also explains some of the basic concepts in deep learning, presents a history of advances in the field, and discusses the current state of the art. He describes the most important deep learning architectures, including autoencoders, recurrent neural networks, and long short-term networks, as well as such recent developments as Generative Adversarial Networks and capsule networks. He also provides a comprehensive (and comprehensible) introduction to the two fundamental algorithms in deep learning: gradient descent and backpropagation. Finally, Kelleher considers the future of deep learning—major trends, possible developments, and significant challenges.

**Automatic Snooker-playing Robot with Speech Recognition Using Deep Learning**

Provides an overview of general deep learning methodology and its applications to a variety of signal and information processing tasks

**Text, Speech, and Dialogue**

Deep learning is the state-of-the-art technique in machine learning with applications in speech recognition. In this study, an efficient system is formulated to process large amounts of speech data within the deep learning framework by harnessing the parallel processing power of High-Performance Computing oriented Graphics Processing Unit (GPU). This thesis focuses on applications of this approach to address stressed speech classification as well as discrimination between different flavors of noise-free speech under Lombard Effect. Different architectures of deep neural networks (DNN) are explored to build state-of-the-art classifiers for detection and classification of stressed speech and Lombard Effect flavors. Furthermore, applications of deep networks are explored to improve current state-of-the-art speaker recognition systems. Further integration of discriminative deep architectures is accomplished for unsupervised methods in training front-ends for Speaker Recognition Evaluation systems.
100 recipes that teach you how to perform various machine learning tasks in the real world. About This Book Understand which algorithms to use in a given context with the help of this exciting recipe-based guide. Learn about perceptrons and see how they are used to build neural networks. Stuck while making sense of images, text, speech, and real estate? This guide will come to your rescue, showing you how to perform machine learning for each one of these using various techniques. Who This Book Is For This book is for Python programmers who are looking to use machine-learning algorithms to create real-world applications. This book is friendly to Python beginners, but familiarity with Python programming would certainly be useful to play around with the code. What You Will Learn Explore classification algorithms and apply them to the income bracket estimation problem. Understand how to perform market segmentation using unsupervised learning. Explore data visualization techniques to interact with your data in diverse ways. Find out how to build a recommendation engine. Understand how to interact with text data and build models to analyze it. Work with speech data and recognize spoken words using Hidden Markov Models. Analyze stock market data using Conditional Random Fields. Work with image data and build systems for image recognition and biometric face recognition. Grasp how to use deep neural networks to build an optical character recognition system. In Detail Machine learning is becoming increasingly pervasive in the modern data-driven world. It is used extensively across many fields such as search engines, robotics, self-driving cars, and more. With this book, you will learn how to perform various machine learning tasks in different environments. We'll start by exploring a range of real-life scenarios where machine learning can be used, and look at various building blocks. Throughout the book, you'll use a wide variety of machine learning algorithms to solve real-world problems and use Python to implement these algorithms. You'll discover how to deal with various types of data and explore the differences between machine learning paradigms such as supervised and unsupervised learning. We also cover a range of regression techniques, classification algorithms, predictive modeling, data visualization techniques, recommendation engines, and more with the help of real-world examples. Style and approach You will explore various real-life scenarios in this book where machine learning can be used, and learn about different building blocks of machine learning using independent recipes in the book.

**Deep Learning**

Abstract: Research on natural language processing, such as for image and speech recognition, is rapidly changing focus from statistical methods to neural networks. In this study, we introduce speech recognition capabilities along with computer vision to allow a robot to play snooker completely by itself. The color of the ball to be pocketed is provided as an audio input using an audio device such as a microphone. The system is able to recognize the color from the input using a trained deep learning network. The system then commands the camera to locate the ball of the identified color on a snooker table by using computer vision. To pocket the target ball, the system then predicts the best shot using an algorithm. This activity can be executed accurately based on the efficiency of the trained deep learning model.

**Make Python Talk**

This book constitutes the proceedings of the 24th International Conference on Text, Speech, and Dialogue, TSD 2021, held in Olomouc, Czech Republic, in September 2021. The 2 keynote speeches and 46 papers presented in this volume were carefully reviewed and selected from 101 submissions. The topical sections "Text", "Speech", and "Dialogue" deal with the following issues: speech recognition; corpora and language resources; speech and spoken language generation; tagging, classification and parsing of text and speech; semantic processing of text and speech; integrating applications of text and speech processing; automatic dialogue systems; multimodal techniques and modelling, and others. Due to the COVID-19 pandemic the conference was held in a "hybrid" mode.

**2020 5th International Conference on Communication and Electronics Systems (ICCES)**

Automatic speech recognition is an active field of study in artificial intelligence and machine learning whose aim is to generate machines that communicate with people via speech. Speech is an information-rich signal that contains paralinguistic information as well as linguistic information. Emotion is one key instance of paralinguistic information that is, in part, conveyed by speech. Developing machines that understand paralinguistic information, such as emotion, facilitates the human-machine communication
as it makes the communication more clear and natural. In the current study, the efficacy of convolutional neural networks in recognition of speech emotions has been investigated. Wide-band spectrograms of the speech signals were used as the input features of the networks. The networks were trained on speech signals that were generated by the actors while acting a specific emotion. The speech databases with different languages were used to train and evaluate our models. The training data on each database were augmented with two levels of augmentations. The dropout technique was implemented to regularize the networks. Our results showed that the gender-independent, language-independent CNN models achieved the state-of-the-art accuracy, outperformed previously reported results in the literature, and emulated or even outperformed human performance over the benchmark databases. Future work is warranted to examine the capability of the deep learning models in speech emotion recognition using daily-life speech signals.

Deep Learning

In recent years, deep learning has fundamentally changed the landscapes of a number of areas in artificial intelligence, including speech, vision, natural language, robotics, and game playing. In particular, the striking success of deep learning in a wide variety of natural language processing (NLP) applications has served as a benchmark for the advances in one of the most important tasks in artificial intelligence. This book reviews the state of the art of deep learning research and its successful applications to major NLP tasks, including speech recognition and understanding, dialogue systems, lexical analysis, parsing, knowledge graphs, machine translation, question answering, sentiment analysis, social computing, and natural language generation from images. Outlining and analyzing various research frontiers of NLP in the deep learning era, it features self-contained, comprehensive chapters written by leading researchers in the field. A glossary of technical terms and commonly used acronyms in the intersection of deep learning and NLP is also provided. The book appeals to advanced undergraduate and graduate students, post-doctoral researchers, lecturers and industrial researchers, as well as anyone interested in deep learning and natural language processing.

Deep Learning for NLP and Speech Recognition

With the widespread adoption of deep learning, natural language processing (NLP), and speech applications in many areas (including Finance, Healthcare, and Government) there is a growing need for one comprehensive resource that maps deep learning techniques to NLP and speech and provides insights into using the tools and libraries for real-world applications. Deep Learning for NLP and Speech Recognition explains recent deep learning methods applicable to NLP and speech, provides state-of-the-art approaches, and offers real-world case studies with code to provide hands-on experience. The book is organized into three parts, aligning to different groups of readers and their expertise. The three parts are: Machine Learning, NLP, and Speech Introduction. The first part has three chapters that introduce readers to the fields of NLP, speech recognition, deep learning and machine learning with basic theory and hands-on case studies using Python-based tools and libraries. Deep Learning Basics. The five chapters in the second part introduce deep learning and various topics that are crucial for speech and text processing, including word embeddings, convolutional neural networks, recurrent neural networks and speech recognition basics. Theory, practical tips, state-of-the-art methods, experimentation and analysis in using the methods discussed in theory on real-world tasks. Advanced Deep Learning Techniques for Text and Speech. The third part has five chapters that discuss the latest and cutting-edge research in the areas of deep learning that intersect with NLP and speech. Topics including attention mechanisms, memory augmented networks, transfer learning, multi-task learning, domain adaptation, reinforcement learning, and end-to-end deep learning for speech recognition are covered using case studies.

From Natural to Artificial Intelligence

Speech Emotion Recognition Using Convolutional Neural Networks

A project-based book that teaches beginning Python programmers how to build working, useful, and fun voice-controlled applications. This fun, hands-on book will take your basic Python skills to the next level as you build voice-controlled apps to use in your daily life. Starting with a Python refresher and an introduction to speech-recognition/text-to-speech functionalities, you’ll soon ease into more
advanced topics, like making your own modules and building working voice-controlled apps. Each chapter scaffolds multiple projects that allow you to see real results from your code at a manageable pace, while end-of-chapter exercises strengthen your understanding of new concepts. You’ll design interactive games, like Connect Four and Tic-Tac-Toe, and create intelligent computer opponents that talk and take commands; you’ll make a real-time language translator, and create voice-activated financial-market apps that track the stocks or cryptocurrencies you are interested in. Finally, you’ll load all of these features into the ultimate virtual personal assistant – a conversational VPA that tells jokes, reads the news, and gives you hands-free control of your email, browser, music player, desktop files, and more. Along the way, you’ll learn how to: ● Build Python modules, implement animations, and integrate live data into an app ● Use web-scraping skills for voice-controlling podcasts, videos, and web searches ● Fine-tune the speech recognition to accept a variety of input ● Associate regular tasks like opening files and accessing the web with speech commands ● Integrate functionality from other programs into a single VPA with computational knowledge engines to answer almost any question. Packed with cross-platform code examples to download, practice activities and exercises, and explainer images, you’ll quickly become proficient in Python coding in general and speech recognition/text to speech in particular.

Speech and Natural Language

Speech Dereverberation gathers together an overview, a mathematical formulation of the problem and the state-of-the-art solutions for dereverberation. Speech Dereverberation presents current approaches to the problem of reverberation. It provides a review of topics in room acoustics and also describes performance measures for dereverberation. The algorithms are then explained with mathematical analysis and examples that enable the reader to see the strengths and weaknesses of the various techniques, as well as giving an understanding of the questions still to be addressed. Techniques rooted in speech enhancement are included, in addition to a treatment of multichannel blind acoustic system identification and inversion. The TRINICON framework is shown in the context of dereverberation to be a generalization of the signal processing for a range of analysis and enhancement techniques. Speech Dereverberation is suitable for students at masters and doctoral level, as well as established researchers.

Video Mining

Deep learning and neural network research has grown significantly in the fields of automatic speech recognition (ASR) and speaker recognition. Compared to traditional methods, deep learning-based approaches are more powerful in learning representation from data and building complex models. In this dissertation, we focus on representation learning and modeling using neural network-based approaches for speech and speaker recognition. In the first part of the dissertation, we present two novel neural network-based methods to learn speaker-specific and phoneme-invariant features for short-utterance speaker verification. We first propose to learn a spectral feature mapping from each speech signal to the corresponding subglottal acoustic signal which has less phoneme variation, using deep neural networks (DNNs). The estimated subglottal features show better speaker-separation ability and provide complementary information when combined with traditional speech features on speaker verification tasks. Additional, we propose another DNN-based mapping model, which maps the speaker representation extracted from short utterances to the speaker representation extracted from long utterances of the same speaker. Two non-linear regression models using an autoencoder are proposed to learn this mapping, and they both improve speaker verification performance significantly. In the second part of the dissertation, we design several new neural network models which take raw speech features (either complex Discrete Fourier Transform (DFT) features or raw waveforms) as input, and perform the feature extraction and phone classification jointly. We first propose a unified deep Highway (HW) network with a time-delayed bottleneck layer (TDB), in the middle, for feature extraction. The TDB-HW networks with complex DFT features as input provide significantly lower error rates compared with hand-designed spectrum features on large-scale keyword spotting tasks. Next, we present a 1-D Convolutional Neural Network (CNN) model, which takes raw waveforms as input and uses convolutional layers to do hierarchical feature extraction. The proposed 1-D CNN model outperforms standard systems with hand-designed features. In order to further reduce the redundancy of the 1-D CNN model, we propose a filter sampling and combination (FSC) technique, which can reduce the model size by 70% and still improve the performance on ASR tasks. In the third part of dissertation, we propose two novel neural-network models for sequence modeling. We first propose an attention mechanism for acoustic sequence modeling. The attention mechanism can
automatically predict the importance of each time step and select the most important information from sequences. Secondly, we present a sequence-to-sequence based spelling correction model for end-to-end ASR. The proposed correction model can effectively correct errors made by the ASR systems.

2019 34th International Technical Conference on Circuits Systems, Computers and Communications (ITC CSCC)

This book focuses on the fundamentals of deep learning along with reporting on the current state-of-the-art research on deep learning. In addition, it provides an insight of deep neural networks in action with illustrative coding examples. Deep learning is a new area of machine learning research which has been introduced with the objective of moving ML closer to one of its original goals, i.e. artificial intelligence. Deep learning was developed as an ML approach to deal with complex input-output mappings. While traditional methods successfully solve problems where final value is a simple function of input data, deep learning techniques are able to capture composite relations between non-immediately related fields, for example between air pressure recordings and English words, millions of pixels and textual description, brand-related news and future stock prices and almost all real world problems. Deep learning is a class of nature inspired machine learning algorithms that uses a cascade of multiple layers of nonlinear processing units for feature extraction and transformation. Each successive layer uses the output from the previous layer as input. The learning may be supervised (e.g. classification) and/or unsupervised (e.g. pattern analysis) manners. These algorithms learn multiple levels of representations that correspond to different levels of abstraction by resorting to some form of gradient descent for training via backpropagation. Layers that have been used in deep learning include hidden layers of an artificial neural network and sets of propositional formulas. They may also include latent variables organized layer-wise in deep generative models such as the nodes in deep belief networks and deep boltzmann machines. Deep learning is part of state-of-the-art systems in various disciplines, particularly computer vision, automatic speech recognition (ASR) and human action recognition.

Multi-access Edge Computing: Software Development at the Network Edge

Explore deep learning applications, such as computer vision, speech recognition, and chatbots, using frameworks such as TensorFlow and Keras. This book helps you to ramp up your practical know-how in a short period of time and focuses you on the domain, models, and algorithms required for deep learning applications. Deep Learning with Applications Using Python covers topics such as chatbots, natural language processing, and face and object recognition. The goal is to equip you with the concepts, techniques, and algorithm implementations needed to create programs capable of performing deep learning. This book covers convolutional neural networks, recurrent neural networks, and multilayer perceptrons. It also discusses popular APIs such as IBM Watson, Microsoft Azure, and scikit-learn. What You Will Learn Work with various deep learning frameworks such as TensorFlow, Keras, and scikit-learn. Use face recognition and face detection capabilities Create speech-to-text and text-to-speech functionality Engage with chatbots using deep learning Who This Book Is For Data scientists and developers who want to adapt and build deep learning applications.

Deepfake Detection and Low-resource Language Speech Recognition Using Deep Learning

Neural Network Based Representation Learning and Modeling for Speech and Speaker Recognition

Solve different problems in modelling deep neural networks using Python, Tensorflow, and Keras with this practical guide About This Book Practical recipes on training different neural network models and tuning them for optimal performance Use Python frameworks like TensorFlow, Caffe, Keras, Theano for Natural Language Processing, Computer Vision, and more A hands-on guide covering the common as well as the not so common problems in deep learning using Python Who This Book Is For This book is intended for machine learning professionals who are looking to use deep learning algorithms to create real-world applications using Python. Thorough understanding of the machine learning concepts and Python libraries such as NumPy, SciPy and scikit-learn is expected. Additionally, basic knowledge in linear algebra and calculus is desired. What You Will Learn Implement different neural network
models in Python. Select the best Python framework for deep learning such as PyTorch, TensorFlow, MXNet and Keras. Apply tips and tricks related to neural networks internals, to boost learning performances. Consolidate machine learning principles and apply them in the deep learning field. Reuse and adapt Python code snippets to everyday problems. Evaluate the cost/benefits and performance implication of each discussed solution. In detail, Deep Learning is revolutionizing a wide range of industries. For many applications, deep learning has proven to outperform humans by making faster and more accurate predictions. This book provides a top-down and bottom-up approach to demonstrate deep learning solutions to real-world problems in different areas. These applications include Computer Vision, Natural Language Processing, Time Series, and Robotics. The Python Deep Learning Cookbook presents technical solutions to the issues presented, along with a detailed explanation of the solutions. Furthermore, a discussion on corresponding pros and cons of implementing the proposed solution using one of the popular frameworks like TensorFlow, PyTorch, Keras and CNTK is provided. The book includes recipes that are related to the basic concepts of neural networks. All techniques, as well as classical networks topologies. The main purpose of this book is to provide Python programmers a detailed list of recipes to apply deep learning to common and not-so-common scenarios. Style and approach: Unique blend of independent recipes arranged in the most logical manner.

**Automatic Speech Recognition**

Intelligent Speech Signal Processing investigates the utilization of speech analytics across several systems and real-world activities, including sharing data analytics, creating collaboration networks between several participants, and implementing video-conferencing in different application areas. Chapters focus on the latest applications of speech data analysis and management tools across different recording systems. The book emphasizes the multidisciplinary nature of the field, presenting different applications and challenges with extensive studies on the design, development and management of intelligent systems, neural networks and related machine learning techniques for speech signal processing. Highlights different data analytics techniques in speech signal processing, including machine learning and data mining. Illustrates different applications and challenges across the design, implementation and management of intelligent systems and neural networks techniques for speech signal processing. Includes coverage of biomodal speech recognition, voice activity detection, spoken language and speech disorder identification, automatic speech to speech summarization, and convolutional neural networks.

**Deep Learning with Applications Using Python**

"Millions of people around the world are diagnosed with neurological disorders like Parkinson’s, Cerebral Palsy or Amyotrophic Lateral Sclerosis. Due to the neurological damage as the disease progresses, the person suffering from the disease loses control of muscles, along with speech deterioration. Speech deterioration is due to neuro motor condition that limits manipulation of the articulators of the vocal tract, the condition collectively called as dysarthria. Even though dysarthric speech is grammatically and syntactically correct, it is difficult for humans to understand and for Automatic Speech Recognition (ASR) systems to decipher. With the emergence of deep learning, speech recognition systems have improved a lot compared to traditional speech recognition systems, which use sophisticated preprocessing techniques to extract speech features. In this digital era, there are still many documents that are handwritten many of which need to be digitized. Offline handwriting recognition involves recognizing handwritten characters from images of handwritten text (i.e. scanned documents). This is an interesting task as it involves sequence learning with computer vision. The task is more difficult than Optical Character Recognition (OCR), because handwritten letters can be written in virtually infinite different styles. This thesis proposes exploiting deep learning techniques like Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) for offline handwriting recognition. For speech recognition, we compare traditional methods for speech recognition with recent deep learning methods. Also, we apply speaker adaptation methods both at feature level and at parameter level to improve recognition of dysarthric speech."—Abstract.

**Speech Dereverberation**

All aspects of Circuits Systems, Computers and Communications
Python Machine Learning Cookbook

Robust Automatic Speech Recognition: A Bridge to Practical Applications establishes a solid foundation for automatic speech recognition that is robust against acoustic environmental distortion. It provides a thorough overview of classical and modern noise-and reverberation robust techniques that have been developed over the past thirty years, with an emphasis on practical methods that have been proven to be successful and which are likely to be further developed for future applications. The strengths and weaknesses of robustness-enhancing speech recognition techniques are carefully analyzed. The book covers noise-robust techniques designed for acoustic models which are based on both Gaussian mixture models and deep neural networks. In addition, a guide to selecting the best methods for practical applications is provided. The reader will: Gain a unified, deep and systematic understanding of the state-of-the-art technologies for robust speech recognition Learn the links and relationship between alternative technologies for robust speech recognition Be able to use the technology analysis and categorization detailed in the book to guide future technology development Be able to develop new noise-robust methods in the current era of deep learning for acoustic modeling in speech recognition The first book that provides a comprehensive review on noise and reverberation robust speech recognition methods in the era of deep neural networks Connects robust speech recognition techniques to machine learning paradigms with rigorous mathematical treatment Provides elegant and structural ways to categorize and analyze noise-robust speech recognition techniques Written by leading researchers who have been actively working on the subject matter in both industrial and academic organizations for many years

Speech Recognition with Probabilistic Transcriptions and End-to-end Systems Using Deep Learning

5th International Conference on Communication and Electronics Systems (ICCES 2020) is being organized on 10-12, June 2020. ICCES will provide an outstanding international forum for sharing knowledge and results in all fields of Engineering and Technology. ICCES provides quality key experts who provide an opportunity in bringing up innovative ideas. Recent updates in the field of technology will be a platform for the upcoming researchers. The conference will be complete, concise, clear, and cohesive in terms of research related to Communication and Electronics systems.

Intelligent Speech Signal Processing

Traditionally, scientific fields have defined boundaries, and scientists work on research problems within those boundaries. However, from time to time those boundaries get shifted or blurred to evolve new fields. For instance, the original goal of computer vision was to understand a single image of a scene, by identifying objects, their structure, and spatial arrangements. This has been referred to as image understanding. Recently, computer vision has gradually been making the transition away from understanding single images to analyzing image sequences, or video. Understanding deals with understanding of video understanding. Sequences, e.g., recognition of gestures, activities, facial expressions, etc. The main shift in the classic paradigm has been from the recognition of static objects in the scene to motion-based recognition of actions and events. Video understanding has overlapping research problems with other fields, therefore blurring the fixed boundaries. Computer graphics, image processing, and video databases have obvious overlap with computer vision. The main goal of computer graphics is to generate and animate realistic looking images, and videos. Researchers in computer graphics are increasingly employing techniques from computer vision to generate the synthetic imagery. A good example of this is image-based rendering and modeling techniques, in which geometry, appearance, and lighting is derived from real images using computer vision techniques. Here the shift is from synthesis to analysis followed by synthesis. Image processing has always overlapped with computer vision because they both inherently work directly with images.

Connectionist Speech Recognition

This book provides a comprehensive overview of the recent advancement in the field of automatic speech recognition with a focus on deep learning models including deep neural networks and many of their variants. This is the first automatic speech recognition book dedicated to the deep learning approach. In addition to the rigorous mathematical treatment of the subject, the book also presents insights and theoretical foundation of a series of highly successful deep learning models.
**Deep Learning for NLP and Speech Recognition**

Connectionist Speech Recognition: A Hybrid Approach describes the theory and implementation of a method to incorporate neural network approaches into state of the art continuous speech recognition systems based on hidden Markov models (HMMs) to improve their performance. In this framework, neural networks (and in particular, multilayer perceptrons or MLPs) have been restricted to well-defined subtasks of the whole system, i.e. HMM emission probability estimation and feature extraction. The book describes a successful five-year international collaboration between the authors. The lessons learned form a case study that demonstrates how hybrid systems can be developed to combine neural networks with more traditional statistical approaches. The book illustrates both the advantages and limitations of neural networks in the framework of a statistical systems. Using standard databases and comparison with some conventional approaches, it is shown that MLP probability estimation can improve recognition performance. Other approaches are discussed, though there is no such unequivocal experimental result for these methods. Connectionist Speech Recognition is of use to anyone intending to use neural networks for speech recognition or within the framework provided by an existing successful statistical approach. This includes research and development groups working in the field of speech recognition, both with standard and neural network approaches, as well as other pattern recognition and/or neural network researchers. The book is also suitable as a text for advanced courses on neural networks or speech processing.

**The Oxford Handbook of Voice Perception**

This book covers the state-of-the-art in deep neural-network-based methods for noise robustness in distant speech recognition applications. It provides insights and detailed descriptions of some of the new concepts and key technologies in the field, including novel architectures for speech enhancement, microphone arrays, robust features, acoustic model adaptation, training data augmentation, and training criteria. The contributed chapters also include descriptions of real-world applications, benchmark tools and datasets widely used in the field. This book is intended for researchers and practitioners working in the field of speech processing and recognition who are interested in the latest deep learning techniques for noise robustness. It will also be of interest to graduate students in electrical engineering or computer science, who will find it a useful guide to this field of research.

**Recent Advances in Computational Intelligence in Defense and Security**

Speech dynamics refer to the temporal characteristics in all stages of the human speech communication process. This speech “chain” starts with the formation of a linguistic message in a speaker’s brain and ends with the arrival of the message in a listener's brain. Given the intricacy of the dynamic speech process and its fundamental importance in human communication, this monograph is intended to provide a comprehensive material on mathematical models of speech dynamics and to address the following issues: How do we make sense of the complex speech process in terms of its functional role of speech communication? How do we quantify the special role of speech timing? How do the dynamics relate to the variability of speech that has often been said to seriously hamper automatic speech recognition? How do we put the dynamic process of speech into a quantitative form to enable detailed analyses? And finally, how can we incorporate the knowledge of speech dynamics into computerized speech analysis and recognition algorithms? The answers to all these questions require building and applying computational models for the dynamic speech process. What are the compelling reasons for carrying out dynamic speech modeling? We provide the answer in two related aspects. First, scientific inquiry into the human speech code has been relentlessly pursued for several decades. As an essential carrier of human intelligence and knowledge, speech is the most natural form of human communication. Embedded in the speech code are linguistic (as well as para-linguistic) messages, which are conveyed through four levels of the speech chain. Underlying the robust encoding and transmission of the linguistic messages are the speech dynamics at all the four levels. Mathematical modeling of speech dynamics provides an effective tool in the scientific methods of studying the speech chain. Such scientific studies help understand why humans speak as they do and how humans exploit redundancy and variability by way of multitiered dynamic processes to enhance the efficiency and effectiveness of human speech communication. Second, advancement of human language technology, especially that in automatic recognition of natural-style human speech is also expected to benefit from comprehensive computational modeling of speech dynamics. The limitations of current speech recognition technology are serious and are well known. A commonly acknowledged
and frequently discussed weakness of the statistical model underlying current speech recognition technology is the lack of adequate dynamic modeling schemes to provide correlation structure across the temporal speech observation sequence. Unfortunately, due to a variety of reasons, the majority of current research activities in this area favor only incremental modifications and improvements to the existing HMM-based state-of-the-art. For example, while the dynamic and correlation modeling is known to be an important topic, most of the systems nevertheless employ only an ultra-weak form of speech dynamics; e.g., differential or delta parameters. Strong-form dynamic speech modeling, which is the focus of this monograph, may serve as an ultimate solution to this problem. After the introduction chapter, the main body of this monograph consists of four chapters. They cover various aspects of theory, algorithms, and applications of dynamic speech models, and provide a comprehensive survey of the research work in this area spanning over past 20 years. This monograph is intended as advanced materials of speech and signal processing for graduate-level teaching, for professionals and engineering practitioners, as well as for seasoned researchers and engineers specialized in speech processing.

**Deep Learning for Natural Language Processing**

The ICTer2019 Conference will focus on important problems and potential solutions in the areas of ICT including but not limited to the following: Natural language processing and text analytics, Cyber physical systems, Big data analytics, Image processing, visualization, modeling and simulation, Autonomic computing, Cloud computing, virtualization, and middleware, Cognitive reasoning, Smart cities and environments, Software reliability and large scale distribution, Parallelisation, HPC and GPU computing, E Science in social science, health, climate, and education, Context aware computing, Cybersecurity and application security, Internet of Things, Mobile and sensor networks, Swarm intelligence, evolutionary computing, and neural networks. ICT for Development, E Learning, ICT Innovation, Blockchain technology, and Cryptocurrencies.

**Dysarthric Speech Recognition and Offline Handwriting Recognition Using Deep Neural Networks**

The textbook covers the main aspects of Edge Computing, from a thorough look at the technology to the standards and industry associations working in the field. The book is conceived as a textbook for graduate students but also functions as a working guide for developers, engineers, and researchers. The book aims not only at providing a comprehensive technology and standard reference overview for students, but also useful research insights and practical exercises for edge software developers and investigators in the area (and for students looking to apply their skills). A particular emphasis is given to Multi-access Edge Computing (MEC) as defined in European Telecommunications Standards Institute (ETSI), in relationship with other standard organizations like 3GPP, thus in alignment with the recent industry efforts to produce harmonized standards for edge computing leveraging both ETSI ISG MEC and 3GPP specifications. Practical examples of Edge Computing implementation from industry groups, associations, companies and edge developers, complete the book and make it useful for students entering the field. The book includes exercises, examples, and quizzes throughout.

**Robust Automatic Speech Recognition**

This textbook explains Deep Learning Architecture, with applications to various NLP Tasks, including Document Classification, Machine Translation, Language Modeling, and Speech Recognition. With the widespread adoption of deep learning, natural language processing (NLP), and speech applications in many areas (including Finance, Healthcare, and Government) there is a growing need for one comprehensive resource that maps deep learning techniques to NLP and speech and provides insights into using the tools and libraries for real-world applications. Deep Learning for NLP and Speech Recognition explains recent deep learning methods applicable to NLP and speech, provides state-of-the-art approaches, and offers real-world case studies with code to provide hands-on experience. Many books focus on deep learning theory or deep learning for NLP-specific tasks while others are cookbooks for tools and libraries, but the constant flux of new algorithms, tools, frameworks, and libraries in a rapidly evolving landscape means that there are few available texts that offer the material in this book. The book is organized into three parts, aligning to different groups of readers and their expertise. The three parts are: Machine Learning, NLP, and Speech Introduction. The first part has three chapters that introduce readers to the fields of NLP, speech recognition, deep learning.
and machine learning with basic theory and hands-on case studies using Python-based tools and libraries. Deep Learning Basics The five chapters in the second part introduce deep learning and various topics that are crucial for speech and text processing, including word embeddings, convolutional neural networks, recurrent neural networks and speech recognition basics. Theory, practical tips, state-of-the-art methods, experimentations and analysis in using the methods discussed in theory on real-world tasks. Advanced Deep Learning Techniques for Text and Speech The third part has five chapters that discuss the latest and cutting-edge research in the areas of deep learning that intersect with NLP and speech. Topics including attention mechanisms, memory augmented networks, transfer learning, multi-task learning, domain adaptation, reinforcement learning, and end-to-end deep learning for speech recognition are covered using case studies.

Deep Learning for NLP and Speech Recognition

Speech perception has been the focus of innumerable studies over the past decades. While our abilities to recognize individuals by their voice state plays a central role in our everyday social interactions, limited scientific attention has been devoted to the perceptual and cerebral mechanisms underlying nonverbal information processing in voices. The Oxford Handbook of Voice Perception takes a comprehensive look at this emerging field and presents a selection of current research in voice perception. The forty chapters summarise the most exciting research from across several disciplines covering acoustical, clinical, evolutionary, cognitive, and computational perspectives. In particular, this handbook offers an invaluable window into the development and evolution of the 'vocal brain', and considers in detail the voice processing abilities of non-human animals or human infants. By providing a full and unique perspective on the recent developments in this burgeoning area of study, this text is an important and interdisciplinary resource for students, researchers, and scientific journalists interested in voice perception.

2019 19th International Conference on Advances in ICT for Emerging Regions (ICTer)

Copyright code : f1f1a245dcce16a587aaaf33399cbb08