Text Document Character Segmentation Matlab Source Code

This is the first comprehensive text on Optical Character Recognition for Indic scripts. It covers many topics and describes OCR systems for eight different scripts—Bangla, Devanagari, Gurmukhi, Gujarti, Kannada, Malayalam, Tamil and Urdu.

Optical character recognition and document image analysis have become very important areas with a fast growing number of researchers in the field. This comprehensive handbook with contributions by eminent experts, presents both the theoretical and practical aspects at an introductory level wherever possible. Contents:Pattern Classification Techniques Based on Function Approximation (U Kressel & J Schürmann)Combination of Multiple Classifier Decisions for Optical Character Recognition (L Lam et al.)Segmentation-Based Cursive Handwriting Recognition (M Shridhar & F Kimura)Handwritten Word Recognition Using Hidden Markov Models (A Kundu)Techniques for Improving OCR Results (A Dengel et al.)Multilingual Document Recognition (A L Spitz)Arabic Character Recognition (A Amin)Interpretation of Engineering Drawings (K Tombre & D Dori)Automatic Reading of Music Notation (D Bainbridge & N Carter)Algorithms for Automatic Signature Verification (G Dimauro et al.)Automatic Reading of Braille Documents (A Antonacopoulos)Information Retrieval and OCR (K Taghva et al.)Benchmarking DIA Systems (T A Nartker et al.)and other papers Readership: Computer scientists and engineers. keywords:

It was estimated that 80% of the information received by human is visual. Image processing is evolving fast and continually. During the past 10 years, there has been a significant research increase in image segmentation. To study a specific object in an image, its boundary can be highlighted by an image segmentation procedure. The objective of the image segmentation is to simplify the representation of pictures into meaningful information by partitioning into image regions. Image segmentation is a technique to locate certain objects or boundaries within an image. There are many algorithms and techniques have been developed to solve image segmentation problems, the research topics in this book such as level set, active contour, AR time series image modeling, Support Vector Machines, Pixon based image segmentations, region similarity metric based technique, statistical ANN and JSEG algorithm were written in details. This book brings together many different aspects of the current research on several fields associated to digital image segmentation. Four parts allowed gathering the 27 chapters around the following topics: Survey of Image Segmentation Algorithms, Image Segmentation methods, Image Segmentation Applications and Hardware Implementation. The readers will find the contents in this book enjoyable and get many helpful ideas and overviews on their own study.

The proceedings includes cutting-edge research articles from the Fourth International Conference on Signal and Image Processing (ICSIP), which is organised by Dr. N.G.P. Institute of Technology, Kalapatti, Coimbatore. The Conference provides academia and industry to discuss and present the latest technological advances and research results in the fields of theoretical, experimental, and application of signal, image and video processing. The book provides latest and most informative content from engineers and scientists in signal, image and video processing from around the world, which will benefit the future research community to work in a more cohesive and collaborative way.

Spread in 133 articles divided in 20 sections the present treatises broadly discusses: Part 1: Image Processing Part 2: Radar and Satellite Image Processing Part 3: Image Filtering Part 4: Content Based Image Retrieval Part 5: Color Image Processing and Video Processing Part 6: Medical Image Processing Part 7: Biometric Part 8: Network Part 9: Mobile Computing Part 10: Pattern Recognition Part 11: Pattern Classification Part 12: Genetic Algorithm Part 13: Data Warehousing and Mining Part 14: Embedded System Part 15: Wavelet Part 16: Signal Processing Part 17: Neural Network Part 18: Nanotechnology and Quantum Computing Part 19: Image Analysis Part 20: Human Computer Interaction

This book constitutes the throughly refereed post-proceedings of the 4th International Conference on Intelligent Data Engineering and Automated Learning, IDEAL 2003, held in Hong Kong, China in March 2003. The 164 revised papers presented were carefully reviewed and selected from 321 submissions; for inclusion in this post-proceedings another round of revision was imposed. The papers are organized in topical sections an agents, automated learning, bioinformatics, data mining, multimedia information, and financial engineering.

This book gathers selected high-quality research papers presented at International Conference on Advanced Computing and Intelligent Technologies (ICACIT 2021) held at NCR New Delhi, India, during March 2021, 2021, jointly organized by Galgotias University, India, and Department of Information Engineering and Mathematics Universita Di Siena, Italy. It discusses emerging topics pertaining to advanced computing, intelligent technologies, and networks including AI and machine learning, data mining, big data analytics, high-performance computing network performance analysis, Internet of things networks, wireless sensor networks, and others. The book offers a valuable asset for researchers from both academia and industries involved in advanced studies.

IAENG Transactions on Engineering Technologies contains forty-nine revised and extended research articles, written by prominent researchers participating in the conference. Topics covered include circuits, engineering mathematics, control theory, communications systems, systems engineering, manufacture engineering, computational biology, chemical engineering, and industrial applications. This book offers the state of art of tremendous advances in engineering technologies and physical science and applications, and also serves as an excellent source of reference for researchers and graduate students working with/on engineering technologies and physical science and applications.

This book constitutes the refereed proceedings of the 6th FIP WG 2.2 International Conference, TCS 2010, held as a part of the 21th World Computer Congress, WCC 2010, in Brisbane, Australia, in September 2010. The 23 revised full papers presented, together with 4 invited talks, were carefully reviewed and selected from 39 submissions. TCS 2010 deals with topics focused at but not limited to algorithms, complexity, models of computation, logic, semantics, specification and verification, power-awareness issues in wireless networks, data mining, knowledge discovery, multiprocessor issues as well as AI issues.

This book brings all the major and frontier topics in the field of document analysis together into a single volume, creating a unique reference source that will be invaluable to a large audience of researchers, lecturers and students working in this field. With chapters written by some of the most distinguished researchers active in this field, this book addresses recent advances in digital document processing research and development.

In this book, differential evolution and its modified variants are applied to the clustering of data and images. Metaheuristics have emerged as potential algorithms for dealing with complex optimization problems, which are otherwise difficult to solve using traditional methods. In this regard, differential evolution is considered to be a highly promising technique for optimization and is being used to solve various real-time problems. The book studies the algorithms in detail, tests them on a range of test images, and carefully analyzes their performance. Accordingly, it offers a valuable reference guide for all researchers, students and practitioners working in the fields of artificial intelligence, optimization and data analytics.

This book constitutes the refereed proceedings of the 5th International Conference on Image Analysis and Recognition, ICIAR 2008, held in Póvoa do Varzim, Portugal, in June 2008. The 110 revised full papers presented together with 2 invited papers were carefully reviewed and selected from 226 submissions. The papers are organized in topical sections on image restoration and enhancement, image and video segmentation, non-linear image processing, image and video coding and encryption, indexing and retrieval, computer vision, feature extraction and classification, shape representation and matching, object recognition, character recognition, texture and motion analysis, tracking, biomedical image analysis, biometrics, face recognition, and a special session on recent advances in multimodal biometric systems and applications.

This book presents a remarkable collection of chapters that cover a wide range of topics in the areas of information and communication technologies and their real-world applications. It gathers the Proceedings of the Future of Information and Communication Conference 2019 (FICC 2019), held in San Francisco, USA from March 14 to 15, 2019. The conference attracted a total of 462 submissions from pioneering researchers, scientists, industrial engineers, and students from all around the world. Following a double-blind peer review process, 160 submissions (including 15 poster papers) were ultimately selected for inclusion in these proceedings. The papers highlight relevant trends in, and the latest research on: Communication, Data Science, Ambient Intelligence, Networking, Computing, Security, and the Internet of Things. Further, they address all aspects of Information Science and communication technologies, from classical to intelligent, and both the theory and applications of the latest technologies and methodologies. Gathering chapters that discuss state-of-the-art intelligent methods and techniques for solving real-world problems, along with future research directions, the book represents both an interesting read and a valuable asset.

The First Book to Describe the Technical and Practical Elements of Chemical Text Mining Explores the development of chemical structure extraction capabilities and how to incorporate these technologies in daily research work For scientific researchers, finding too much information on a subject, not finding enough information, or not being able to access full text documents often costs them time, money, and quality. Addressing these concerns, Chemical Information Mining: Facilitating Literature-Based Discovery presents strategic ideas for properly selecting and successfully using the best text mining tools for scientific research. Links chemical and biological entities at the heart of life science research The book focuses on information extraction issues, highlights available solutions, and underscores the value of these solutions to academic and commercial scientists. After introducing the drivers behind chemical text mining, it discusses chemical semantics. The contributors describe the tools that identify and convert chemical names and images to structure-searchable information. They also explain natural language processing, name entity recognition concepts, and semantic web technologies. Following a section on current trends in the field, the book looks at where information mining approaches fit into the research needs within the life sciences. Shaping the future of scientific information and knowledge management By building knowledge and competency in the growing area of literature-based discovery, this book shows how text mining of the chemical literature can increase drug discovery opportunities and enhance life science research.

This book discusses email spam detection and its challenges such as text classification and categorization. The book proposes an efficient spam detection technique that is a combination of Character Segmentation and Recognition and Classification (CSRC). The author describes how this can detect whether an email (text and image based) is a spam mail or not. The book presents four solutions: first, to extract the text character from the image by segmentation process which includes a combination of Discrete Wavelet Transform (DWT) and skew detection. Second, text characters are via text recognition and visual feature extraction approach which relies on contour analysis with improved Local Binary Pattern (LBP). Third, extracted text features are classified using improvised K-Nearest Neighbor search (KNN) and Support Vector Machine (SVM). Fourth, the performance of the proposed method is validated by the measure of metric named as sensitivity, specificity, precision, recall, F-measure, accuracy, error rate and correct rate. Presents solutions to email spam detection and discusses its challenges such as text classification and categorization; Analyzes the proposed techniques' performance using precision, F-measure, recall and accuracy; Evaluates the limitations of the proposed research thereby recommending future research.

Innovative Data Communication Technologies and ApplicationICIDCA 2019Springer Nature

This book presents emerging concepts in data mining, big data analysis, communication, and networking technologies, and discusses the state-of-the-art in data engineering practices to tackle massive data distributions in smart networked environments. It also provides insights into potential data distribution challenges in ubiquitous data-driven networks, highlighting research on the theoretical and systematic framework for analyzing, testing and designing intelligent data analysis models for evolving communication frameworks. Further, the book showcases the latest developments in wireless sensor networks, cloud computing, mobile network, autonomous systems, cryptography, automation, and other communication and networking technologies. In addition, it addresses data security, privacy and trust, wireless networks, data classification, data prediction, performance analysis, data validation and verification models, machine learning, sentiment analysis, and various data analysis techniques.

This book constitutes the thoroughly refereed post-workshop-proceedings of the 4th International Workshop on Camera-Based Document Analysis and Recognition, CBDAR 2011, held in Beijing, China, in September 2011. The 13 revised full papers presented were carefully selected during a second round of reviewing and improvement from numerous original submissions. Intended to give a snapshot of the state-of-the-art research in the field of camera based document analysis and recognition, the papers are organized in topical sections on text detection and recognition in scene images, camera-based systems, and datasets and evaluation.

This book comprises the refereed proceedings of the International Conference, AIM/CCPE 2012, held in Bangalore, India, in April 2012. The papers presented were carefully reviewed and selected from numerous submissions and focus on the various aspects of research and development activities in computer science, information technology, computational engineering, mobile communication, control and instrumentation, communication system, power electronics and power engineering.

The international multi-topic conference IMTIC 2008 was held in Pakistan during April 11–12, 2008. It was a joint venture between Mehran University, Jamshoro, Sindh and Aalborg University, Esbjerg, Denmark. Apart from the two-day main event, two workshops were also held: the Workshop on Creating Social Semantic Web 2.0 Information Spaces and the Workshop on Wireless Sensor Networks. Two hundred participants registered for the main conference from 24 countries and 43 papers were presented; the two workshops had overwhelming support and over 400 delegates registered. IMTIC 2008

served as a platform for international scientists and the engineering community in general, and in particular for local scientists and the engineering c- munity to share and cooperate in various fields of interest. The topics presented had a reasonable balance between theory and practice in multidisciplinary topics. The c- ference also had excellent topics covered by the keynote speeches keeping in view the local requirements, which served as a stimulus for students as well as experienced participants. The Program Committee and various other committees were experts in their areas and each paper went through a double-blind peer review process. The c- ference received 135 submissions of which only 46 papers were selected for presen- tion: an acceptance rate of 34%.

This book constitutes the thoroughly refereed post-conference proceedings of the 5th International Conference on Mobile Computing, Applications, and Services (MobiCASE 2013) held in Paris, France, in November 2013. The 13 full, 5 short and 9 poster papers were carefully reviewed and selected from 64 submissions, and are presented together with 3 papers from the Workshop on Near Field Communication for Mobile Applications (NFS). The conference papers are covering mobile applications development, mobile social networking, novel user experience and interfaces, mobile services and platforms such as Android, iOS, BlackBerry OS, Windows phone, Bada, mobile software engineering and mobile Web, mobile payments and M2M infrastructure, mobile services such as novel hardware add-ons, energy aware services or tools, NFC-based services, authentication services.

Security and authentication issues are surging to the forefront of the research realm in global society. As technology continues to evolve, individuals are finding it easier to infiltrate various forums and facilities where they can illegally obtain information and access. By implementing biometric authentications to these forums, users are able to prevent attacks on their privacy and security. Biometrics: Concepts, Methodologies, Tools, and Applications is a multi-volume publication highlighting critical topics related to access control, user identification, and surveillance technologies. Featuring emergent research on the issues and challenges in security and privacy, various forms of user authentication, biometric applications to image processing and computer vision, and security applications within the field, this publication is an ideal reference source for researchers, engineers, technology developers, students, and security specialists.

This book constitutes the thoroughly refereed proceedings of the Eleventh International Symposium on Natural Language Processing (SNLP-2016), held in Phranakhon Si Ayutthaya, Thailand on February 10–12, 2016. The SNLP promotes research in natural language processing and related fields, and provides a unique opportunity for researchers, professionals and practitioners to discuss various current and advanced issues of interest in NLP. The 2016 symposium was expanded to include the First Workshop in Intelligent Informatics and Smart Technology. Of the 66 high-quality papers accepted, this book presents twelve from the Symposium on Natural Language Processing track and ten from the Workshop in Intelligent Informatics and Smart Technology track (SSAI: Special Session on Artificial Intelligence).

The two volume set LNCS 8047 and 8048 constitutes the refereed proceedings of the 15th International Conference on Computer Analysis of Images and Patterns, CAIP 2013, held in York, UK, in August 2013. The 142 papers presented were carefully reviewed and selected from 243 submissions. The scope of the conference spans the following areas: 3D TV, biometrics, color and texture, document analysis, graph-based methods, image and video indexing and database retrieval, image and video processing, image-based modeling, kernel methods, medical imaging, mobile multimedia, model-based vision approaches, motion analysis, natural computation for digital imagery, segmentation and grouping, and shape representation and analysis.

This book includes selected papers from the 4th International Conference on Computational Vision and Bio Inspired Computing (ICCVBIC 2020), held in Coimbatore, India, from November 19 to 20, 2020. This proceedings book presents state-of-the-art research innovations in computational vision and bio-inspired techniques. The book reveals the theoretical and practical aspects of bio-inspired computing techniques, like machine learning, sensor-based models, evolutionary optimization and big data modeling and management that make use of effectual computing processes in the bio-inspired systems. As such it contributes to the novel research that focuses on developing bio-inspired computing solutions for various domains, such as humancomputer interaction, image processing, sensor-based single processing, recommender systems and facial recognition, which play an indispensable part in smart agriculture, smart city, biomedical and business intelligence applications.

This book presents a systematic introduction to the latest developments in video text detection. Opening with a discussion of the underlying theory and a brief history of video text detection, the text proceeds to cover pre-processing and post-processing techniques, character segmentation and recognition, identification of non-English scripts, techniques for multi-modal analysis and performance evaluation. The detection of text from both natural video scenes and artificially inserted captions is examined. Various applications of the technology are also reviewed, from license plate recognition and road navigation assistance, to sports analysis and video advertising systems. Features: explains the fundamental theory in a succinct manner, supplemented with references for further reading; highlights practical techniques to help the reader understand and develop their own video text detection systems and applications; serves as an easy-to-navigate reference, presenting the material in self-contained chapters.

This conference proceedings summarizes invited publications from the two IDES (Institute of Doctors Engineers and Scientists) International conferences, both held in Bangalore/ India.

Affordable and mainstream manipulation of multimedia data types will lead to tremendous growth in imaging and multimedia data in general computing environments. Multimedia and imaging applications can now provide benefits to common business applications by integrating voice, sound, images, animation and digitized video. Ultimately, it will be possible to convert all information that is currently stored on paper, video and film into a digitized environment. This will allow users to organize, search and route multimedia objects over local and wide area networks in real time. The authors' introductory level presentation of this new class of data types supplies the database technology required for effective manipulation and storage. Multimedia and database experts, Khoshafian and Baker aptly illustrate the ability of multimedia database systems to concurrently share, access, and query large collections of multimedia information. They introduce the elemental concepts of object and relational databases and then apply them to multimedia and imaging databases. Fundamental database topics discussed include querying, transaction support, recovery, security, and storage. This book provides information essential to the incorporation of multimedia databases that will improve the quantity and quality of information manipulated by computer users in many areas including medicine, computer aided design, and information retrieval systems.

This book presents the latest research in hybrid intelligent systems. It includes 57 carefully selected papers from the 16th International Conference on Hybrid Intelligent Systems (HIS 2016) and the 8th World Congress on Nature and Biologically Inspired Computing (NaBIC 2016), held on November 21–23, 2016 in Marrakech, Morocco. HIS - NaBIC 2016 was jointly organized by the Machine Intelligence Research Labs (MIR Labs), USA; Hassan 1st University, Settat, Morocco and University of Sfax, Tunisia. Hybridization of intelligent systems is a promising research field in modern artificial/computational intelligence

and is concerned with the development of the next generation of intelligent systems. The conference's main aim is to inspire further exploration of the intriguing potential of hybrid intelligent systems and bio-inspired computing. As such, the book is a valuable resource for practicing engineers /scientists and researchers working in the field of computational intelligence and artificial intelligence.

The book is about all aspects of computing, communication, general sciences and educational research covered at the Second International Conference on Computer & Communication Technologies held during 24-26 July 2015 at Hyderabad. It hosted by CMR Technical Campus in association with Division – V (Education & Research) CSI, India. After a rigorous review only quality papers are selected and included in this book. The entire book is divided into three volumes. Three volumes cover a variety of topics which include medical imaging, networks, data mining, intelligent computing, software design, image processing, mobile computing, digital signals and speech processing, video surveillance and processing, web mining, wireless sensor networks, circuit analysis, fuzzy systems, antenna and communication systems, biomedical signal processing and applications, cloud computing, embedded systems applications and cyber security and digital forensic. The readers of these volumes will be highly benefited from the technical contents of the topics.

Easy access to digital information in every form is something which has become indispensable given our ever-increasing reliance on digital technology. But such access would not be possible without the reliable and effective infrastructure which has led to the large-scale development of web technologies. This book presents the 27 papers delivered at the 6th International Conference on Applications of Digital Information and Web Technologies (ICADIWT), held in February 2015, at the University of Macau, Macau. The book is divided into seven sections: Internet communication, human-computer interaction, adaptive web applications, data communication, cloud computing, systems engineering, and data mining. Since each paper is a survey contributed by different experts from very many countries, this book can be seen as a collection of the current research trends in the field and hence it will be of interest to all those whose work involves digital information and web technology.

This book constitutes the proceedings of the First International Conference on Mining Intelligence and Knowledge Exploration, MIKE 2013, held in Tamil Nadu, India on December 2013. The 82 papers presented were carefully reviewed and selected from 334 submissions. The papers cover the topics such as feature selection, classification, clustering, image processing, network security, speech processing, machine learning, information retrieval, recommender systems, natural language processing, language, cognition and computation and other certain problems in dynamical systems.

This two-volume set (CCIS 1075 and CCIS 1076) constitutes the refereed proceedings of the Third International Conference on Advanced Informatics for Computing Research, ICAICR 2019, held in Shimla, India, in June 2019. The 78 revised full papers presented were carefully reviewed and selected from 382 submissions. The papers are organized in topical sections on computing methodologies; hardware; information systems; networks; software and its engineering.

This book constitutes the refereed proceedings of the International Conference on Information Systems for Indian Languages, ICISIL 2011, held in Patiala, India, in March 2011. The 63 revised papers presented were carefully reviewed and selected from 126 paper submissions (full papers as well as poster papers) and 25 demo submissions. The papers address all current aspects on localization, e-governance, Web content accessibility, search engine and information retrieval systems, online and offline OCR, handwriting recognition, machine translation and transliteration, and text-to-speech and speech recognition - all with a particular focus on Indic scripts and languages.

The rapid proliferation of Multimedia and Network Information Systems is one of the key features of our times. What is also important is that the pace of change is ever increasing. University projects of today will form the core of consumer products of tomorrow. Therefore, it is very important to have a broad view of the recent scientific investigation in that area. This was the primary reason for gathering this collection of carefully selected and hopefully representative research projects, found solutions, and finally applications. They are the achievements of scientific teams from several countries. The contents of the monograph has been divided into four parts: 1) Multimedia Information Technology, 2) Information Systems Specification, 3) Information Systems Applications, 4) Web Systems and Network Technologies. The book presents up to date research from the diverse fields of multimedia and Internet data processing.