ICMV

CONFERENCE PROGRAM

2020 The 13th International Conference on Machine Vision

NOV. 2-6, 2020 | ROME, ITALY

GMT+1

CONTENTS

PART ONE
General Conference Information

PART TWO
Keynote & Plenary Speakers,
Special Session Chairs

PART THREE
Technical Program

PART FOUR
Presentation Abstracts

WELCOME ADDRESS

Welcome to 2020 The 13th International Conference on Machine Vision (ICMV 2020), to be held during November 2-6, 2020, in virtual way via ZOOM tool, due to current global epidemic situation.

Here on behalf of the organizing committee, including our international advisory chair, general chairs, chapter chairs, local chair, program chairs, technical committees, we would like to convey our appreciation of your participations despite this unprecedented time.

This year the conference will be held in full web streaming. We are creating a virtual environment where the attendees can present the papers and can participate in all the sessions that will be organized in a full immersive experience

We are confident that you will find the technical program of ICMV 2020 very stimulating and inspiring. We warmly invite all of you to join us in interacting with the keynotes and plenary speakers, and more importantly, interacting with all of your peer attendees. A technical program like this would not have been possible without the hard work and devotion of many of our organizing committee members and volunteers. We would like to especially thank the Special Session Chairs for their exceptional work. We would also like to express our sincere thanks to all the TPC members and reviewers for their help in the paper review process.

Finally, we wish to thank all the authors and attendees for participating in the conference. We hope you will have a fruitful and memorable experience at ICMV 2020.

ICMV Conference Organizing Group

/01

PART I

General Conference Information

COMMITTEES

International Advisory Chair
Antanas Verikas, Halmstad University, Sweden

General Chairs
Wolfgang Osten, University of Stuttgart, Germany
Dmitry Nikolaev, Institute for Information Transmission Problems, Russia

Chapter Chairs
Alexander Bernstein, Skolkovo Institute of Science and Technology (Skoltech), Russia Federation
Petia Radeva, University of Barcelona, Spain
Dorra Sellami, Sfax University, Tunisia
Sei-ichiro Kamata, Waseda University, Japan

Local Chair
Alessia Cedola, The Sapienza University of Rome, Rome unit Nanotec CNR, Italy

Program Chairs
Johan Debayle, Ecole Nationale Supérieure des Mines, France
Vladimir Arlazarov, Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Russia
Pino Caballero-Gil, University of La Laguna, Spain
Brian W. Surgenor, Queen's University, Canada
Jianhong Zhou, University of Electronic Science and Technology of China, China

COMMITTEES

Technical Committees

A. I. Chulichkov
Annette Stahl
Aristidis Likas
Asadchikov V.
Ashish Khare
Assia Kourgli
Baihua Li
Batouche Mohamed
Dmitry V. Polevoy
Egor Bondarev
Egor Ershov
Elena Battini Sonmez
Evaldas Vaičiukynas
Evgeni Magid
Evgeny Burnaev
Evgeny Shvets
Gernot Stübl
Haluk Küçük
Hareesha K. S
Henryk Palus
James Hahn
Jean-Louis Dillenseger
Jianjun Yi
Jose M. Massa
Jun-Dong Cho
Junqiao Zhao
Kishor Bhurchandi
Liang Gao
M. Iqbal b. Saripan
Manshan Zhou
Manuel F. González Penedo
Marcos Orgega
Maya Dimitrova
Mehmet Cunkas
Mohamed Elhelw
Mohamed Rizon
Mounîm A. El-Yacoubi
Mourad Zaied
M-Tahar Kechadi
Oleg Slavin
Pabitra Mitra
Panagiotis Barmpoutis
Pınar Akdemir Özışık
Qieshi Zhang
Qujiang Lei
Reyer Zwiggelaar
Ridha Ejbali
Seokwon Yeom
Sergey Gladilin
Sergii Mashtalir
Seridi Hamid
Szidónia Lefkovits
Theodore Pachidis
Thomas B. Moeslund
Udo Birk
Viacheslav Voronin
Victoria A. Sablina
Vladimir A. Fursov

COMMITTEES

Technical Committees

Vladmimir, Khryashchev
Wafa' AlSharafat
Walid Hariri

Wanda Benesova
Xose M. Pardo Lopez
Y. Benezeth

Ye Duan
Zhan Ma
Zineddine Kouahla

IMPORTANT TIPS

GMT+1

- All formal session arranged by Rome Time..
- Join the room 15 Mins before session start.
- Connect with Stable WiFi or wired net work.

Audio & Video

- Turn on your Audio & Video.
- Use headset to enhance the audio effect and avoid the speaker echo or howling.
- Stay in a quite place without noise.
- Join **TEST DAY on Nov. 1.**

Zoom ID

- Virtual Room A: **976 5819 4637**
- Virtual Room B: **680 4475 3785**
- Virtual Room C: **659 3075 7574**

ZOOM

- ZOOM download link: zoom.us
- Download link for Chinese authors: https://zoom.com.cn/download
- ZOOM Operation guideline via : http://icmv.org/zoom.html

Rename

- For authors, please rename with Session No.+ Paper ID+ Name, before you join the room. (eg. S1+V001+Alice LI)
- For KN/PS or SC, please rename with KN/PS/SC+ Name..

Presentation

- Send me a Video within 10/15 Mins as back-up.
- Stay online at least during KN&PS speech time and your own session for Q&A.
- English used only.
- Certificate & receipt will be emailed after event.

Program Overview

Virtual Room A: 976 5819 4637; Virtual Room B: 680 4475 3785
For Test arrangement please see EXCEL form!

MON., 2 Nov. 2020	TUE., 3 Nov. 2020	WED., 4 Nov. 2020	THU., 5 Nov. 2020		FRI., 6 Nov. 2020	
Virtual Room A	Virtual Room A	Virtual Room A	Virtual Room A	Virtual Room B	Virtual Room A	Virtual Room B
09:00-09:05 **Opening Remarks** **Wolfgang Osten, University of Stuttgart, Germany** Chair: Dmitry Nikolaev, Institute for Information Transmission Problems, Russia						
09:05-09:55 **Keynote I: Aaron Hertzmann** **Adobe Research, USA**	**09:30-10:10** **Plenary II: Sergei Gorlatch** **University of Muenster, Germany**	09:30-12:00 Special Session III Machine Vision for Autonomous Driven Cars under Harsh Environmental Conditions (10:30-10:40, Break Time)	9:30-11:15 Session 6 Artificial Intelligence and Intelligent Computing	9:30-11:30 Session 7 Software and Information Engineering	10:00-11:30 Session 12 Intelligent Recognition Technology and Algorithm	10:00-11:45 Session 13 Digital Image Processing and Application
09:55-10:45 **Keynote II: Zhi-Hua Zhou** **Nanjing University, China**	**10:10-10:50** **Plenary III: Alexander Bernstein** **Skolkovo Institute of Science and Technology, Russia**					
10:45-11:00 Break Time	10:50-11:05 Break Time					
11:00-11:40 **Plenary I: Wolfgang Osten, University of Stuttgart, Germany**	**11:05-11:45** **Plenary IV: Petia Radeva** **University of Barcelona, Spain**					
11:40-13:30 Lunch Break	11:45-13:30 Lunch Break	12:00-13:00 Lunch Break	11:30-13:00 Lunch Break		11:30-13:00 Lunch Break	
13:30-17:00 Special Session I Camera Based and Mobile Recognition (15:10-15:20,Break Time)	13:30-16:40 Special Session II Advanced Imaging and Tomography (15:10-15:20,Break Time)	13:00-14:50 Special Session IV Computer optics Journal	13:00-14:30 Session 8 Target Detection	13:00-14:45 Session 9 Artificial Intelligence and Data Engineering	13:00-14:30 Session 14 Digital Image Processing and Application	13:00-14:30 Session 15 Computer Vision and Visualization
		14:50-15:10 Break Time	14:30-15:00 Break Time	14:45-15:00 Break Time	14:30-15:00 Break Time	14:30-15:00 Break Time
		15:10-17:30 Track Color computer vision	15:00-16:30 Session 10 Image Analysis and Methods	15:00-17:00 Session 11 Computer and Information Security	15:00-16:30 Session 16 Computer Photography and Image Processing	15:00-16:30 Session 17 Intelligent Image Analysis & Key Technologies
					16:30-17:00 **Closing Ceremony**	

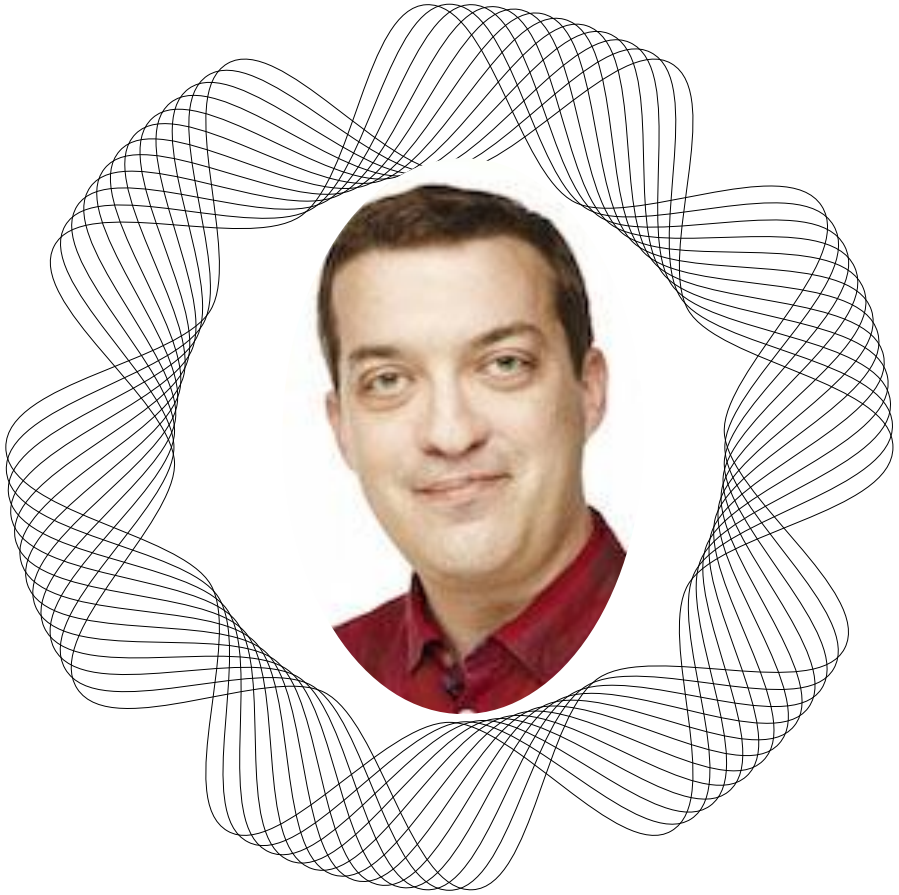
/02

PART II

Keynote & Plenary Speakers, and Special Session Chairs Information

KEYNOTE SPEAKER I

Monday, Nov. 2, 2020
Speech Time: 09:05-09:55
ZOOM ID: 976 5819 4637



Prof. Aaron Hertzmann

Adobe Research, USA
Fellow of the ACM, IEEE, H-index-59

Aaron Hertzmann received a BA in computer science and art & art history from Rice University in 1996, and a PhD in computer science from New York University in 2001. He was a Professor at University of Toronto for 10 years, and has also worked at Pixar Animation Studios, University of Washington, Microsoft Research, Mitsubishi Electric Research Lab, and Interval Research Corporation.

He is an Affiliate Professor at University of Washington, an ACM Fellow, an IEEE Fellow, and the Editor-in-Chief of Foundations and Trends in Computer Graphics and Vision.

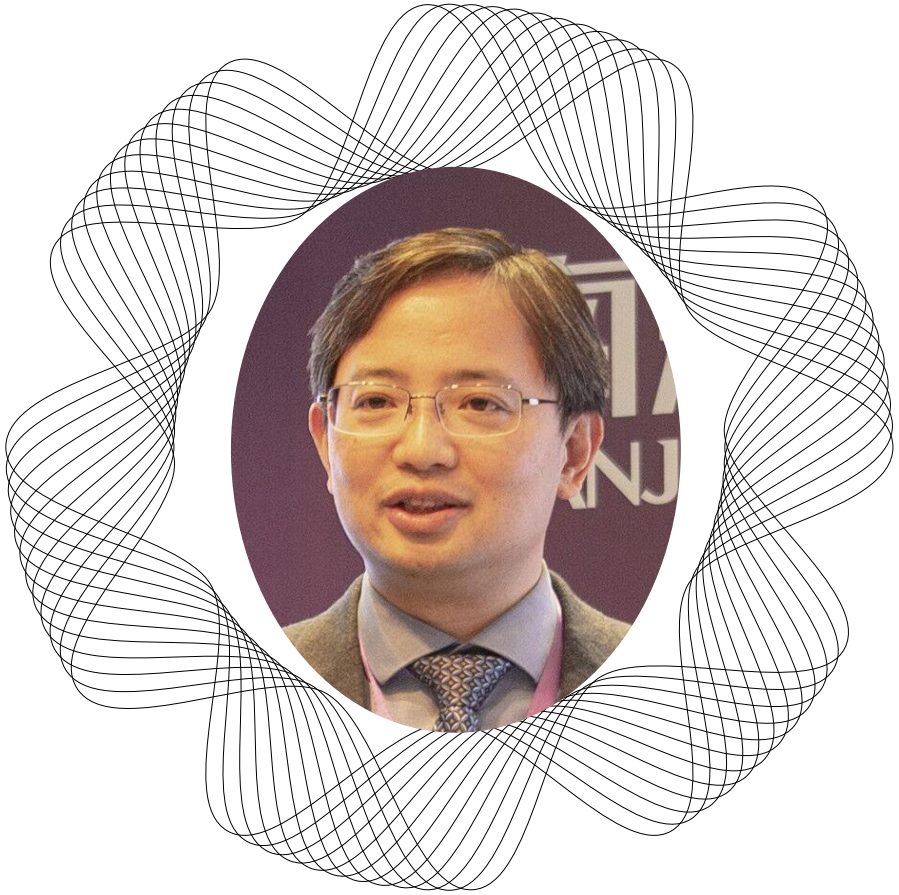
KEYNOTE SPEECH I

TITLE: Can Computers Create Art?

I will discuss whether computers, using Artificial Intelligence (AI), could create art. I cover the history of automation in art, examining the hype and reality of AI tools for art together with predictions about how they will be used. I will also discuss different scenarios for how an algorithm could be considered the author of an artwork, which, I argue, comes down to questions of why we create and appreciate artwork.

KEYNOTE SPEAKER II

Monday, Nov. 2, 2020
Speech Time: 09:55-10:45
ZOOM ID: 976 5819 4637



Prof. Zhi-Hua Zhou

Nanjing University, China

Fellow of the ACM, AAAI, AAAS, IEEE, IAPR, H-index 101

About Prof. Zhi-Hua Zhou: He is a Professor, Head of the Department of Computer Science and Technology, Dean of the School of Artificial Intelligence, and Founding Director of the LAMDA Group, Nanjing University. His main research interests are in artificial intelligence, machine learning and data mining. He authored the books "Ensemble Methods: Foundations and Algorithms (2012)", "Machine Learning (in Chinese, 2016)", etc., and published more than 200 papers in top-tier international journals/conferences. According to Google Scholar, his publications have received 50,000+ citations, with an H-index of 101. He also holds 24 patents and has rich experiences in industrial applications. He has received awards including the National Natural Science Award of China, the IEEE CS Edward J. McCluskey Technical Achievement Award, ACML Distinguished Contribution Award, etc. He serves as the Editor-in-Chief of Frontiers of Computer Science, Associate Editor-in-Chief of Science China Information Sciences, and Action/Associate Editor of Machine Learning, IEEE PAMI, ACM TKDD, etc. He founded ACML (Asian Conference on Machine Learning) and served as Chair for many premier conferences, such as Program Chair of IJCAI 2021 and AAAI 2019, General Chair of ICDM 2016, Senior Area Chair of NeurIPS, etc. He is a Fellow of the ACM, AAAI, AAAS, IEEE, IAPR.

KEYNOTE SPEECH II

TITLE: Leveraging Unlabeled Data: From Self-training to Abductive Learning

In many real AI application scenarios, such as some special machine vision tasks, it can be expensive or even infeasible to collect a huge amount of labeled training data, and therefore, leveraging unlabeled data is attracting much more attention. In this talk, we will briefly introduce the efforts of the machine learning community for leveraging unlabeled data, from self-training which generally suffers from noisy pseudo-labels, to semi-supervised learning which requires nice labeled data, active learning which requires human-in-loop, and the recent proposed abductive learning which offers a balanced way to exploit learning and knowledge reasoning.

PLENARY SPEAKER I

Monday, Nov. 2, 2020
Speech Time: 11:00-11:40
ZOOM ID: 976 5819 4637

Prof. Wolfgang Osten

University of Stuttgart, Germany

About Prof. Wolfgang Osten: he received the MSc/Diploma in Physics from the Friedrich-Schiller-University Jena in 1979. From 1979 to 1984 he was a member of the Institute of Mechanics in Berlin working in the field of experimental stress analysis and optical metrology. In 1983 he received the PhD degree from the Martin-Luther-University Halle-Wittenberg for his thesis in the field of holographic interferometry. From 1984 to 1991 he was employed at the Central Institute of Cybernetics and Information Processes ZKI in Berlin making investigations in digital image processing and computer vision. Between 1988 and 1991 he was heading the Institute for Digital Image Processing at the ZKI. In 1991 he joined the Bremen Institute of Applied Beam Technology (BIAS) to establish and to direct the Department Optical 3D-Metrology till 2002. Since September 2002 he has been a full professor at the University of Stuttgart and director of the Institute for Applied Optics. From 2006 till 2010 he was the vice rector for research and technology transfer of the Stuttgart University where he is currently the vice chair of the university council. His research work is focused on new concepts for industrial inspection and metrology by combining modern principles of optical metrology, sensor technology and image processing. Special attention is directed to the development of resolution enhanced technologies for the investigation of micro and nano structures.

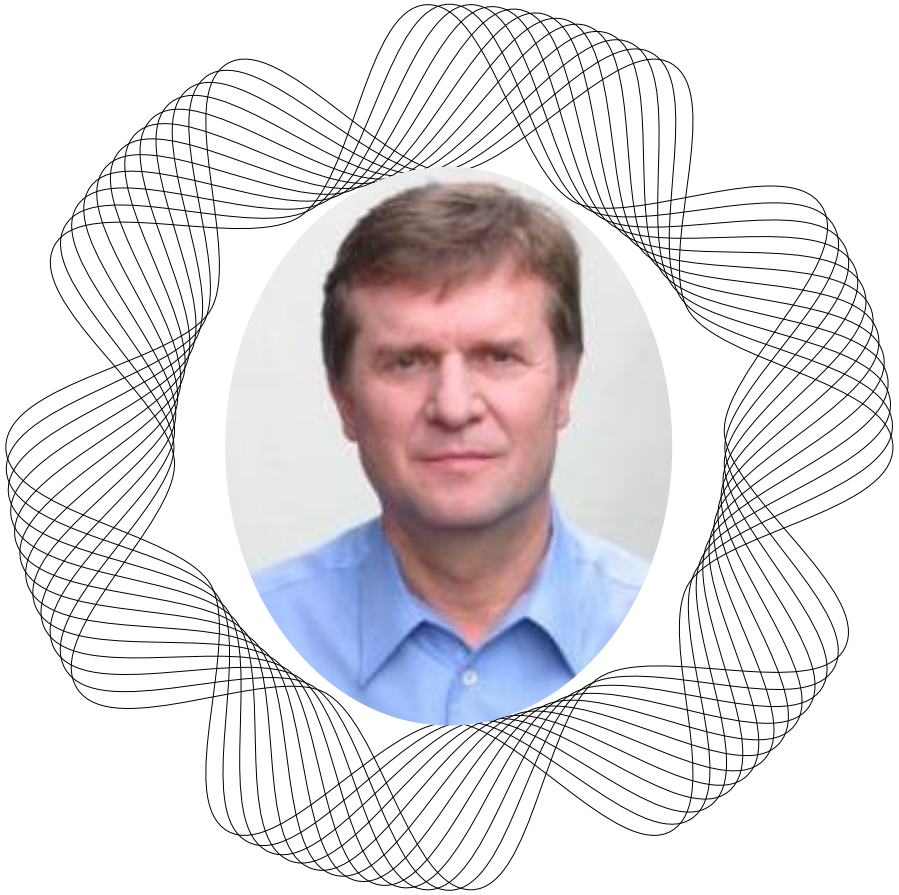
PLENARY SPEECH I

TITLE: How to Design an Optical Measurement Systems for Machine Vision with Outstanding Performance: Methods and Examples

In the context of measurement technology, optical methods have a number of unique features. These features include in particular the non-contact and high speed interaction with the object under test, the largely free scalability of the dimension of the probing tool, the high resolution of the data, the diversity of information channels in the light field, and the flexible adaptability of the comparative standard – the wavelength. On the other hand due to increased demands from production, the user is confronted with a number of serious challenges. Two of the biggest challenges that currently attract high attention in both the technical as well as life sciences, relate to exceeding the physical limits of resolution and to improve the precision of the measurement. Therefore optical measurement methods applied in machine vision are subject to constant improvement. The characteristics that give rise to improve the performance of the systems are obviously dependent on the purpose of the measurement and the object under test. But there are also general features that can be used to assess the performance of a measurement system. Here we refer to the spatial and temporal resolution, the area related resolution, the precision and trueness of the results, the robustness, the degree of automation, the process capability and the ability to work as close as possible to the process. In this contribution we describe the current challenges for measurement systems. Based on this we discuss general and application dependent features for the assessment of modern optical measurement systems. Afterwards, we describe measures to assess and to improve their performance. Finally, we show an advanced optical measurement system where several of these features are considered with regard to ensuring a high performance.

PLENARY SPEAKER II

Tuesday, Nov. 3, 2020
Speech Time: 09:30-10:10
ZOOM ID: 976 5819 4637



Prof. Sergei Gorlatch

University of Muenster, Germany

Sergei Gorlatch is Full Professor of Computer Science at the University of Muenster (Germany) since 2003. Earlier he was Associate Professor at the Technical University of Berlin, Assistant Professor at the University of Passau, and Humboldt Research Fellow at the Technical University of Munich, all in Germany. Prof. Gorlatch has more than 200 peer-reviewed publications in renowned international books, journals and conferences. He was principal investigator in several international research and development projects in the field of software for parallel, distributed, Grid and Cloud systems and networking, funded by the European Community and by German national bodies.

PLENARY SPEECH II

TITLE: Distributed Applications Based on Mobile Cloud and Software-Defined Networks

We consider an emerging class of challenging software applications called Real-Time Online Interactive Applications (ROIA). ROIA are networked applications connecting a potentially very high number of users who interact with the application and with each other in real time, i.e., a response to a user's action happens virtually immediately. Typical representatives of ROIA are multiplayer online computer games, advanced simulation-based e-learning, and serious gaming. All these applications are characterized by high performance and QoS requirements, such as: short response times to user inputs (about 0.1-1.5 s); frequent state updates (up to 100 Hz); large and frequently changing numbers of users in a single application instance (up to tens of thousands simultaneous users).
This talk will address two challenging aspects of software for future Internet-based ROIA applications: a) using Mobile Cloud Computing for allowing high application performance when a ROIA application is accessed from multiple mobile devices, and b) managing dynamic QoS requirements of ROIA applications by employing the emerging technology of Software-Defined Networking (SDN).

PLENARY SPEAKER III

Tuesday, Nov. 3, 2020
Speech Time: 10:10-10:50
ZOOM ID: 976 5819 4637

Prof. Alexander Bernstein

Skolkovo Institute of Science and Technology, Russia

Prof. Alexander Bernstein is a Principal Research Scientist in Skolkovo Institute of Science and Technology (Skoltech), Center for Computational and Data-Intensive Science and Engineering. Prof. Bernstein started his career at the Research Institute of Automatic Equipment in 1969, where he was developing mathematical models and algorithms for computer networks. At the same time, he was engaged in mathematical statistics. For research in this field, he received the Candidate of Sciences Degree in Math from Steklov Mathematical Institute of USSR Academy of Sciences in 1973 and the Doctor of Sciences degree in Math in 1987 from the Department of Computational Mathematics and Cybernetics of Moscow State University In the year of 1991, the Higher Attestation Commission awarded Alexander with the academic rank of Professor in the field of Intelligent Technologies and Systems. In 2002, Alexander joined the Software Engineering Center of the Russian Academy of Sciences to lead the projects in developing data analysis and applied mathematics software. Prior to joining Skoltech, Alexander held the positions of Chief Researcher at the Institute for System Analysis RAS and Lead Scientist in the Data Analysis Research Lab at the Institute for Information Transmission Problems RAS. At the same time, he had part-time full professor positions at National Research University Higher School of Economics and Moscow Institute of Physics and technology.

PLENARY SPEECH III

TITLE: Semantic Segmentation in Machine Vision

Scene understanding, which is a core high-level machine vision problem for inferring knowledge from imagery, is based usually on the solutions of a few machine vision task: image classification (categorizing the entire image into a class), object detection / localization (detecting objects within the image and localizing them by centroids or bounding boxes) and segmentation (delineating the boundary of the objects in the image). The speech is devoted to the Image segmentation which lays the basis for performing object detection and classification by labeling specific semantically meaningful regions of an image according to what's being shown (to classify each region into one of the pre-determined classes). We will consider a semantic image segmentation task whose goal is to label each pixel of an image with a corresponding class of what is being represented. The speech will also talk about further development of full-pixel semantic segmentation, such as instance segmentation (separate labels for different objects of the same class) and even part-based segmentation (low-level decomposition of already segmented classes into their components).

Image segmentation techniques that were commonly used in the past (clustering, histogram-based image segmentation, edge detection, etc.) are applicable only for segmentation of one image, use rigid algorithms, and require human intervention and expertise. Therefore, the report will describe in detail about modern image segmentation techniques which are powered by deep learning technologies.

Typical semantic segmentation algorithms based on training deep learning model require fully supervised annotated large datasets in which every pixel of every image must be labeled with a class of interest label or background in order to train the deep learning model. Coming up with such labels comes at a heavy cost in time and money. Therefore, the speech will present also the modern methods of so-called Weakly-supervised semantic segmentation in which a pixel-wise segmentation of an image is performed on the basis of only the labels of existing semantic objects in the image.

PLENARY SPEAKER IV

Tuesday, Nov. 3, 2020
Speech Time: 11:05-11:45
ZOOM ID: 976 5819 4637

Prof. Petia Radeva

University of Barcelona, Spain

Prof. Petia Radeva completed her undergraduate study on Applied Mathematics at the University of Sofia, Bulgaria, in 1989. In 1996, she received a Ph.D. degree in Computer Vision at UAB. In 2007, she moved as Tenured Associate professor at the Universitat de Barcelona (UB), Department of Mathematics and Informatics, where from 2009 to 2013 she was Director of Computer Science Undergraduate Studies. Petia Radeva is Head of the Consolidated Group Computer Vision at the University of Barcelona (CVUB) at UB (www.ub.edu/cvub) and Head of the Medical Imaging Laboratory of Computer Vision Center (www.cvc.uab.es). Petia Radeva's research interests are on Development of learning-based approaches (specially, deep learning) for computer vision, and their application to health. Currently, she is involved on projects that study the application of wearable cameras and life-logging, to extract visual diary of individuals to be used for memory reinforcement of patients with mental diseases (e.g. Mild cognitive impairment). Moreover, she is exploring how to extract semantically meaningful events that characterize lifestyle and healthy habits of people from egocentric data. She is associate editor of Pattern Recognition journal and International Journal of Visual Communication and Image Representation. She obtained the ICREA award from the Catalonian Government for her scientific merits in 2014, the international award "Aurora Pons Porrata" from CIARP in 2016 and the Prize "Antonio Caparrós" for the best technology transfer project of 2013.

PLENARY SPEECH IV

TITLE:

Special Session I:
Camera Based and Mobile Recognition

Monday, Nov. 2, 2020
Speech Time: 13:30-17:00
ZOOM ID: 976 5819 4637

Chair: Prof. Vladimir Arlazarov

Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Russia

Vladimir holds a Ph.D. in technical sciences as a specialist in development, application of recognition systems and data mining. He has worked in the IT fields since 1997, leading and managing more than 70 projects related to creation and embedding of document recognition systems for Pension Funds, Ministries, Banks, Public Transportation and Oil & Gas companies. Under his leadership several generations of industrial OCR and document capture systems were developed, including the industrial system for identity documents recognition. Tremendous visionary entrepreneur, Vladimir is the author of several scientific patents about artificial intelligence and he is focusing on advancing the company R&D leadership in computer vision software. At 2019 he has 50 publications and patents.

SPECIAL SESSION I

TOPIC: Camera Based and Mobile Recognition

Image recognition as a service differs from industrial machine vision in many ways. Not only the environment is uncontrolled, but camera hardware is unknown too. In the case of on-device recognition computational complexity of recognition methods and size of binary data (including CNN models) are critical. Situation is better in the presence of specialized processing hardware, but this implies proper software adaptation. Users cooperativity sighnifically improves recognition quality. On the other hand, random user can feed system with irrelevant data. These points clearly should not be ignored. We invite to this track specialists both in software and hardware to discuss mentioned problems and their possible solutions.

Special Session II:
Advanced Imaging and Tomography

Tuesday, Nov. 3, 2020
Speech Time: 13:30-16:40
ZOOM ID: 976 5819 4637

Chair: Prof. Alessia Cedola

The Sapienza University of Rome, Rome unit
Nanotec CNR, Italy

About Prof. Alessia Cedola: Actual position: permanent scientist at Institute of Nanotecnology Laboratory for Soft and Living Matter of the National Research Council (CNR) in Rome. Enabled Associate Professor of Experimental Physics. Academic Studies: [1999] PhD at the University "Joseph Fourier" in Grenoble (France), summa cum laude. Experimental thesis at European Synchrotron Radiation Facility (ESRF). Main research interests: X-ray imaging, X-ray Phase Contrast Imaging and Tomography, Biomedical applications. Supervisor of several Laurea and PhD Thesis, Teachers of National and International schools of Physics, Responsible for several scientific contracts for young researchers, Chair of several symposia and workshops. Scientific roles: Responsible of the X-ray physics group since 2012, Member of the Management Committee of two ESF COST projects, Member of the Scientific Committee of several workshops and conferences, 2015 - Member of the Editorial Board of Scientific Reports

SPECIAL SESSION II

TOPIC: Advanced Imaging and Tomography

Machine vision is one of the most powerful nondestructive methods of object exploration, but it is limited by object surface. For internal structure analysis (in medicine, geology, quality control, etc.) one requires the methods of tomography. Advanced imaging and tomography track will provide a forum to present and discuss the latest advances and state of the art approaches in tomography. We intend to organize a fruitful exchange of opinions and ideas between specialists in machine vision and tomography, as well as encourage and facilitate interdisciplinary communication amongst university researches and industry professionals involved in tomographic software and hardware development.

Special Session III

Machine Vision for Autonomous Driven Cars under Harsh Environmental Conditions

Wednesday, Nov. 4, 2020
Speech Time: 09:30-12:00
ZOOM ID: 976 5819 4637

Chair: Prof. Wolfgang Osten

University of Stuttgart, Germany

About Prof. Wolfgang Osten: he received the MSc/Diploma in Physics from the Friedrich-Schiller-University Jena in 1979. From 1979 to 1984 he was a member of the Institute of Mechanics in Berlin working in the field of experimental stress analysis and optical metrology. In 1983 he received the PhD degree from the Martin-Luther-University Halle-Wittenberg for his thesis in the field of holographic interferometry. From 1984 to 1991 he was employed at the Central Institute of Cybernetics and Information Processes ZKI in Berlin making investigations in digital image processing and computer vision. Between 1988 and 1991 he was heading the Institute for Digital Image Processing at the ZKI. In 1991 he joined the Bremen Institute of Applied Beam Technology (BIAS) to establish and to direct the Department Optical 3D-Metrology till 2002. Since September 2002 he has been a full professor at the University of Stuttgart and director of the Institute for Applied Optics. From 2006 till 2010 he was the vice rector for research and technology transfer of the Stuttgart University where he is currently the vice chair of the university council. His research work is focused on new concepts for industrial inspection and metrology by combining modern principles of optical metrology, sensor technology and image processing. Special attention is directed to the development of resolution enhanced technologies for the investigation of micro and nano structures.

SPECIAL SESSION III

TOPIC: Machine Vision for Autonomous Driven Cars under Harsh Environmental Conditions

The development of autonomous driven cars is a hot topic taken up not only by industry but increasingly by research institutions. The topic is not new. Already in the 80th big research projects such as the European EUREKA project “Prometheus” (PROgraMme for a European Traffic of Highest Efficiency and Unprecedented Safety, 1986–1994) were aimed at finding new solutions for increased road safety. However, the key finding gained almost 30 years ago was not so much an increase in autonomy as improved driver assistance, in order to provide more support to the driver in more complex situations. But new algorithms and implementations of Artificial Intelligence, sophisticated sensors and sensor fusion techniques are currently promising a significant improvement in the reliability of autonomous vehicles. In that context a lot of challenging problems are waiting on an answer. To them belong especially questions directed to changing environmental conditions during the autonomous drive such as rain, fog, snow, and in general difficult visibility conditions. If vehicles are to drive fully autonomously in the future, systems are required which scan the environment with high precision, high spatial and temporal resolution and guarantee the necessary reliability even in these difficult situations.

Today, a wealth of different sensors are integrated and combined. The sensor and data fusion can be described as inadequate in current configurations, since the individual data streams are processed and interpreted in parallel. A combination of the results takes place very late in the process chain. Inadequate results characterized by latency and blurring are the result. The special session is dedicated to that challenge. Various lectures address new approaches for better dealing with such complex situations. To them belong new ways of machine Learning for infrastructure monitoring, new LiDAR sensors for obstacle recognition, and two new sensors for the improvement of object recognition through scattering media.

SPECIAL SESSION IV

Wednesday, Nov. 4, 2020
Speech Time: 13:00-14:50
ZOOM ID: 976 5819 4637

TOPIC: Computer Optics Journal

Chair: Prof. Vladislav Sergeev

Samara National Research University, Russia

Computer Optics Journal is abstracted and indexed in international Science Databases, like Scopus (Elsevier), Ei Compendex (Elsevier), Web of Science: Esci(Clarivate Analytics).
The journal is indexed in Key Russian databases, like RISC (Russian Index of Science Citation), "Cyberleninka" (sciencetific e-libreary), Open Academic Journal Index (OAJI), MathNet (All-Russian Mathematical Portal).

The Major Journal Metrics:
CiteScore 2019---3.9 SJR 2019-0.586 SNIP 2019-2.376 H-index 2019-27

Journal's Subject areas and categories (SJR)
Engineering—Q1 Computer Science—Q2 Physics and Astronomy—Q2

Subjects Covered:
Diffrative Optics/Information Optical Technology/ Nanophotonics and Optics Nanostructures/ Image Analysis& Understanding/ Digital Signal Processing/ Information Coding& Security/ Earth Remote Sensing Technology/ Hyperspectral Data Analysis/ Numerial Methods for Optics and Image Processing /Intelligent Video Analysis

Track

Wednesday, Nov. 4, 2020
Speech Time: 15:10-17:30
ZOOM ID: 976 5819 4637

TOPIC: Color Computer Vision

Chair: Dr. Egor Ershov

Vision Systems Lab., Institute of Information Transmission Problems, Russian Academy of Science, Russia

Institute for Information Transmission Problems of the Russian Academy of Sciences (Kharkevich Institute or IITP RAS), in cooperation with University of Zagreb, invites research groups and individual enthusiasts to participate in the 2nd International Illumination Estimation Challenge (IEC).

The main goal of the challenge in this year is to develop novel algorithms for estimation of multiple light sources scene illumination and demonstrate its effectiveness using large and diverse image dataset.

This year IEC event will supplement the regular program of the 13th ICMV, which bring together leading experts in computer vision and image processing.

/03

PART III

Technical Program

HIGHLIGHT GUIDE

Special Session I

Camera Based and Mobile Recognition

13:30-17:00, 2nd Nov.

Special Session II

Advanced Imaging and Tomography

13:30-16:40, 3rd Nov.

Special Session III

Machine Vision for Autonomous Driven Cars under Harsh Environmental Conditions

09:30-12:00, 4th Nov.

Special Session IV

Computer Optics Journal

15:10-17:30, 4th Nov.

Track

Color Computer Vision

13:30-17:00, 4th Nov.

Closing Ceremony

16:30-17:00, 6th Nov.

TECHNICAL PROGRAM

Special Session I-a

Camera Based and Mobile Recognition

Monday, Nov. 2, 2020
13:30-17:00
ZOOM ID: 976 5819 4637

Chair: Prof. Vladimir Arlazarov, Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, Russia

13:30-13:50	V051	Choosing the best image of the document owner's photograph in the video stream on the mobile device Polevoy D.V., **Mr. Aliev M.A.,** Nikolaev D.P. Smart Engines Service LLC, Russia
13:50-14:10	V022	Tracking and identification for football video analysis using deep learning **Dr. Shreedhar Rangappa**, Baihua Li, Ruiling Qian Loughborough University, UK
14:10-14:30	V068	A modification of a stopping method for text recognition in a video stream with best frame selection **Mr. Ilya Tolstov**, Stanislav Martynov, Vera Farsobina, Konstantin Bulatov Keldysh Institute of Applied Mathematics of Russian Academy of Sciences, Russia
14:30-14:50	V044	About Viola-Jones image classifier structure in the problem of stamp detection in document images **Mr. Daniil P. Matalov**, Sergey A. Usilin, Vladimir V. Arlazarov Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, Russia
14:50-15:10	V089	Distance-based online pairs generation method for metric networks training **Mr. Ivan Kondrashev**, Alexander Sheshkus, Vladimir V. Arlazarov Bauman Moscow State Technical University, Russia
15:10-15:20		Break Time

TECHNICAL PROGRAM

Special Session I-b

Camera Based and Mobile Recognition

Chair: Prof. Vladimir Arlazarov, Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, Russia

15:20-15:40	V092	Block convolutional layer for position dependent features calculation **Mr. Sergey A. Ilyuhin** , Alexander V. Sheshkus, Vladimir L. Arlazarov Smart Engines Service LLC; Moscow Institute of Physics and Technology, Russia
15:40-16:00	V104	Improved algorithm of ID card detection by a priori knowledge of the document aspect ratio **Mr. Daniil V. Tropin**, Ivan A. Konovalenko, Natalya S. Skoryukina, Dmitry P. Nikolaev, Vladimir V. Arlazarov MIPT (NRU), Russia
16:00-16:20	V060	Image stacking versus per-frame results combination for video text recognition **Mr. Savelyev**, Konstantin Bulatov, Vladimir L. Arlazarov Smart Engines, Moscow, Russia; Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, Russia
16:20-16:40	V075	Generative approach for 1D barcode dataset population for mobile-based recognition **Mr. Milovzorov A.N**, Bezmaternykh P.V., Arlazarov V.V. National University of Science and Technology “MISIS”, Russia
16:40-17:00	V084	Chess recognition using 3D patterned illumination camera Lars Brunner, Mario Salvator, Philipp Roebrock, **Prof. Udo J. Birk** University of Applied Sciences of the Grisons, Switzerland

TECHNICAL PROGRAM

Special Session II-a

Advanced Imaging and Tomography

Tuesday, Nov. 3, 2020
13:30-16:40
ZOOM ID: 976 5819 4637

Chair: Prof. Alessia Cedola, The Sapienza University of Rome, Rome unit Nanotec CNR, Italy		
13:30-13:50	V016	Improving the quality of tomographic images by the methods of the computer-aided measuring systems **Prof. Chulichkov A.I**., Molkov E.M., Prutjan S.I. Lomonosov Moscow State University, Russia
13:50-14:10	V021	Artifacts suppression in biomedical images using a guided filter **Dr. I. Bukreeva**, A. Ingacheva, M. Fratini, A. Cedola, O. Junemann, E. Longo, F. Wilde, J. Moosmann, A. Buzmakov, Y. Krivonosov, D. Zolotov, S. Saveliev, V. Asadchikov and M. Chukalina 'Sapienza' University, Italy
14:10-14:30	V026	Blind CT images quality assessment of cupping artifacts **Ms. Anastasia Ingacheva**, Marina Chukalina, Dmitry Nikolaev Institute for Information Transmission Problems (Kharkevich Institute) RAS, Russia
14:30-14:50	V024	Processing and understanding of images in spectral tomography **Dr. Marina Chukalina**, Anastasiya Ingacheva, Alexey Buzmakov, Dmitry Nikolaev FSRC «Crystallography and photonics» RAS, Russia
14:50-15:10	V054	Modelling the point source function for collimators employed in radio guided surgery and SPECT **Dr. Ya. L. Shabelnikova**, S. I. Zaitsev, V. B. Obolenskiy Institute of microelectronic technology problems and high purity materials RAS, Russia
15:10-15:20		Break Time

TECHNICAL PROGRAM

Special Session II-b

Advanced Imaging and Tomography

Chair: Prof. Alessia Cedola, The Sapienza University of Rome, Rome unit Nanotec CNR, Italy

15:20-15:40	V100	Empirical analysis of the optimality of RSRE-based stopping rules for monitored reconstruction Konstantin Bulatov, **Mr. Arseniy Mukovozov**, Vladimir V. Arlazarov Smart Engines Service LLC, Russia
15:40-16:00	V108	CT images GAN-based augmentation with AdaIN for lung nodules detection Maksim Kryuchkov, **Ms. Natalia Khanzhina**, Ilya Osmakov and Pavel Ulyanov ITMO University, Russia
16:00-16:20	V4001	Lightweight denoising filtering neural network for FBP algorithm **Mr. Andrei Yamaev**, Marina Chukalina, Dmitry Nikolaev, Alexander Sheshkus, Alexey Chulichkov MSU, Russia
16:20-16:40	V047	Iterative reconstruction of incomplete tomography data: Application cases **Dr. Alexey Buzmakov**, Denis Zolotov, Marina Chukalina, Anastasia Ingacheva, Victor Asadchikov, Dmirty Nikolaev, Yuri Krivonosov, Irina Dyachkova, Inna Bukreeva FSRC “Crystallography and photonics” RAS, Russia

TECHNICAL PROGRAM

Special Session III-a

Machine Vision for Autonomous Driven Cars under Harsh Environmental Conditions

Wednesday, Nov. 4, 2020
09:30-12:00
ZOOM ID: 976 5819 4637

Chair: Prof. Wolfgang Osten, University of Stuttgart, Germany		
09:30-09:50	V053	Local path planning algorithm for autonomous vehicle based on multi-objective trajectory optimization in state lattice **Mr. Ivan I. Kornev**, Vladislav I. Kibalov, Oleg S. Shipitko Institute for Information Transmission Problems – IITP RAS, Bol'shoy Karetnyy Pereulok 19, Russia
09:50-10:10	V065	Deep car detection by fusing grayscale image and weighted upsampled LiDAR depth **Mr. Meisam Jamshidi Seikavandi**, Kamal Nasrollahi, Thomas B. Moeslund Khaje Nasir University of Technology, Iran
10:10-10:30	V061	An approach to road scene text recognition with per-frame accumulation and dynamic stopping decision **Dr. Konstantin Bulatov**, Nadezhda Fedotova, Vladimir V. Arlazarov Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Russia
10:30-10:40		Break Time

TECHNICAL PROGRAM

Special Session III-b
Machine Vision for Autonomous Driven Cars under Harsh Environmental Conditions

Chair: Prof. Wolfgang Osten, University of Stuttgart, Germany		
10:40-11:00	V069	A machine vision based smart conveyor system Mengchao Zhang, **Dr. Vedang Chauhan** and Manshan Zhang Western New England University, US
11:00-11:20	V091	A target-free calibration method for automotive augmented reality head-up displays **Mr. Xiang Gao**, Kai Wu, Marc Necker, Wilhelm Stork, Adnane Jadid, Gudrun Klinker Mercedes-Benz AG & Karlsruhe Institute of Technology, Germany
11:20-11:40	V094	Fast hough transform-based road markings detection for autonomous vehicle **Mr. Oleg Shipitko**, Ekaterina Panfilova and Irina Kunina Smart Engines Service LLC, Russia
11:40-12:00	V059	LIDAR-based parking spot search algorithm **Mr. Dinir Imameev**, Aufar Zakiev, Tatyana Tsoy, Yang Bai, Mikhail Svinin, Evgeni Magid Kazan Federal University, Russia

TECHNICAL PROGRAM

Special Session IV

Computer Optics Journal

Wednesday, Nov. 4, 2020
13:00-14:50
ZOOM ID: 976 5819 4637

Prof. Vladislav Sergeev, Samara National Research University, Russia

13:00-13:20	V5001-A	Algorithm for post-processing of tomographic images to calculate the dimension-geometric features of porous structures **M. V. Chukalina**, A. V. Khafizov, Mr. V. V. Kokhan, A. V. Buzmakov, R. A. Senin, V. I. Uvarov, M. V. Grigorev Smart Engines LLC, Institute for Information Transmission Problems RAS, Russia
13:20-13:40	V5002-A	Weighted combination of per-frame recognition results for text recognition in a video stream **Ms. Olga Petrova**, Konstantin Bulatov, Vladimir V. Arlazarov, Vladimir L. Arlazarov Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, Russia
13:40-14:00	V5003-A	A generalization of Otsu's Method for linear separation of two unbalanced classes in document image binarization E.I. Ershov, S.A. Korchagin, **Mr. V.V. Kokhan**, P.V. Bezmaternykh Smart Engines LLC, Institute for Information Transmission Problems RAS, Russia
14:00-14:20	V5004-A	Algorithm for choosing the best frame in a video stream in the task of identity document recognition Aliev M.A., Kunina I.A., **Mr. Kazbekov A.V**., Arlazarov V.L. Smart Engines Service LLC, Moscow, Russia
14:20-14:40	V5005-A	Optimal affine image normalization approach for optical character recognition I.A. Konovalenko, **Mr. V.V. Kokhan**, D.P. Nikolaev Smart Engines LLC, Institute for Information Transmission Problems RAS, Russia
14:40-14:50		Break Time

TECHNICAL PROGRAM

Track

Color Computer Vision

Wednesday, Nov. 4, 2020
14:50-17:20
ZOOM ID: 976 5819 4637

Chair: Prof. Egor Ershov, Vision Systems Lab., Institute of Information Transmission Problems, Russian Academy of Science, Russia

14:50-15:00		Speech
15:00-15:20	V082	Consensus-driven Illuminant Estimation with GANs **Dr. Marco Buzzelli**, Riccardo Riva, Simone Bianco, Raimondo Schettini University of Milano – Bicocca
15:20-15:40	V077	Robust white balance estimation using joint attention and angular loss optimization **Mr. Zhihao Li,** Zhan Ma Nanjing University, China
15:40-16:00	V072	SDE-AWB: a generic solution for 2nd international illumination estimation challenge **Dr. Yanlin Qian**, Sibo Feng, Kang Qian, Miaofeng Wang Tampere University, Finland
16:00-16:20	V057	Color correction of the document owner's photograph image during recognition on mobile device Dmitry Polevoy, **Ms. Ekaterina Panfilova**, Egor Ershov, Dmitry Nikolaev Smart Engines Service LLC ; V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences; The Institute for Information Transmission Problems of Russian Academy of Sciences, Russia
16:20-16:30		Break Time

TECHNICAL PROGRAM

Track

Color Computer Vision

Chair: Prof. Egor Ershovn, Vision Systems Lab., Institute of Information Transmission Problems, Russian Academy of Science, Russia

16:30-16:50	V006	A low-complexity yet accurate calibration method for automotive augmented reality head-up displays **Mr. Xiang Gao**, Marc Necker, Wilhelm Stork Mercedes-Benz AG & Karlsruhe Institute of Technology, Germany
16:50-17:10	V020	Camera auto-calibration for complex scenes **Mr. Anas Ali** and Pavel Smrz Brno University of Technology, Czech
17:10-17:30	V045	A dynamic programming inspired outlier rejection algorithm for image mosaicing problem Christopher Smith and **Dr. Semih Dinc** Auburn University at Montgomery, USA

TECHNICAL PROGRAM

Session 6

Artificial Intelligence and Intelligent Computing

Thursday, Nov. 5, 2020
09:30-11:15
ZOOM ID: 976 5819 4637

Chair: Prof. Johan DEBAYLE, Ecole Nationale Supérieure des Mines de Saint-Etienne, France

09:30-09:45	V017	Model selection for support-vector machines through metaheuristic optimization algorithms Oumeima Ghnimi, **Dr. Sofiane Kharbech**, Akram Belazi, and Ammar Bouallegue Tunis El Manar University, Tunisia / University of Lille, France
09:45-10:00	V030	A semi-supervised learning approach for CBIR systems with relevance feedback Marco Brighi, **Dr. Annalisa Franco**, Dario Maio University of Bologna, Italy
10:00-10:15	V042	Deep convolutional neural network based autonomous drone navigation **Mr. Karim Amer**, Mohamed Samy, Mahmoud Shaker, and Mohamed Elhelw Nile University, Egypt
10:15-10:30	V070	Spatiotemporal feature based convolutional neural network for violence detection **Ms. Amira Ben Mabrouk**, Ezzeddine Zagrouba Higher Institute of Computer Science, University of Tunis EL Manar, Tunisia
10:30-10:45	V088	Almost indirect 8-bit convolution for QNNs **Mr. Anton Trusov**, Elena Limonova, Sergey Usilin Moscow Institute of Physics and Technology, Russia
10:45-11:00	V095	DVAE-SR: Denoiser variational auto-encoder and super-resolution to counter adversarial attacks **Mr. Omar Dardour**, Mourad Zaied, Petia Radeva RTIM-Lab, Tunisia
11:00-11:15	V101	Abstractive meeting summarization based on an attentional neural model **Ms. Nouha Dammak**, Yassine BenAyed University of Sousse, Tunisia

TECHNICAL PROGRAM

Session 7

Software and Information Engineering

Thursday, Nov. 5, 2020
09:30-11:30
ZOOM ID: 680 4475 3785

Chair: Prof. Hiroshi Hosobe, Hosei University, Japan

09:30-09:45	SE020	Continuous development and testing of access and usage control: A systematic literature review Said Daoudagh, **Dr. Francesca Lonetti**, Eda Marchetti ISTI-CNR, Italy
09:45-10:00	SE030	Development frameworks for microservice-based applications: Evaluation and comparison **Mr. Hai Dinh-Tuan**, Maria Mora-Martinez, Felix Beierle, Sandro Rodriguez Garzon Technische Universität Berlin, Germany
10:00-10:15	SE031	The link between transformational and servant leadership in devops-oriented organisations **Mr. Krikor Maroukian**, Stephen R. Gulliver Microsoft, Greece
10:15-10:30	SE024	Maintainability metrics for android applications in Kotlin: An evaluation of tools **Ms. Lisa-Marie Andrä**, Bernhard Taufner, Sigrid Schefer-Wenzl, Igor Miladinovic FH Campus Vienna, Austria
10:30-10:45	SE029	Testing Event-Driven Programs in Processing **Prof. Hiroshi Hosobe** Hosei University, Japan
10:45-11:00	SE009	Communication challenges in agile teams from the communication theory prospective **Ms. Sofiia Yermolaieva** Innopolis University, Russia
11:00-11:15	SE003	A dashboard to support decision-making processes in learning ecosystems: A metamodel integration **Prof. Francisco José-García Peñalvo**, Andrea Vázquez-Ingelmo, Alicia García-Holgado, Roberto Therón University of Salamanca, Spain

TECHNICAL PROGRAM

Session 7

Software and Information Engineering

Chair:		
11:15-11:30	SE010	Empirical study on commonly used combinations of estimation techniques in software development planning **Mr. Arman Kialbekov** Innopolis University, Russia

TECHNICAL PROGRAM

Session 8

Target Detection

Thursday, Nov. 5, 2020
13:00-14:30
ZOOM ID: 976 5819 4637

Chair:		
13:00-13:15	V003	Violent scenes detection based on connected component analysis Samira Labbadi, Mariem Gnouma, Ridha EJbali and **Prof. Mourad Zaied** Research Team in Intelligent Machines, Tunisia
13:15-13:30	V010	Fatigue detection based on non-contact respiratory detection **Mr. Xing Chen**, Lumei Su, Min Xu, Bo Deng, Yuanxuan Zhu Xiamen University of Technology, China
13:30-13:45	V041	Robust real-time pedestrian detection on embedded devices Mohamed Afifi, Yara Ali, **Mr. Karim Amer**, Mahmoud Shaker, and Mohamed Elhelw Nile University, Egypt
13:45-14:00	V043	Line detection via a lightweight CNN with a Hough Layer **Mr. Lev Teplyakov**, Kirill Kaymakov, Evgeny Shvets and Dmitry Nikolaev IITP RAS, Russia
14:00-14:15	V050	Maximizing object detection using suas Curtis Manore, **Mr. Pratheek Manjunath**, Dominic Larkin United States Military Academy, USA
14:15-14:30	V096	Multimodal features for shots boundary detection **Mr. Mohamed Bouyahi**, Yassine Ben Ayed MIRACL laboratory, Tunisia

TECHNICAL PROGRAM

Session 9

Artificial Intelligence and Data Engineering

Thursday, Nov. 5, 2020
13:00-14:45
ZOOM ID: 680 4475 3785

Chair: Assoc. Prof. Stephen R. Gulliver, University of Reading, UK

13:00-13:15	S027	A neural-assessment system based on Emirates (QFE) Fatima Alzahraa Kouka, Lamees Mohammad Dalbah, Shaimaa Mahmood Mounir, **Prof. Raed Abu Zitar** Ajman University, United Arab Emirates
13:15-13:30	SE022	Specifying key-properties to improve the recognition skills of neural networks **Mr. Benjamin Jahić**, Guelfi Nicolas, Ries Benoît University of Luxembourg, Luxembourg
13:30-13:45	SE023	Dynamic data consistency tests using a CRUD Matrix as an underlying model **Assoc. Prof. Miroslav Bures**, Vaclav Rechtberger Czech Technical University in Prague, Czech Republic
13:45-14:00	SE042	A framework based on model driven engineering and model weaving to support data-driven interoperability for smart grid Applications **Mr. Eder Matheus Silveira Felix**, Denivaldo Cicero Lopes, Osvaldo Silva Sousa Jr. Federal University of Maranhão, Brazil
14:00-14:15	S001	Constructing neural networks by extending the optimization field **Prof. Wenyuan Zhang** Huazhong University of Science and Technology, China
14:15-14:30	S020	On the relationship between input sparsity and noise robustness in Hierarchical Temporal Memory Spatial Pooler **Dr. Damir Dobric**, Andreas Pech, Bogdan Ghita, Thomas Wennekers University of Plymouth, Germany
14:30-14:45	S024	Improving fairness in speaker recognition Gianni Fenu, **Dr. Giacomo Medda**, Mirko Marras, Giacomo Meloni University of Cagliari, Italy

TECHNICAL PROGRAM

Session 10

Image Analysis and Methods

Thursday, Nov. 5, 2020
15:00-16:30
ZOOM ID: 976 5819 4637

Chair: Prof. Udo J. Birk, University of Applied Sciences of the Grisons, Switzerland

15:00-15:15	V027	Keypoint-based Static Object Removal from Photographs Alexandr Volkov, **Ms. Valeria Efimova**, Viacheslav Shalamov and Andrey Filchenkov ITMO University, Russia
15:15-15:30	V033	Improvement of U-Net architecture for image binarization with activation functions replacement **Mr. Alexander Gayer**, Alexander Sheshkus, Dmitri P. Nikolaev, Vladimir V. Arlazarov NUST “MISIS”, Russia
15:30-15:45	V037	On the effectiveness of adversarial unsupervised domain adaptation for iris presentation attack detection in mobile devices **Mrs. Yomna Safaa-El-Din**, Mohamed Moustafa and Hani Mahdi Ain Shams University – Cairo, Egypt
15:45-16:00	V038	The method of search for falsifications in copies of contractual documents based on N-grams Oleg Slavin, **Ms. Elena Andreeva**, Vladimir V. Arlazarov Smart Engines Service Ltd., Russia
16:00-16:15	V049	LRA-Net: Local region attention network for 3D point cloud completion **Mr. Hang Wu**, Yubin Miao Shanghai Jiao Tong University, China
16:15-16:30	V087	Slope detection criterion robust to sparse 2D data **Mr. Dmitry Bocharov**, Alexey Kroshnin, Dmitry Nikolaev Institute for Information Transmission Problems, RAS, Russia

TECHNICAL PROGRAM

Session 11

Computer and Information Security

Thursday, Nov. 5, 2020
15:00-17:00
ZOOM ID: 680 4475 3785

Chair:		
15:00-15:15	SE012	A cost model for decoder decision trees **Ms. Lillian Tadros** Technical University of Dortmund, Germany
15:15-15:30	SE021	A generic framework for capturing reliability in cyber-physical systems **Mr. Nazakat Ali**, Manzoor Hussain, YoungJae Kim and Jang-Eui Hong Chungbuk National University, Republic of South Korea
15:30-15:45	SE025	LDA categorization of security bug reports in chromium projects **Mr. Wajdi Mohammed Aljedaani**, Yasir Javed, Mamdouh Alenezi Al-Kharj College of Technology, Saudi Arabia
15:45-16:00	SE026	Supporting contact tracing by privacy-friendly registration at catering facilities **Dr. Michiel Willocx**, Dave Singelée, Jorn Lapon, Vincent Naessens Imec-DistriNet, Belgium
16:00-16:15	S025	CRATOS: Cognition of reliable algorithm for time-series optimal solution **Dr. Ziling Wu**, Ping Liu, Zheng Hu, Bocheng Li, Jun Wang Sun Yat-sen University, China
16:15-16:30	S002	A survey on information diffusion in online social networks **Dr. Yujie YANG** Wuhan University, China

TECHNICAL PROGRAM

Session 11

Computer and Information Security

Chair:		
16:30-16:45	SE028	An approach based on model driven engineering to support the development of web of things **Ms. Rayanne Silva de Oliveira**, Denivaldo Lopes Federal University of MaranhãoUFMA, Brazil
16:45-17:00	SE044	An empirical investigation of spikes in agile software development **Mr. Hussein Al Hashimi**, Abdullah Altaleb, Andrew Gravell University of Southampton, UK

TECHNICAL PROGRAM

Session 12

Intelligent Recognition Technology and Algorithm

Friday, Nov. 6, 2020
10:00-11:30
ZOOM ID: 976 5819 4637

Chair:		
10:00-10:15	V005	A novel approach to improve the social acceptance of autonomous driving vehicles by recognizing the emotions of passengers **Dr. Antonio Costantino Marceddu**, Jacopo Sini, Massimo Violante, Bartolomeo Montrucchio Polytechnic of Turin, Italy
10:15-10:30	V015	Steel pipe counting system based on image recognition Yuxin Jia, Jie Xu, **Dr. Qujiang Lei** , Xiuhao Li, Guangchao Gui, Weijun Wang Guangzhou Institute of Advanced Technology, Chinese Academy of Sciences, China
10:30-10:45	V018	the impact of pre-processing algorithms in facial expression recognition **Mr. Daniel Canedo**, António J. R. Neves University of Aveiro, Portugal
10:45-11:00	V052	Precise localization of synchronization patterns for Aztec Code matrix extraction **Mr. Aleksey D. Bursikov**, Pavel V. Bezmaternykh and Vitaliy M. Kliatskine Moscow Institute for Physics and Technology, Russia
11:00-11:15	V073	Application of shared backbone DNNs in ADAS perception systems **Mr. Mikhail G. Lobanov**, Dmitry L. Sholomov Cognitive Robotics Ltd., Russia
11:15-11:30	V103	Fingerspelling recognition using synthetic images and deep transfer learning **Mr. Nguyen Tu Nam**, Shinji Sako, Bogdan Kwolek Nagoya Institute of Technology, Japan

TECHNICAL PROGRAM

Session 13

Digital Image Processing and Application

Friday, Nov. 6, 2020
10:00-11:45
ZOOM ID: 680 4475 3785

Chair:		
10:00-10:15	S003	Deep image compositing Shivangi Aneja, **Dr. Soham Mazumder** Technical University of Munich, Germany
10:15-10:30	S101	Dog and cat classification with deep residual network **Ms. Yao Yuchen** Software School of North University China, China
10:30-10:45	S018	Generation of datasets for semantic segmentation from 3D scanned data to train a classifier for visual navigation **Dr. Hayato Komatsuzaki**, Raimu Yokota, Shogo Sakata, Miho Adachi, Ryusuke Miyamoto Meiji University, Japan
10:45-11:00	S013	Attention guided multi-scale regression for scene text detection **Dr. Zhiwei Zheng** Huazhong University of Science and Technology, China
11:00-11:15	S021	Breast lesion detection from mammograms using deep convolutional neural networks **Dr. Gloria Gonella**, Marco Paracchini, Elisabetta Binaghi, Marco Marcon University of Insubria, Italy
11:15-11:30	S022	Development of a unified and dynamic geometric framework for modelling plant leaf spots **Dr. Duaa Alshadli**, Kambiz Borna, Cesar Lador Unitec Institute of Technology, New Zealand
11:30-11:45	S008	A fabric image segmentation pipeline to find the yarn weaving pattern from a single photograph **Dr. Cihan Bal**, Mehmet K. Baran Marmara University,Turkey

TECHNICAL PROGRAM

Session 14

Digital Image Processing and Application

Friday, Nov. 6, 2020
13:00-14:30
ZOOM ID: 976 5819 4637

Chair: Prof. Alexander Bernstein , Skolkovo Institute of Science and Technology, Russia

13:00-13:15	V008	Event correlation for deception detection in long video **Ms. Liqian Gao**, Jianbang Qin, Chunhui Du, Wei Guo Shanghai Jiao Tong University, China
13:15-13:30	V009	Weak anomaly-reinforced autoencoder for unsupervised anomaly detection **Mr. Xinqiang Chen**, Lumei Su, Guansen Deng, Mingyong Huang, Jiajun Wu and Yanqing Peng Xiamen University of Technology, China
13:30-13:45	V014	Automated synthetic datasets construction for part semantic segmentation of non-cooperative satellites **Mr. Hongkai Ding**, Jianjun Yi, Zhuoran Wang, Yajun Zhang, Hailei Wu, Shuqing Cao East China University of Science and Technology, China
13:45-14:00	V040	An application of geometric aspects of variational autoencoder model to forgery detection of scanned documents Igor Janiszewski, **Mr. Dmitry Slugin** and Elena Andreeva Federal Research Center Computer Science and Control of Russian Academy of Sciences, Russia
14:00-14:15	V067	Fader networks for domain adaptation on fMRI: ABIDE-II study Marina Pominova, **Mrs. Ekaterina Kondrateva**, Maksim Sharaev, Alexander Bernstein and Evgeny Burnaev Skolkovo University of Science and Technology
14:15-14:30	V085	Occlusion aware unsupervised learning of optical flow from video **Mr. Jianfeng Li**, Junqiao Zhao, Tiantian Feng Tongji University, China

TECHNICAL PROGRAM

Session 15

Computer Vision and Visualization

Friday, Nov. 6, 2020
13:00-14:30
ZOOM ID: 680 4475 3785

Chair:		
13:00-13:15	V039	How good MVSNets are at depth fusion Oleg Voynov, Alexander Safin, Savva Ignatyev and **Prof. Evgeny Burnaev** Skolkovo Institute of Science and Technology, Russia
13:15-13:30	V074-A	Combining compressed sensing and deep learning to create a more efficient machine vision setup **Mr. Alexander Birk**, Karsten Frenner, Wolfgang Osten University of Stuttgart, Germany
13:30-13:45	V066	Domain shift in computer vision models for MRI data analysis: An overview **Mrs. Ekaterina Kondrateva**, Marina Pominova, Elena Popova, Maksim Sharaev, Alexander Bernstein and Evgeny Burnaev Skolkovo Institute of Science and Technology, Russia
13:45-14:00	V076	DOME-T: Adversarial computer vision attack on deep learning models based on Tchebichef image moments **Mr. T. Maliamanis** and G.A. Papakostas International Hellenic University, Greece
14:00-14:15	V007	Scalable multi-view stereo using CMA-ES and distance transform-based depth map refinement **Mr. Nirmal S. Nair**, Madhu S. Nair University of Kerala, India
14:15-14:30	V048	Real-time vineyard trunk detection for a grapes harvesting robot via deep learning Eftichia Badeka, **Mr. Theofanis Kalampokas**, Eleni Vrochidou, Konstantinos Tziridis, George A. Papakostas, Theodore Pachidis, Vassilis G. Kaburlasos International Hellenic University, Greece

TECHNICAL PROGRAM

Session 16

Computer Photography and Image Processing

Friday, Nov. 6, 2020
15:00-16:30
ZOOM ID: 976 5819 4637

Chair:		
15:00-15:15	V097	Bipolar morphological U-Net for document binarization **Ms. Elena Limonova**, Dmitry Nikolaev and Vladimir Arlazarov Institute for Systems Analysis, FRC CSC RAS, Russia
15:15-15:30	V079	An instance segmentation framework for in-situ plankton taxa assessment **Dr. Aya Saad**, Sondre Bergum and Annette Stahl The Norwegian University of Science and Technology, Norway
15:30-15:45	V080	thermal image processing for feature extraction from encapsulated phase change materials Brian Whinery, Yuri Gulak, **Dr. Vedang Chauhan**, Jingzhou Zhao, Jingru Benner, Feng Ye Western New England University, US
15:45-16:00	V081	Application of fractional bio-inspired filter for salient color detection **Mr. Juan Anaya-Jaimes**, Angie Garcıá -Castro, J.A. Tenreiro-Machado and R.E. Gutiérrez-Carvajal Universidad Militar Nueva Granada, Colombia
16:00-16:15	V093	DNNs for multi-map semantic segmentation **Mr. Pavel Kurnikov** and Dmitry Sholomov Cognitive Robotics Ltd., Russia
16:30-17:00		**Closing Ceremony**

TECHNICAL PROGRAM

Session 17

Intelligent Image Analysis and Key Technologies

Friday, Nov. 6, 2020
15:00-16:30
ZOOM ID: 680 4475 3785

Chair: Assoc. Prof. Andrey Kuznetsov, Samara National Research University, Russia

15:00-15:15	V098	Fast and accurate mobile-aided screening system of moderate diabetic retinopathy **Dr. Yaroub Elloumi**, Manef Ben Mbarek, Rahma Boukadida, Mohamed Akil, Mohamed Hedi Bedoui University of Monastir, Tunisia
15:15-15:30	V058	Shape-aware generative adversarial networks for attribute transfer **Mr. Lei Luo**, William Hsu, Shangxian Wang Kansas State University, United States
15:30-15:45	V011	Robust technique for representative volume element identification in noisy microtomography images of porous materials based on pores morphology and their spatial distribution Grigoriev M., **Mr. Khafizov A.**, Kokhan V., Asadchikov V. FSRC "Crystallography and photonics" RAS, Russia
15:45-16:00	V064	Memory consumption reduction for identity document classification with local and global features combination **Ms. Natalya Skoryukina**, Vladimir Arlazarov, Artemiy Milovzorov Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, Russia
16:00-16:15	V063	Language of Gleam: Impressionism artwork automatic caption generation for people with visual impairments **Mr. Dongmyeong Lee**, Hyegyeong Hwang, Muhammad Shahid Jabbar, Jun-Dong Cho Sungkyunkwan University, South Korea
16:15-16:30	V4007	Feathers dataset for fine-grained visual categorization **Ms. Alina Belko**, Konstantin Dobratulin, Andrey Kuznetsov Samara University, Russia
16:30-17:00		**Closing Ceremony in Room A**

/04

PART IV

Presentation Abstracts

Monday, 2 Nov. 2020

Special Session I
Camera Based and Mobile Recognition

Virtual Room A:
976 5819 4637

Chair: Vladimir Arlazarov

Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, Russia

13:30-13:50 V051

Choosing the best image of the document owner's photograph in the video stream on the mobile device

Polevoy D.V., Aliev M.A., Nikolaev D.P.
Smart Engines Service LLC, Russia

One of the business tasks of personal documents recognition using mobile devices is to obtain a high quality image of document owner’s photograph. Such photographs are used to verify and identify the owner of the document. For example, in remote self-service systems, the image of a photo can be compared to a selfie.
When a document is captured with a mobile device camera in uncontrolled conditions, the photograph’s image quality varies greatly from frame to frame. In this paper, factors influencing the image quality of a photograph are considered: features of personal documents, capture and recognition processes. A method for choosing the best photograph image is proposed. The quality of the method is assessed on real data by the method of stochastic modeling.

13:50-14:10 V022

Tracking and Identification for Football Video Analysis using Deep Learning

Shreedhar Rangappa, Baihua Li, Ruiling Qian
Loughborough University, UK

The algorithm for 3D vector image reconstruction from a set of spectral tomographic projections collected with CT set-up completed with an optical element or elements inside the optical path behind the sample is proposed. The purpose of their placement into the optical path is to divide the integral polychromatic projection into a series of monochromatic projections, i.e., to get a multi-channel image. Understanding of the reconstruction results in the monochromatic case is beyond question, the relationship between the reconstructed spatial distribution of the linear attenuation coefficient and the discrete description of the elemental structure of the probed object is linear. In difference with monochromatic case the result of the reconstruction from polychromatic projections is a spatial distribution of the so-called effective or average attenuation coefficient, its connection to a discrete description of the elemental structure is nontrivial. However, if the distribution of the averaged coefficient is supplemented by distributions of linear coefficients for several energies, then it is possible to estimate of the local composition of the object. We present a model for the formation of spectral multi-channel projection based on crystal analyzer usage and describe the steps needed to solve the tomography inverse problem.

14:10-14:30 V068

A modification of a stopping method for text recognition in a video stream with best frame selection

Ilya Tolstov, Stanislav Martynov, Vera Farsobina, Konstantin Bulatov
Keldysh Institute of Applied Mathematics of Russian Academy of Sciences, Russia

One of the most important problem in constructing computer vision systems for embedded and mobile devices is offline recognition of text strings. In this paper, we analyze the problem of text strings recognition process in a video stream using best frame selection. This method allows to incorporate the information from multiple views of the same target object, thus increasing the overall extraction accuracy. A stopping method is proposed, which allows to make an automatic stopping decision, i.e. to terminate the process at the optimal time in order to maximize the responsiveness of the system. Experimental evaluation on open identity document datasets MIDV-500 and MIDV-2019 show that the proposed stopping rule allows to decrease mean error level of the text recognition results in comparison with a baseline approach which stops after a fixed amount of processed frames.

14:30-14:50 V044

About Viola-Jones image classifier structure in the problem of stamp detection in document images

Daniil P. Matalov, Sergey A. Usilin, Vladimir V. Arlazarov
Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, Russia

Monday, 2 Nov. 2020

In this paper we explore a set of modifications of the cascade structure of the Viola-Jones detector on the example of solving stamp detection problem. The experiments on the public “SPODS” dataset for various document attributes extraction problems with extremely limited training set are presented. The positive training set is augmented by applying various image processing algorithms relevant to the stamp model to an available in a single instance image for each stamp type. We describe and analyze such structures of the Viola and Jones classifiers as the original cascade structure, tree, soft cascade, and perform the training experiments. Experimental results show that each modification of the cascade structure of the classifier has its own advantages and disadvantages, and the choice of the Viola-Jones classifier design significantly affects the quality of solving object detection problem.

14:50-15:10 V089

Distance-based online pairs generation method for metric networks training

Ivan Kondrashev, Alexander Sheshkus, Vladimir V. Arlazarov
Bauman Moscow State Technical University, Russia

In this work, we consider the pairs generation algorithm based on the distances between elements in metric space. The right generation of training data is an actual issue, and its solution leads to better neural network learning. Understanding the properties of the source data, we can select pairs for training in such a way that the network will pay more attention to elements that are close in the metric space and have different classes. However, the problem arises when these properties are difficult to extract from the data and a more universal pairs generation method is needed. Our method generates pairs using the results of the network from previous iterations, in parallel with the training process itself. Thus, we do not need to evaluate the properties of elements ourselves, and we can use absolutely any data as learning objects. We demonstrate this approach using the example of Korean character recognition, and also compare it with other commonly used pair generation methods.

15:20-15:40 V092

Block convolutional layer for position dependent features calculation

Sergey A. Ilyuhin , Alexander V. Sheshkus, Vladimir L. Arlazarov
Smart Engines Service LLC; Moscow Institute of Physics and Technology, Russia

Image recognition includes problems where special features can be found only in a specific area of an image. This fact suggests us to apply different filters to different areas of input images. Convolutional networks have only fully-connected and locally-connected layers to make it. A Fully-connected layer erases the position factor for every output and a locally connected layer storage an enormous number of parameters. We need a layer that can apply different convolution kernels for different areas of an input image and not carry so many parameters as a locally-connected layer for high scale resolution images. This is why in this paper, we introduce a new type of convolutional layer - a block layer, and a way to construct a neural network using block convolutional layers to achieve better performance in the image classification problem. The influence of block layers on the quality of the neural network classifier is shown in this paper. We also provide a comparison with neural network architecture LeNet-5 as a baseline. The research was conducted on open datasets: MNIST, CIFAR-10, Fashion MNIST. The results of our research prove that this layer can increase the accuracy of neural network classifiers without increasing the number of operations for the neural network.

15:40-16:00 V104

Improved algorithm of ID card detection by a priori knowledge of the document aspect ratio

Daniil V. Tropin, Ivan A. Konovalenko, Natalya S. Skoryukina, Dmitry P. Nikolaev, Vladimir V. Arlazarov
MIPT (NRU), Russia

In this work, we consider a problem of quadrilateral document borders detection in images captured by a mobile device’s camera. State-of-the-art algorithms for the quadrilateral document borders detection are not designed for cases when one of the document borders is either completely out of the frame, obscured, or of low contrast. We propose the algorithm which correctly processes the image in such cases. It is built on the classical contour-based algorithm. We modify the latter using the document’s aspect ratio which is known a priori. We demonstrate that this modification reduces the number of incorrect detections by 34% on an open dataset MIDV-500.

16:00-16:20 V060

Image stacking versus per-frame results combination for video text recognition

Boris Savelyev, Konstantin Bulatov, Vladimir L. Arlazarov
Smart Engines, Moscow, Russia; Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, Russia

Monday, 2 Nov. 2020

In the field of document analysis and recognition using mobile devices for capturing, and the field of object recognition in a video stream, it is important to be able to combine the information received from different frames, since the quality of text recognition depends on the effectiveness of collecting the maximal amount of information about the target object. This paper examines and compares the effectiveness of two different combination approaches, namely pre-combination of images before recognition and the combination of recognition results. The combination methods are briefly described. The quality of the combined results obtained using different methods was measured and compared on the MIDV-500 dataset. The results show that the approach with a combination of text strings recognition results is more effective in comparison with the preliminary combination of images. It can be concluded that simple image stacking with projective alignment does not allow to achieve a comparable recognition results combination quality, and thus in order to include the information about per-frame changes of the text images more sophisticated image combination algorithms need to be employed.

16:20-16:40 V075

Generative approach for 1D barcode dataset population for mobile-based recognition

Milovzorov A.N, Bezmaternykh P.V., Arlazarov V.V.
National University of Science and Technology “MISIS”, Russia

Barcode recognition via the mobile device camera is an actual problem which occurs in various fields. To receive experimental baselines and provide methods comparison researchers require some carefully annotated data samples. For those purposes, some datasets have already been collected and published in the public domain. But they suffer from different disadvantages. In this paper, we present a novel challenging dataset designed for barcode reading quality evaluation. It is
populated using the generative approach with well-established image augmentation technique application. Among them, different geometrical transformations, kind of noises, brightness and lighting variance were used. The generative approach also allowed to populate ground-truth automatically and exclude errors introduced during the manual process. The dataset consists of training (41184 images) and validation parts (10296 images). It contains seven most popular symbologies: CODABAR, CODE-39, CODE-93, CODE-128, EAN-13, UPC-A, UPC-E. As an experimental baseline, the reading quality obtained with the open-source Zxing library is provided. The dataset with all the supplementary materials is available at ftp://smartengines.com/barcode.

16:40-17:00 V084

Chess Recognition using 3D Patterned Illumination Camera

Lars Brunner, Mario Salvator, Philipp Roebrock, Udo J. Birk
University of Applied Sciences of the Grisons, Switzerland

Computer Vision has been applied to augment traditional board games such as Chess for a number of reasons. While augmented reality enhances the gaming experience, the required additional hardware (e.g. head gear) is still not widely accepted in everyday leisure activities, and therefore, camera based methods have been developed to interface the computer with the real-life chess board. However, traditional 2D camera approaches suffer from ill-defined environmental conditions (lighting, viewing angle) and are therefore severely limited in their application. To answer this issue, we have incorporated a consumer-grade depth camera based on patterned illumination. We could show that in combination with traditional 2D color images, the recognition of chess pieces is made easier, which allows seamless integration of the real-life chess pieces with the computer program. Our method uses a fusion approach from depth and RGB camera data and is suitable for two distant players to play against each other, using two physical sets of chess.

Tuesday, 3 Nov. 2020

Special Session II
Advanced Imaging and Tomography

Virtual Room A:
976 5819 4637

Chair: Alessia Cedola — The Sapienza University of Rome, Rome unit Nanotec CNR, Italy

13:30-13:50 V016

Improving the quality of tomographic images by the methods of the computer-aided measuring systems

Chulichkov A.I., Molkov E.M., Prutjan S.I.
Lomonosov Moscow State University, Russia

The development of methods for improving the quality of tomographic images is an urgent task, as the presence of artifacts and insufficient sharpness of tomography result can cause erroneous decisions in medical diagnostics, when analyzing the structure of geological cores, etc. It is assumed that several artifacts arise due to inadequate a priori information involved in constructing the result of tomography, and a two-stage method is proposed for constructing an estimate of the distribution of the absorption coefficient of the sample. At the first stage, without using a priori information about the internal structure, the methods of the theory of computer-aided measuring systems construct an estimate of the absorption coefficient, which results in a blurred image of the internal structure of the object. At the second stage, the resolution of this image is increased by the method of moving averages with coefficients calculated from the condition of maximum accuracy for estimating the brightness of the central pixel of the window. Further, this estimate is refined from the natural conditions of the non-negativeness of the estimated brightness. Examples of the application of the method for assessing the structure of a child's tooth are given.

13:50-14:10 V021

Artifacts suppression in biomedical images using a guided filter

I. Bukreeva, A. Ingacheva, M. Fratini, A. Cedola, O. Junemann, E. Longo, F. Wilde, J. Moosmann, A. Buzmakov, Y. Krivonosov, D. Zolotov, S. Saveliev, V. Asadchikov and M. Chukalina
Nanotec-CNR, c/o Phys. Dep. at ‘Sapienza’ University, Italy

Despite significant progress in computer vision, pattern recognition, and image analysis, artifacts in imaging still hampers the progress in many scientific fields relying on the results of image analysis. We here present an advanced image-based artifacts suppression algorithm for high-resolution tomography. The algorithm is based on guided filtering of a reconstructed image mapped from the Cartesian to the polar coordinates space. This post-processing method efficiently reduces both ring- and radial streak artifacts in a reconstructed image. Radial streak artifacts can appear in tomography with an off-center rotation of a large object over 360 degrees used to increase the reconstruction field of view. We successfully applied the developed algorithm for improving x-ray phase-contrast images of human post-mortem pineal gland and olfactory bulbs.

14:10-14:30 V026

Blind CT images quality assessment of cupping artifacts

Anastasia Ingacheva, Marina Chukalina, Dmitry Nikolaev
Institute for Information Transmission Problems (Kharkevich Institute) RAS, Russia

In Computed tomography (CT) usage of common reconstruction algorithms to the projection data acquired with polychromatic probing radiation leads to the appearance of a cup-like distortion. CT image quality can be improved by adjusting the CT scanner or the reconstruction algorithm, but for this purpose assessment of cupping artifacts evaluation needs to be done. Existing assessment methods either rely on expert opinion or require an object binary mask, which can be unavailable. In this paper, we propose a method for blind assessment of cupping artifacts that do not require any prior information. The main idea of the proposed method is to evaluate the degree of change in intensity near automatically found edges of optically dense objects. We prove the applicability of the method on the collected dataset with cupping artifacts. The results show a monotonic dependency between the severity of cupping artifacts and the calculated with the proposed method value.

14:30-14:50 V024

Processing and Understanding of Images in Spectral Tomography

Marina Chukalina, Anastasiya Ingacheva, Alexey Buzmakov, Dmitry Nikolaev
FSRC «Crystallography and photonics» RAS, Russia

The algorithm for 3D vector image reconstruction from a set of spectral tomographic projections collected with CT set-up completed with an optical element or elements inside the optical path behind the sample is proposed. The purpose of their placement into the optical path is to divide the integral polychromatic projection into a series of monochromatic projections, i.e., to get a multi-channel image. Understanding of the reconstruction results in the monochromatic case is beyond

Tuesday, 3 Nov. 2020

question, the relationship between the reconstructed spatial distribution of the linear attenuation coefficient and the discrete description of the elemental structure of the probed object is linear. In difference with monochromatic case the result of the reconstruction from polychromatic projections is a spatial distribution of the so-called effective or average attenuation coefficient, its connection to a discrete description of the elemental structure is nontrivial. However, if the distribution of the averaged coefficient is supplemented by distributions of linear coefficients for several energies, then it is possible to estimate of the local composition of the object. We present a model for the formation of spectral multi-channel projection based on crystal analyzer usage and describe the steps needed to solve the tomography inverse problem.

14:50-15:10 V054

Modelling the Point Source Function for Collimators employed in Radio Guided Surgery and SPECT

Ya. L. Shabelnikova, S. I. Zaitsev, V. B. Obolenskiy
Institute of microelectronic technology problems and high purity materials RAS, Russia

For the multi-channel collimators employed in SPECT measurements and in gamma probes (for radio guided surgery) an approach to modeling the functioning of collimator is developed. This approach involves the calculation of point source function and the sensitivity field for the considering device. The effect of finite absorption in collimator material is taken into account. The results of modelling for confocal collimator are presented. The calculations were performed for the attenuation length corresponding to the device 3D-printed from metal and in the assumption of infinite absorption. It makes possible to estimate quantitatively the effect of gamma photons partial propagation through the matrix of collimator on the sensitivity and resolution for the specified material and the geometry of the device.

15:20-15:40 V100

Empirical analysis of the optimality of RSRE-based stopping rules for monitored reconstruction

Konstantin Bulatov, Arseniy Mukovozov, Vladimir V. Arlazarov
Smart Engines Service LLC, Russia

One of the challenges present in the field of tomographic imaging is the reduction of the radiation dose imparted to the object. Monitored reconstruction is one of the approaches to reduce the dose by means of dynamic stopping of the scanning process. In this paper, analysis was performed for the RSRE-based stopping rules for monitored reconstruction, using the previously analyzed RSRE-based reconstruction quality metrics, a widely-used PSNR metric, as well as quality metrics designed to mimic human perception, such as SSIM and ISSIM. It was shown that the unnormalized RSRE-based stopping rule performs better than the baseline for SSIM, ISSIM, and PSNR, and that for the latter the best result was achieved using the stopping rule designed for the reconstruction quality metric normalized to the Radon invariant. The stopping rules were compared with a synthetic a-posteriori “perfect” stopping rule and it was shown that the RSRE-based stopping rules closely approach the perfect stopping rule for the RSRE metric normalized on the Radon invariant, as well as for the ISSIM metric.

15:40-16:00 V108

CT Images GAN-based Augmentation with AdaIN for Lung Nodules Detection

Maksim Kryuchkov, Natalia Khanzhina, Ilya Osmakov and Pavel Ulyanov
ITMO University, Russia

In this work we made an attempt to improve 3D detection of pulmonary nodules on CT images using Conditional GANs extended with Adaptive Instance Normalization and combined Wasserstein Loss for data augmentation (DA). Nodule generating GAN model used for DA was built upon an open-source CT-GAN network which provides high and reproducible results in the nodule generation task. For the evaluation purpose we used DeepSEED model, which is a 3D end-to-end one-stage detector. We tested our approach on the LUNA16 dataset, the subset of LIDC-IDRI. The proposed model outperformed the baseline detection model trained on the original dataset by 3% of average sensitivity. The augmentation helped achieve a remarkable classification rate: 91% of sensitivity and 86% of specificity.

16:00-16:20 V4001

Lightweight Denoising Filtering Neural Network For FBP Algorithm

Andrei Yamaev, Marina Chukalina, Dmitry Nikolaev, Alexander Sheshkus, Alexey Chulichkov
MSU, Russia

In that paper, we a suggest lightweight filtering neural network, which implements the filtering stage in the Filtered Back-Projection algorithm (FBP), but good reconstruction results are achieved not

only in ideal data but also in noisy data, which a usual FBP algorithm cannot achieve. Thus, our neural network is not an only variation of Ramp filter, which is usually used then FBP algorithm, but also a denoising filter. The neural network architecture was inspired with the idea of the possibility of the Ramp filtering operation's approximation with sufficient accuracy. The efficiency of our network was shown on the synthetic data, which imitate tomographic projections collected with low exposition. In the generation of synthetic data, we have taken into account the quantum nature of X-ray radiation, exposition time of one frame, and non-linear detector response. The FBP reconstruction time with our neural network was 13 times faster than the time of reconstruction neural network from Learned Primal-Dual Reconstruction, and our reconstruction quality 0.906 by SSIM metric, which is enough to identify most significant objects.

16:20-16:40 V047

Iterative Reconstruction of Incomplete Tomography Data: Application Cases

Alexey Buzmakov, Denis Zolotov, Marina Chukalina, Anastasia Ingacheva, Victor Asadchikov, Dmirty Nikolaev, Yuri Krivonosov, Irina Dyachkova, Inna Bukreeva
FSRC "Crystallography and photonics" RAS, Russia

The method of Computed Tomography (CT) has progressed throughout the past decade with advances in CT apparatus and program parts that have resulted in an increasing number of CT applications. Today innovative CT X-ray detectors have high spatial resolution till a tenth or hundredth of a micron. However, its field of view is significantly limited. The object being scanned with a high resolution does not always completely enter in (covered by) the field of view of the detector. The collected projections data may be incomplete. The use of incomplete data in classical reconstruction methods leads to image quality loss. This paper provides a new advanced reconstruction method that demonstrates image quality improvements compared with classical methods when incomplete data collected. The method uses the hypothesis about the consistency of object description in sinogram space and reconstruction space. Input data for the algorithm proposed are incomplete data, and the output data are the reconstructed image and the confidence values for all pixels of the image (reconstruction reliability). A detailed description of the algorithm is presented. Its quality characteristics are based on Shepp-Logan phantom studies.

Wednesday, 4 Nov. 2020

Special Session III
Machine Vision for Autonomous Driven Cars under Harsh Environmental Conditions

Virtual Room A:
976 5819 4637

Chair: Wolfgang Osten — University of Stuttgart, Germany

09:30-09:50 V053

Local Path Planning Algorithm for Autonomous Vehicle Based on Multi-objective Trajectory Optimization in State Lattice

Ivan I. Kornev, Vladislav I. Kibalov, Oleg S. Shipitko
Institute for Information Transmission Problems – IITP RAS, Russia

The paper presents an algorithm for constructing a local path for a vehicle with nonholonomic kinematics of an automobile type. A local path is a sequence of transitions in the graph of possible maneuvers that minimizes a given cost function. The graph is constructed by duplicating along the global path pre-calculated in a curvilinear coordinate system set of kinematically feasible motion primitives. The use of pre-computed motion primitives significantly reduces the time of graph construction. The weight of each maneuver – the edge of the transition graph – is calculated as a weighted sum of costs based on several criteria. The specified cost function minimizes maneuvering and maintains a safe distance to static obstacles. The information about obstacles is extracted from an occupancy grid map. Dijkstra's algorithm is used to search a path in the weighted directed graph. The algorithm was tested on a dataset containing real road scenes. Each scene represents a given global path and a static environment model where a safe local path must be found. Local path search is performed in real-time. Experiments have shown that safe local paths have been found in all scenes where it was physically possible. At the same time, the obtained local paths were on average only on 1:3% longer than the given global paths which demonstrate the high applicability of the proposed algorithm.

09:50-10:10 V065

Deep car detection by fusing grayscale image and weighted upsampled LiDAR depth

Meisam Jamshidi Seikavandi, Kamal Nasrollahi, Thomas B. Moeslund
Khaje Nasir University of Technology, Iran

Recent advances have shown sensor-fusion's vital role in accurate detection, especially for advanced driver assistance systems. We introduce a novel procedure for depth upsampling and sensor-fusion that together lead to an improved detection performance, compared to state-of-the-art results for detecting cars. Upsampling is generally based on combining data from an image to compensate for the low resolution of a LiDAR (Light Detector and Ranging). This paper, on the other hand, presents a framework to obtain dense depth map solely from a single LiDAR point cloud that makes it possible to use just one deep network for both LiDAR and image modalities. The produced full-depth map is added to the grayscale version of the image to produce a two-channel input for a deep neural network. The simple preprocessing structure is efficiently competent in filing cars' shapes, which helps the fusion framework to outperforms the state-of-the-art on the KITTI object detection for the Car class.

10:10-10:30 V061

An approach to road scene text recognition with per-frame accumulation and dynamic stopping decision

Konstantin Bulatov, Nadezhda Fedotova, Vladimir V. Arlazarov
Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Russia

Camera-based road scene analysis is an important task for building driving assistance systems and autonomous vehicles. An crucial component of road scene analysis is detection, tracking, and recognition of text object. In this paper, we consider the recognition of road scene text objects in sequences of video frames, and propose an approach to per-frame recognition results accumulation with a dynamic stopping decision. Experimental evaluation on an open dataset RoadText-1K showed that the proposed approach allows to achieve mean lower recognition error for the same mean number of processed frames, and significantly reduce the number of text objects which have to be recognized in each frame, thus relieving the load on the computational unit.

11:40-11:00 V069

A Machine Vision Based Smart Conveyor System

Mengchao Zhang, Vedang Chauhan and Manshan Zhang
Western New England University, US

Wednesday, 4 Nov. 2020

Material flow detection on conveyor based on machine vision is the research topic of this paper. A belt conveyor system equipped with a camera and micro-controller is used as the test apparatus. The purpose of this experiment is to obtain the quantity of material on the conveyor belt using machine vision and then develop an intelligent speed adjustment system for belt conveyor according to the quantity, so as to avoid waste of energy and reduce the wear of the conveyor. Three image processing algorithms that developed, applied and compared were: 1) Background Subtraction; 2) Canny edge detection and morphological operations; 3) Particle analysis using. It is observed that all three methods perform well for material detection on the conveyor belt. However, the particle analysis method resulted in higher reliability and accuracy with faster processing speed. The research provides new developmental ideas for intelligent conveyor systems.

11:00-11:20 V091

A Target-free Calibration Method for Automotive Augmented Reality Head-Up Displays

Xiang Gao, Kai Wu, Marc Necker, Wilhelm Stork, Adnane Jadid, Gudrun Klinker
Mercedes-Benz AG & Karlsruhe Institute of Technology, Germany

In this paper, we introduce a newly developed target-free calibration method for automotive augmented reality head-up displays (AR-HUDs), which can be performed fully automatically using a smartphone camera. Our method requires no calibration target to be set up in front of the vehicle. Instead, it utilizes feature points of the environment, which makes it robust against misplaced targets and allows for an easy deployment, i.e. in garages. Under the pinhole model assumption, we decouple the perspective projection matrix into three parts: intrinsic matrix, relative pose between the vehicle's 3D sensor and the smartphone camera, and then rotation between the camera space and the HUD field of view (HUD-FOV). Based on the epipolar constraint, we acquire the relative pose. The determination of intrinsic and rotation matrices is also accomplished without any pre-designed calibration target. The calibration itself takes less than 5 minutes for an eye box with 9 different training viewpoints. With our new approach, we achieve a competitive average reprojection error of 6.7mm at a distance of 7.5 m, which is comparable to the previous work that applied targets.

11:20-11:40 V094

Fast Hough Transform-Based Road Markings Detection For Autonomous Vehicle

Oleg Shipitko, Ekaterina Panfilova and Irina Kunina
Smart Engines Service LLC, Russia

The paper presents an algorithm for road markings detection in the image. The road markings are approximated by with a restricted maximum curvature angle. To detect a marking segments an image is processed by a sliding window and for each window position, a straight line is detected by calculating Fast Hough Transform (FHT). Further, detected segments are grouped by relative position. Segments groups are then approximated by polylines. The algorithm was tested on real data collected from the front-looking camera of the autonomous vehicle driving at the experimental area "Kalibr" (Moscow). The road marking dataset used to evaluate the algorithm is publicly available at ftp://vis.iitp.ru/road markup dataset/. The precision of road markings detector was evaluated as 43%, and the recall as 73% which is sufficient for the autonomous vehicle precise positioning as demonstrated in [1].

11:40-12:00 V059

LIDAR-based parking spot search algorithm

Dinir Imameev, Aufar Zakiev, Tatyana Tsoy, Yang Bai, Mikhail Svinin, Evgeni Magid
Kazan Federal University, Russia

Autonomous driving considers issues related to a car driving in different real world situations. This work addresses a parking task and describes a new LIDAR-based parking spot search algorithm. The proposed approach was successfully validated in virtual experiments within the Gazebo simulator in a parking area with a perpendicular parking setup. HDBScan, OPTICS, and Gaussian Mixture clustering methods were compared for LIDAR data clustering in the parking spot search task, and the HDBScan clustering demonstrated best prediction and performance results.

Wednesday, 4 Nov. 2020

Special Session IV
Computer Optics Journal

Virtual Room A:
976 5819 4637

Chair: Vladislav Sergeev — Samara National Research University, Russia

13:00-13:20 V5001-A

Algorithm for post-processing of tomographic images to calculate the dimension-geometric features of porous structures

M. V. Chukalina, A. V. Khafizov, V. V. Kokhan, A. V. Buzmakov, R. A. Senin, V. I. Uvarov, M. V. Grigorev
Smart Engines LLC, Institute for Information Transmission Problems RAS, Russia

Algorithm for post-processing of the grayscale 3D tomographic (CT) images of porous structures with the automatic selection of filtering parameters is proposed. The determination of parameters is carried out on a representative part of the analyzed image. A criterion for the search for optimal filtration parameters, based on the calculation of voxels of the so-called "levitating stones" is described. The stages of CT image filtering and its binarization are performed sequentially. Bilateral and anisotropic diffuse filtering are implemented; the Otsu method for unbalanced classes is chosen for binarization. Verification of the proposed algorithm was carried out on model data. To create model porous structures, we used our image generator which implements the function of anisotropic porous structures generation. The results of the post-processing of real CT images containing noise and reconstruction artifacts by the proposed method are discussed.

13:20-13:40 V5002-A

Weighted combination of per-frame recognition results for text recognition in a video stream

Olga Petrova, Konstantin Bulatov, Vladimir V. Arlazarov, Vladimir L. Arlazarov
Research Center “Computer Science and Control” of Russian Academy of Sciences, Russia

The use scope for document recognition has extended, and as a result, the recognition technologies without specialized equipment have become more relevant. Among such technologies, document recognition using mobile devices is of interest. However, it is not always possible to ensure controlled capturing conditions and, consequentially, high quality of input images. Unlike specialized scanners, mobile cameras allow using a video stream as an input, thus obtaining several images of a recognized object, captured with various characteristics. In such case, a problem of combining the information from multiple input frames arises. In this paper, we pro-pose a weighing model for the process of combining the per-frame recognition results, two approaches to the weighted combination of the text recognition results, and two weighing criteria. The effectiveness of the proposed approaches is tested using datasets of identity documents captured with a mobile device camera in different conditions, including perspective distortion of the document image and low lighting conditions. Experimental results show that the weighting combination can improve the text recognition result quality in the video stream, and the per-character weighting method with input image focus estimation as a base criterion allows to achieve the best results on the analyzed datasets.

13:40-14:00 V5003-A

A generalization of Otsu's Method for linear separation of two unbalanced classes in document image binarization

E.I. Ershov, S.A. Korchagin, V.V. Kokhan, P.V. Bezmaternykh
Smart Engines LLC, Institute for Information Transmission Problems RAS, Russia

The classical Otsu’s method is a common tool in document image binarization. Often, two classes, text and background, are unbalanced, thus the assumption of the classical Otsu’s method is not met. In this work, we considered the unbalanced pixel classes of background and text. We experimentally demonstrated that the employment of the criterion, which takes into account the unbalance of the classes, allows for a higher binarization accuracy. We described the generalization of the criteria for a two-parametric model. For the latter, the algorithm for the optimal linear separation search via fast linear clustering is proposed. We also demonstrated that the two-parametric model with the proposed separation allows for the image binarization accuracy increase for the documents with complex background or spots.

14:00-14:20 V5004-A

Algorithm for choosing the best frame in a video stream in the task of identity document recognition

Aliev M.A., Kunina I.A., Kazbekov A.V., Arlazarov V.L.
Smart Engines Service LLC, Moscow, Russia

Wednesday, 4 Nov. 2020

During the document recognition in a video stream using a mobile device camera, the image quality of the document varies greatly from frame to frame. Sometimes the recognition system is required not only to recognize all the specified attributes of the document, but also to select final document image of the best quality. This is necessary, for example, for archiving or providing various services; in some countries it can be required by law. In this case, the recognition system needs to assess the quality of frames in the video stream and choose the “best” frame. In this paper we consider the solution to such a problem where the “best” frame means the presence of all the specified attributes in a readable form on the document image. The method was set up on a private dataset, and then tested on documents from the open MIDV-2019 dataset. A practically applicable result was obtained for use in recognition systems.

14:20-14:40 V5005-A

Optimal affine image normalization approach for optical character recognition

I.A. Konovalenko, V.V. Kokhan, D.P. Nikolaev
Smart Engines LLC, Institute for Information Transmission Problems RAS, Russia

Optical character recognition (OCR) on images captured from arbitrary angles requires preliminary normalization, i.e. geometric transformation resulting in an image as if it was captured from the angle suitable for OCR. In most cases, a surface containing characters can be considered flat, and a pinhole model can be adopted for a camera. Thus, in theory, the normalization should be projective. Usually, the camera optical axis is approximately perpendicular to the document surface, so projective normalization can be replaced with affine one without significant loss of accuracy. The affine image transformation is performed significantly faster than the projective normalization, which is important for OCR on mobile devices. In this work, we propose a fast approach for image normalization. It utilizes the affine normalization instead of projective if there is no significant loss of accuracy. The approach is based on a proposed criterion for the normalization accuracy: root mean square (RMS) coordinate discrepancies over the region of interest (ROI). The problem of optimal affine normalization according to this criterion is considered. We have established that this unconstrained optimization is quadratic and can be reduced to the problem of fractional quadratic functions integration over the ROI. The latter was solved analytically in the case of OCR where the ROI consists of rectangles. The proposed approach is generalized for various cases when instead of affine transform its special cases are used: scaling, translation, shearing and their superposition, which allows for the additional acceleration of image normalization.

Wednesday, 4 Nov. 2020

Track
Color Computer Vision

Virtual Room A:
976 5819 4637

Chair: Egor Ershov

Vision Systems Lab., Institute of Information Transmission Problems, Russian Academy of Science, Russia

15:00-15:20 V082

Consensus-driven Illuminant Estimation with GANs

Marco Buzzelli, Riccardo Riva, Simone Bianco, Raimondo Schettini
University of Milano – Bicocca, Italy

We present a method for illuminant estimation that exploits a generative adversarial network architecture to generate a spatially-varying illuminant map. This map is then transformed by consensus into a global illuminant estimation, in the form of a single RGB triplet. To this end, different consensus strategies are designed and compared in this paper. The best solution won second place in the 2nd International Illumination Estimation Challenge, specifically for the indoor track.

15:20-15:40 V077

Robust White Balance Estimation Using Joint Attention and Angular Loss Optimization

Zhihao Li, Zhan Ma
Nanjing University, China

White balance estimation (WBE) is one of the most fundamental and crucial steps in modern Image Signal Processor (ISP). Recent years have witnessed the advancements of deep-learning based WBE. However, existing models were mostly trained on individual datasets with limited samples captured using various camera sensors, making it hard for model generalization. In this paper, we propose a novel Channel-Attention optimized U-net model, in which an angular loss is embedded, to accurately estimate the white balance. We demonstrate our approach on recently released large-scale dataset “Cube Plus” captured using the same camera sensor, offering state-of-the-art performance.

15:40-16:00 V072

SDE-AWB: a Generic Solution for 2nd International Illumination Estimation Challenge

Yanlin Qian, Sibo Feng, Kang Qian, Miaofeng Wang
Tampere University, Finland

We propose a neural network-based solution for three different tracks of 2nd International Illumination Estimation Challenge (chromaticity.iitp.ru). Our method is built on pre-trained Squeeze-Net backbone, differential 2D chroma histogram layer and a shallow MLP utilizing Exif information. By combining semantic feature, color feature and Exif metadata, the resulting method – SDE-AWB – obtains 1st place in both indoor and two-illuminant tracks and 2nd place in general track.

16:00-16:20 V057

Color correction of the document owner's photograph image during recognition on mobile device

Dmitry Polevoy, Ekaterina Panfilova, Egor Ershov, Dmitry Nikolaev
Smart Engines Service LLC Moscow, Russia; V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia; The Institute for Information Transmission Problems of Russian Academy of Sciences, Moscow, Russia

The growing popularity of mobile services increases the risks of financial and other losses from fraudulent user actions. To reduce the number of illegal actions and comply with the law when using mobile services, it is often scheme that user presents it’s identity document. In case of remote access via a mobile device, this means receiving and analyzing a video of a document image. One of the criteria for the authenticity of the captured security document is the presence of security optically variable devices (kinegrams, holograms). Reliable determination of the presence or absence of such security elements from video stream frames is greatly complicated by changes in color between frames.
The paper discusses the possibility of using a priori information about the monochrome photograph of the document owner to compensate changes in color between frames. Color distributions are investigated on the example of black-and-white photographs. A new method for automatic white balance correction is proposed. The results of the method are tested on real data obtained with a mobile device.

Wednesday, 4 Nov. 2020

16:30-16:50 V006

A Low-Complexity yet Accurate Calibration Method for Automotive Augmented Reality Head-Up Displays

Xiang Gao, Marc Necker, Wilhelm Stork
Mercedes-Benz AG & Karlsruhe Institute of Technology, Germany

Automotive augmented reality head-up displays (AR-HUDs) superimpose driving related information with the real world in the direct sight of the driver. A key prerequisite for an immersive AR experience is a highly precise calibration. State-of-the-art methods require large targets and a lot of space in front of the vehicle, or special complex equipment, which is inconvenient in both factories and workshops. In this paper, we propose a low-complexity yet accurate calibration method using only a small sheet of patterned paper as the target, which is laid directly on the windscreen. The full field of view (FOV) can be calibrated, with the optical distortion corrected by extracted warping maps. The changing views of drivers are considered by interpolating both projection parameters and distortion models. The angular reprojection error falls within 0.04°, while the run-time is limited up to 1 minute per viewpoint. Our method shows high applicability in the automotive industry because of both reduced target complexity and competitive reprojection errors. Moreover, due to the reduced effort and simplified equipment, our method opens a way for customers to recalibrate their AR-HUDs themselves.

16:50-17:10 V020

Camera Auto-calibration for Complex Scenes

Anas Ali and Pavel Smrz
Brno University of Technology, Czech

In this paper, we propose a novel method for automatic camera calibration based on pedestrians' observations. Our proposed method is capable of estimating calibration parameters for complex scenes having more than one ground plane. Unlike existing methods that require time-consuming optimization step, our method uses real-time re-estimation step based on fuzzy logic while relaxing the assumption on the number of ground planes in the scene. Furthermore, we propose a dominant ground plane detection step for better calibration parameter estimation on complex scenes. To evaluate our proposed method, we run comprehensive testing using 5 different datasets covering varieties of calibration parameters and scene properties, we also conduct tests on a synthetic dataset for more detailed analysis. Test results show that our proposed method outperforms existing state-of-the-art methods in both performance accuracy and time complexity.

17:10-17:30 V045

A Dynamic Programming Inspired Outlier Rejection Algorithm for Image Mosaicing Problem

Christopher Smith and Semih Dinc
Auburn University at Montgomery, USA

Image mosaicing is a challenging problem when there is one (or more) corrupt image(s) in the input sequence. Since the transformation of later images rely on the previous transformation calculations, one miscalculation error caused by the corrupt image propagates to other image transformations, which fails the whole process. It is not a trivial task to detect and remove these corrupt images, which we call “outliers”. In this paper, we propose a dynamic programming inspired outlier rejection algorithm to identify and remove the corrupt image(s) from the sequence. Our approach stores the previously calculated transformation matrices in a 2D array and determines the validity of composite transformations based on a decision criterion. This criterion identifies the transformations coming from corrupt images by comparing the direct registration of images and the composite transformations from existing matrices in the array. We have performed experiments on both synthetic and real datasets. Visual and numerical results show that the proposed algorithm is an efficient tool for detecting and rejecting the outlier image from the mosaic image.

Thursday, 5 Nov. 2020

Session 6
Artificial Intelligence and Intelligent Computing

Virtual Room A:
976 5819 4637

Chair: Johan DEBAYLE — Ecole Nationale Supérieure des Mines de Saint-Etienne, France

09:30-09:45 V017

Model Selection for Support-Vector Machines through Metaheuristic Optimization Algorithms

Oumeima Ghnimi, Sofiane Kharbech, Akram Belazi, and Ammar Bouallegue
unis El Manar University, Tunisia / University of Lille, France

A machine learning algorithm aims at designing a mathematical model based on a given training data set. Generally, the built model has a set of parameters that need to be adjusted. Since the performance of a given model depends on its settings, the parameters have to be carefully chosen through a fine-tuning step. A good model selection not only boosts performance but also allows a well-generalized model, i.e., a model that works sound on unseen data. In this paper, we assess the effectiveness of some metaheuristic optimization algorithms for support-vector machines (SVM) model selection. Computer simulations show that optimization algorithms that overall outperforms other algorithms using benchmark functions can be, further, definitely used for an efficient SVM model selection for classification. Thus, we show that Teaching–Learning-Based Optimization algorithm is faster and also enables the most accurate classification, even against other proposed methods in the literature for SVM model selection.

09:45-10:00 V030

A semi-supervised learning approach for CBIR systems with relevance feedback

Marco Brighi, Annalisa Franco, Dario Maio
University of Bologna, Italy

WSemi-supervised learning techniques are gaining importance in the scenario of constantly growing data collections. CBIR systems must be able to autonomously analyze the patterns available, to fully exploit unlabeled data with the final objective of identifying an optimal representation space where data belonging to the same semantic class are close to each other. In this work we propose to adopt relevance feedback as a mean of collecting information about the semantic classes perceived by the user and to exploit this information for a long-term learning process where a more effective feature space can be obtained by a proper metric learning technique and class labels can be automatically assigned to unlabeled patterns. The process can iterate as new data become available thus providing a tool for successfully managing new incoming data. The experimental results will confirm the advantages of the proposed learning approach.

10:00-10:15 V042

Deep Convolutional Neural Network Based Autonomous Drone Navigation

Karim Amer, Mohamed Samy, Mahmoud Shaker, and Mohamed Elhelw
Center for Informatics Science, Nile University, Giza, Egypt

This paper presents a novel approach for aerial drone autonomous navigation along predetermined paths using only visual input form an onboard camera and without reliance on a Global Positioning System (GPS). It is based on using a deep Convolutional Neural Network (CNN) combined with a regressor to output the drone steering commands. Furthermore, multiple auxiliary navigation paths that form a ‘navigation envelope’ are used for data augmentation to make the system adaptable to real-life deployment scenarios. The approach is suitable for automating drone navigation in applications that exhibit regular trips or visits to same locations such as environmental and desertification monitoring, parcel/aid delivery and drone-based wireless internet delivery. In this case, the proposed algorithm replaces human operators, enhances accuracy of GPS-based map navigation, alleviates problems related to GPS-spoofing and enables navigation in GPS-denied environments. Our system is tested in two scenarios using the Unreal Engine-based AirSim [32] plugin for drone simulation with promising results of average cross track distance less than 1.4 meters and mean waypoints minimum distance of less than 1 meter.

10:15-10:30 V070

Spatiotemporal feature based Convolutional Neural Network for violence detection

Amira Ben Mabrouk, Ezzeddine Zagrouba
Higher Institute of Computer Science, University of Tunis EL Manar, Tunisia

Thursday, 5 Nov. 2020

Automatic violence detection is one of the most interesting tasks in video surveillance applications. It is essential to ensure human safety by preventing fatal accidents. In this paper, we propose a spatiotemporal feature based on convolutional neural network and optical flow information for violence detection. Our descriptor, named distribution-based CNN feature (DCNN), estimates first the joint distribution of the optical flow magnitude and orientation around STIP points in the aim of modeling the motion structure. Then, it extracts deep features using the pretrained ResNet50 network. Our DCNN feature is finally fed into the SVM classifier for training. We evaluate the proposed descriptor on two challenging benchmark datasets designed for violence detection in both uncrowded and crowded scenes. For the two datasets, our DCNN feature shows significant improvement compared to the state-of-the-art descriptors.

10:30-10:45 V088

Almost Indirect 8-bit Convolution for QNNs

Anton Trusov, Elena Limonova, Sergey Usilin
Moscow Institute of Physics and Technology, Russia

The implementations of the convolution operation in neural networks are usually based on convolution-to-GeMM (General Matrix Multiplication) transformation. However, this transformation requires a big intermediate buffer (called im2col or im2row), and its initialization is both memory and time-consuming. To overcome this problem, one may use the Indirect Convolution Algorithm. This algorithm replaces the im2row buffer with a much smaller buffer of pointers, called indirection buffer. However, it limits our flexibility in the choice of multiplication micro-kernel, making matrix multiplication slightly less efficient than in the classical GeMM algorithm. To overcome this problem, we propose the Almost Indirect Convolution Algorithm, which initializes small specifically ordered block of values, which is used in matrix multiplication, via indirection buffer, the same way GeMM Algorithms initializes one block from im2row buffer. Our approach allows us to combine computational efficiency and flexibility in shape of GeMM micro-kernels with a small memory footprint of the Indirect Convolution Algorithm. Experiments with convolutions of 8-bit matrices on ARM processors show that our convolution works 14-24% faster than Indirect for a small number of channels and 10-20% faster than classical GeMM-based. This proves that it is perfectly suitable for computing inference of 8-bit quantized networks on mobile devices.

11:00-11:15 V095

DVAE-SR: Denoiser Variational auto-encoder and super-resolution to counter adversarial attacks

Omar Dardour, Mourad Zaied, Petia Radeva
RTIM-Lab, Tunisia

Recently, adversarial examples become one of the most dangerous risks in deep learning, which affects applications of real world such as robotics, cyber-security and computer vision. In image classification, adversarial attacks showed the ability to fool classifiers with small imperceptible perturbations added to the input. In this paper, we present an efficient defense mechanism, we call DVAE-SR that combine variational autoencoder and super-resolution to eliminate adversarial perturbation from image input before feeding it to the CNN classifier. The DVAE-SR can successfully defend against both white-box and black-box attacks without retraining CNN classifier and it recovers better accuracy than Defense-GAN and Defense-VAE..

11:15-11:30 V101

Abstractive Meeting Summarization based on an Attentional Neural Model

Nouha Dammak, Yassine BenAyed
University of Sousse, Tunisia

Through the ages, in all nations, at all times, people spend a lot of their time on discussing new and important issues either on meetings or in conferences. With the evolution and the abundance of Automatic Speech Recognition (ASR) frameworks, automatic transcripts and even automatic meeting summarization are getting more and more interesting. Recently, automatic summarization faces deeper progresses on speech summarization. Neural models had been introduced to tackle with many difficulties of abstractive summarization. Our contribution in this paper focuses on these weaknesses of neural abstractive meeting summarization and suggests an encoder-decoder model based on an attentional algorithm on the decoding sequence. We proposed a deep encoder-decoder model based on attention mechanism (DEDA) for ASR transcripts. Experiments on the AMI Dataset demonstrates that our proposed method ensured competitive results with the state of the art even on extractive or abstractive models. The experimental analyses also put the stress on the performance of the summarized utterances as well as the reduction of the occurrence repetition in summaries.

Thursday, 5 Nov. 2020

Session 7
Software and Information Engineering

Virtual Room B:
680 4476 3785

Chair: Hiroshi Hosobe — Hosei University, Japan

09:30-09:45 SE020

Continuous Development and Testing of Access and Usage Control: A Systematic Literature Review

Said Daoudagh, Francesca Lonetti, Eda Marchetti
ISTI-CNR, Italy

Development and testing of access/usage control systems is a growing research area. With new trends in software development such as DevOps, the development of access/usage control also has to evolve. Objective: The main aim of this paper is to provide an overview of research proposals in the area of continuous development and testing of access and usage control systems. Method: The paper uses a Systematic Literature Review as a research method to define the research questions and answer them following a systematic approach. With the specified search string, 210 studies were retrieved. After applying the inclusion and exclusion criteria in two phases, a final set of 20 primary studies was selected for this review. Results: Results show that primary studies are mostly published in security venues followed by software engineering venues. Furthermore, most of the studies are based on the standard XACML access control language. In addition, a significant portion of the proposals for development and testing is automated with test assessment and generation the most targeted areas. Some general guidelines for leveraging continuous developing and testing of the usage and access control systems inside the DevOps process are also provided.

09:45-10:00 SE030

Development Frameworks for Microservice-based Applications: Evaluation and Comparison

Hai Dinh-Tuan, Maria Mora-Martinez, Felix Beierle, Sandro Rodriguez Garzon
Technische Universität Berlin, Berlin, Germany

The microservice architectural style has gained much attention from both academia and industry recently as a novel way to design, develop, and deploy cloud-native applications.This concept encourages the decomposition of a monolith into multiple independently deployable units. A typical microservices-based application is formed of two service types: functional services, which provide the core business logic, and infrastructure services, which provide essential functionalities for a microservices ecosystem. To improve developers' productivity, many software frameworks have been developed to provide those reusable infrastructure services, allowing programmers to focus on implementing microservices in arbitrary ways. In this work, we made use of four open source frameworks to develop a cloud-based application in order to compare and evaluate their usability and practicability. While all selected frameworks promote asynchronous microservice design in general, there are differences in the ways each implements services. This leads to interoperability issues, such as message topic naming convention. Additionally, a key finding is the long startup times of JVM-based services that might reduce application's resiliency and portability. Some other advantages come directly from the programming language, such as the ability of Go to generate native binary executables, which results in very small and compact Docker images (up to 78% smaller compared to other languages).

10:00-10:15 SE031

The Link Between Transformational and Servant Leadership in Devops-Oriented Organisations

Krikor Maroukian, Stephen R. Gulliver
Microsoft, Greece

DevOps is a set of agile and lean practices and principles in the context of software product development aiming to decrease mean time-to-market and mean time-to-recover-from-failure through a shift in organizational mindset-skillset-toolset. There is literature to suggest that adopting DevOps has been challenging in practice and that a particular leadership style is necessary to lead DevOps adoption. There are studies to suggest that DevOps leadership is mainly related to transformational leadership characteristics. In this research, a mixed methods approach is used. Initially, semi-structured interviews are conducted with 30 EMEA (Europe, Middle-East and Africa) agile and lean practitioners holding more than 10 years of practitioner experience (81%) from the private and public sectors. The contribution also includes an analysis and evaluation of a survey completed by 250 participants of which 93% works in Europe and Middle East and 76% has held previous leadership positions. By looking to recent literature we identified agile, lean and DevOps practices and principles. In addition, we identify benefits and inhibitors to DevOps adoption and its leadership. Our results suggest that deep rooted organizational culture and lack of DevOps definition clarity are usually considered impediments to DevOps adoption followed by poor communication and collaboration. Our results also show that certain DevOps adoption leadership characteristics are relevant to transformational leadership and servant leadership. The research results also indicate that the DevOps adoption leadership role is linked to certain metrics.

Thursday, 5 Nov. 2020

10:15-10:30 SE024

Maintainability Metrics for Android Applications in Kotlin: An Evaluation of Tools

Lisa-Marie Andrä, Bernhard Taufner, Sigrid Schefer-Wenzl, Igor Miladinovic
FH Campus Vienna, Austria

Over the last decade, Android has become the largest mobile platform, providing a huge and ever increasing amount of apps in the Google Play store. Since 2017, a growing number of Android apps have been written using the Kotlin programming language. Maintainability is an important attribute of software quality. It determines how well maintenance and code modification can be done. This is particularly important for Android apps that are frequently updated to meet changing user requirements.
Several metrics can provide information about the maintainability of a system. They can be calculated with different analysis tools. This paper analyzes the maintainability of Kotlin applications. In particular, we investigate which metrics are relevant for the maintainability of Kotlin applications and to what extent analysis tools support the calculation of these metrics. Our results show that these tools are very limited with respect to the supported programming languages. Especially for Android applications written in Kotlin, the majority of maintainability metrics cannot be calculated with current tools.

10:30-10:45 SE029

Testing Event-Driven Programs in Processing

Hiroshi Hosobe
Hosei University, Japan

Event-driven programming is a paradigm that is widely used in many fields. Processing is a set of programming languages and environments specialized in event-driven programming for interactive graphical applications. It provides only low-level event-handling functions, which imposes difficulty on novice programmers in programming complex behaviors. This paper proposes a method for unit-testing event-driven Processing programs. It allows writing testable Processing programs and test programs in Java. To demonstrate how it works, this paper presents case studies on testing whether mouse and key events are correctly handled.

11:45-11:00 SE009

Communication Challenges in Agile Teams from The Communication Theory Prospective

Sofiia Yermolaieva
Innopolis University, Russia

People are the only constant component of any software project. Diversity of cultures, behavioral patterns, and life scripts leads to different understanding and meaning of interactions and information exchange among team members. As processes that follow Agile methodology approaches consider individuals and interaction as the highest priority, it was decided to analyze communication issues in teams who work by such processes. The work presents an overview of obstacles for effective communication based on literature review. The survey among 53 respondents was conducted to evaluate the most common communication challenges which Agile teams face in their work. These challenges are further mapped with communication theories. On the basis of the mapping, we developed recommendations for overcoming problems described in each of the theories. The results of the work could be used by people with managerial roles in Agile teams to learn more about the root causes of ineffective communication among team members. The solutions and recommendations could be used to overcome these issues.

11:00-11:15 SE003

A Dashboard to Support Decision-Making Processes in Learning Ecosystems: A Metamodel Integration

Francisco José-García Peñalvo, Andrea Vázquez-Ingelmo, Alicia García-Holgado, Roberto Therón
University of Salamanca, Spain

There are software solutions to solve most of the problems related to information management in any company or institutions, but still, there is a problem for transforming information into knowledge. Technological ecosystems emerge as a solution to combine existing tools and human resources to solve different problems of knowledge management. In particular, when the ecosystem is focused on learning processes associated with knowledge are named learning ecosystems. The learning ecosystem metamodel defined in previous works solves several problems related to the definition and implementation of these solutions. However, there are still challenges associated with improving the analysis and visualization of information as a way to discover knowledge and support decision making processes. On the other hand, there is a metamodel proposal to define customized dashboards for supporting decision-making processes. This proposal aims to integrate both metamodels as a way to improve the definition of learning ecosystems.

Thursday, 5 Nov. 2020

11:15-11:30 SE010

Empirical Study on Commonly Used Combinations of Estimation Techniques in Software Development Planning

Mr. Arman Kialbekov
University, Russia

Precise estimation of time and effort in the lifecycle of a project plays an important role in delivering a final product on time within an established budget, and with expected quality. Inappropriate estimation of effort may lead to overestimation or underestimation of required resources. As people remain to be a constant part of the software development process they should also be considered as an influencing factor in software development estimation process. This work presents the most popular estimation techniques that are used by the survey sample. The questionnaire was completed by 51 participants with industrial experience in Innopolis. The goal was to find differences between roles in software development teams and the techniques that are used to estimate efforts. The result of the survey used to derive main groups of the respondents and map them to estimation techniques. These mappings were analyzed and compared between two main groups to find tendencies, differences, and similarities. The study is considered to help understand patterns of effort estimation among software engineers and revising existing estimation techniques.

Thursday, 5 Nov. 2020

Session 8
Target Detection

Virtual Room A:
976 5819 4637

Chair:

13:00-13:15 V003

Violent scenes detection based on connected component analysis

Samira Labbadi, Mariem Gnouma, Ridha EJbali and Mourad Zaied
Research Team in Intelligent Machines, Tunisia

Violent action represents a threat to public security, thus intelligent violence detection became one of the important and challenging topics in video surveillance scenarios for this reason there is a growing appeal of video-surveillance systems. Hence, it's mandatory for the detection of violent or abnormal activities to avert any casualties which could cause any damages. Distinctly, in this paper, it is possible to create a network to learn spatial-temporal information on all subjects of violence rather than going through each concept separately.
In order to construct a new concept for violence detection system, we rely on a strategy of a dynamic frame skipping to reduce the complexity of calculation. However, following the regions of interest in the frame, the overall complexity of the calculation is decreased. Withal, the History of Binary Motion Image for n successive images is used for features extraction to facilitate to model the human behaviors. Then, the biggest regions of interest are extracted in order to find the maximum component represented violence action. Finally, deep neural networks involve three stacked Auto-encoders and a Softmax are adopted as an exterior layer for classification.

13:15-13:30 V010

Fatigue detection based on non-contact respiratory detection

Xing Chen, Lumei Su, Min Xu, Bo Deng, Yuanxuan Zhu
Xiamen University of Technology, China

Most of current methods for fatigue detection are based on fatigue features extracted from face or head. This paper proposes a novel fatigue detection mode based on non-contact respiratory detection. This paper proposes to use the respiratory rate feature for fatigue detection. Since the appearance of the respiratory rate fatigue feature will be earlier than the facial fatigue feature, fatigue can be predicted early by the respiratory rate. The difficulty of respiratory rate fatigue detection lies in the extraction of respiratory rate features and detection accuracy, so this paper adopts the non-contact respiratory detection based on dense optical flow method to solve the above problems. In this paper, a large number of experimental comparisons between respiratory rate fatigue detection and facial fatigue detection have confirmed that fatigue can be detected faster through respiratory rate fatigue characteristics, but there will still be a small deviation in the detection of respiratory rate, so this paper proposes to compare facial fatigue characteristics with breathing The fatigue detection method combined with the rate of fatigue characteristics, this method greatly improves the accuracy and speed of fatigue detection.

13:30-13:45 V041

Robust Real-Time Pedestrian Detection on Embedded Devices

Mohamed Afifi, Yara Ali, Karim Amer, Mahmoud Shaker, and Mohamed Elhelw
Nile University, Egypt

Detection of pedestrians on embedded devices, such as those on-board of robots and drones, has many applications including road intersection monitoring, security, crowd monitoring and surveillance, to name a few. However, the problem can be challenging due to continuously-changing camera viewpoint and varying object appearances as well as the need for lightweight algorithms suitable for embedded systems. This paper proposes a robust framework for pedestrian detection in many footages. The framework performs fine and coarse detections on different image regions and exploits temporal and spatial characteristics to attain enhanced accuracy and real time performance on embedded boards. The framework uses the Yolo-v3 object detection [1] as its backbone detector and runs on the Nvidia Jetson TX2 embedded board, however other detectors and/or boards can be used as well. The performance of the framework is demonstrated on two established datasets and its achievement of the second place in CVPR 2019 Embedded Real-Time Inference (ERTI) Challenge.

13:45-14:00 V043

Line detection via a lightweight CNN with a Hough Layer

Lev Teplyakov, Kirill Kaymakov, Evgeny Shvets and Dmitry Nikolaev
IITP RAS, Russia

Thursday, 5 Nov. 2020

Line detection is an important computer vision task traditionally solved by Hough Transform. With the advance of deep learning, however, trainable approaches to line detection became popular. In this paper we propose a lightweight CNN for line detection with an embedded parameter-free Hough layer, which allows the network neurons to have global strip-like receptive fields. We argue that traditional convolutional networks have two inherent problems when applied to the task of line detection and show how insertion of a Hough layer into the network solves them. Additionally, we point out some major inconsistencies in the current datasets used for line detection.

14:00-14:15 V050

Maximizing Object Detection Using Suas

Curtis Manore, Pratheek Manjunath, Dominic Larkin
United States Military Academy, USA

This paper examines optimal look-angles for a camera which is mounted on a small unmanned aerial system (sUAS), that provides for maximized object detection on the ground. Using a generic convolutional neural network (CNN), this research identifies the best angle for detecting a ground target from an aerial perspective. The study involves altering camera angles on an sUAS that is flown along a fixed trajectory and then determining the angle which provides the highest detection rate of predefined objects, which are emplaced at known locations on the ground. The experiment is conducted in simulation and validated on a physical quadcopter. The results of this paper directly influence the U.S. Army's research efforts on training neural networks and developing object detection algorithms.

14:15-14:30 V096

Multimodal Features for Shots Boundary Detection

Mohamed Bouyahi, Yassine Ben Ayed
MIRACL laboratory, Tunisia

Shot Boundary Detection (SBD) also known as a temporal video segmentation is a preprocessing task for multiple videos applications, such as indexing and retrieval. The SBD output provides coherent temporal units which are easy to manipulate. The Most previous works implement theirs frameworks based on visual features to measure similarity for transition detection task. However, the video is very enriched by data which could be beneficial. In this paper, referring to recent multimodal works, we propose to introduce the audio components to increase the SBD task. Firstly, we worked on candidate segments obtained by measuring similarity between low features (SURF, HSF) from original video. Then we used deep features obtained from trained model (Resnet-50) for visual similarity and we introduced the audio segmentation based on Power Spectrum Density (PSD) to contribute for transition detection. The proposed method is evaluated on the clip shots dataset. Experiments on this data show that the proposed multimodal approach can achieve a better performance compared with the state-of-the-art of methods that used visualapproach.

Thursday, 5 Nov. 2020

Session 9
Artificial Intelligence and Data Engineering

Virtual Room B:
680 4475 3785

Chair: Stephen R. Gulliver — University of Reading, UK

13:00-13:15 S027

A Neural-Assessment System Based on Emirates (QFE)

Fatima Alzahraa Kouka, Lamees Mohammad Dalbah, Shaimaa Mahmood Mounir, Raed Abu Zitar
Ajman University, United Arab Emirates

In order to strengthen the teaching and learning phase it is assumed that the assessment of course results dependent upon student grades is important. Our analysis methods and workflows leverage the benefits of AI, for example the capacity to evaluate vast data sets and detect correlations more accurately than humans would use artificial intelligence technologies to help classify large data. It would be used to assess the course learning results based on QF-Emirates (Qualifiers Frame of the United Arab Emirates) criteria with actual data and use it to recommend teaching and learning interventions. We investigate and validate the right neural networks architecture that produces full performance. To that end, a modern algorithm has been improvised. Application to a database to store data and provide data regarding the review of course learning results would be deployed for our suggested recommendation framework. The suggestion method is evaluated and findings are promising as a machine-learning framework. Our neural network based system was able to generate solutions for new cases and provide support in the assessment of courses learning outcomes.

13:15-13:30 SE022

Specifying Key-Properties to Improve the Recognition Skills of Neural Networks

Benjamin Jahić, Guelfi Nicolas, Ries Benoît
University of Luxembourg, Luxembourg

Software engineers are increasingly asked to build datasets for engineering neural network-based software systems. These datasets are used to train neural networks to recognise data. Traditionally, data scientists build datasets consisting of random collected or generated data. Their approaches are often costly, inefficient and time-consuming. Software engineers rely on these traditional approaches that do not support precise data selection criteria based on customer's requirements. We introduced a software engineering method for dataset augmentation to improve neural networks by satisfying the customer's requirements. In this paper, we introduce the notion of key-properties to describe the neural network's recognition skills. Key-properties are used all along the engineering process for developing the neural network in cooperation with the customer. We propose a rigorous process for augmenting datasets based on the analysis and specification of the key-properties. We conducted an experimentation on a case study on the recognition of the state of a digital meter counter. We demonstrate an informal specification of the neural network's key-properties and a successful improvement of a neural network's recognition of the meter counter state.

13:30-13:45 SE023

Dynamic Data Consistency Tests Using a CRUD Matrix as an Underlying Model

Miroslav Bures, Vaclav Rechtberger
Czech Technical University in Prague, Czech Republic

In testing of software and Internet of Things (IoT) systems, one of necessary type of tests has to verify the consistency of data that are processed and stored in the system. The Data Cycle Test technique can effectively do such tests. The goal of this technique is to verify that the system processes data entities in a system under test in a correct way and that they remain in a consistent state after operations such as create, read, update and delete. Create, read, update and delete (CRUD) matrices are used for this purpose. In this paper, we propose an extension of the Data Cycle Test design technique, which is described in the TMap methodology and related literature. This extension includes a more exact definition of the test coverage, a reflection of the relationships between the tested data entities, an exact algorithm to select and combine read and update operations in test cases for a particular data entity, and verification of the consistency of the produced test cases. As verified by our experiments, in comparison to the original Data Cycle Test technique, this proposed extension helps test designers to produce more consistent test cases that reduce the number of undetected potential data consistency defects.

13:45-14:00 SE042

A Framework Based on Model Driven Engineering and model Weaving to Support Data-Driven Interoperability for Smart Grid Applications

Eder Matheus Silveira Felix, Denivaldo Cicero Lopes, Osvaldo Silva Sousa Jr.
Federal University of Maranhão, Brazil

Thursday, 5 Nov. 2020

The smart grid is envisioned as a new and improved power grid. The high complexity of this system of systems is raising the need of engineering methodologies that manages this complexity. The Model Driven Engineering (MDE) approach has been proposed in literature to increase the productivity and reduce the software development costs. The MDE deals with complexity by using models as the main artifacts of the software development process. Moreover, the use of standard-compliant models to represent data in the power system domain its being strongly encouraged to improve interoperability. In this paper, we propose a framework based on Model Driven Engineering and the weaving of models to support the development of interoperable applications for the smart grid domain. Our approach separates the modeling of the application from the modeling of the data that is shared between applications. The elements of those models are related using the weaving technique. The proposed approach was implemented on Eclipse platform, and two applications were developed and deployed to validate the approach.

14:00-14:15 S001

Constructing Neural Networks By Extending The Optimization Field

Wenyuan Zhang
Huazhong University of Science and Technology, China

The performance of the neural network is not unilaterally increased with the training parameters of the neural network, and it is possible that the parameters of the training ground become more and the performance is degraded. This paper explains what the optimization field is and proposes a new way to solve this degradation problem. The new method is to built a neural network which includes the previous optimization field while adding parameters and extends the new optimization field. From this point, the paper explains the excellence of Inception, ResNet and DenseNet networks. Furthermore, this paper also proposes a new method for constructing neural networks that can improve the performance of neural networks with the increase of training parameters.

14:15-14:30 S020

On the relationship between input sparsity and noise robustness in Hierarchical Temporal Memory Spatial Pooler

Damir Dobric, Andreas Pech, Bogdan Ghita, Thomas Wennekers
University of Plymouth, Germany

Hierarchical Temporal Memory - Spatial Pooler is a cortical learning algorithm inspired by the biological functioning of the neocortex. It is responsible for the sparse encoding of spatial patterns used as an input for further processing inside of a Hierarchical Temporal Memory (HTM). During the learning process, the Spatial Pooler groups spatially similar inputs into the same sparse distributed representation (SDR) memorized as a set of active mini-columns. The role of SDR generated by the learning process of the Spatial Pooler is to provide an input for learning of sequences inside of the HTM. One of the features of the Spatial Pooler is also the robustness to noise in the input. This paper summarizes the work in progress, which analyses the relationship between the encoding of the input pattern and the robustness of the memorized pattern against noise. In this work many synthetic input patterns with different sparsity were used to set the hypothesis, which claims that SP robustness against the noise in the input depends on the sparsity of the input. To validate the hypothesis, many random input vectors with a large portion of noise were generated. Then the change of SDR output was compared with the change of input by the given portion of noise. It was shown that the SDR output change is very small in comparison to change of the input by adding of a large portion of the noise in the input. By adding of a significant portion of the noise to the input, the learned output remains almost unchanged. This indicates a great robustness to noise. Experiments show that the robustness against noise of the Spatial Pooler directly depends on the sparsity of the input pattern. Preliminary tests suggest implementation of a boosting mechanism of the input to improve the robustness against noise.

14:30-14:45 S024

Improving Fairness in Speaker Recognition

Gianni Fenu, Giacomo Medda, Mirko Marras, Giacomo Meloni
University of Cagliari, Italy

The human voice conveys unique characteristics of an individual, making voice biometrics a key technology for verifying identities in various industries. Despite the impressive progress of speaker recognition systems in terms of accuracy, a number of ethical and legal concerns has been raised, specifically relating to the fairness of such systems. In this paper, we aim to explore the disparity in performance achieved by state-of-the-art deep speaker recognition systems, when different groups of individuals characterized by a common sensitive attribute (e.g., gender) are considered. In order to mitigate the unfairness we uncovered by means of an exploratory study, we investigate whether balancing the representation of the different groups of individuals in the training set can lead to a more equal treatment of these demographic groups. Experiments on two state-of-the-art neural architectures and a large-scale public dataset show that models trained with demographically-balanced training sets exhibit a fairer behavior on different groups, while still being accurate. Our study is expected to provide a solid basis for instilling beyond-accuracy objectives (e.g., fairness) in speaker recognition.

Thursday, 5 Nov. 2020

Session 10
Image Analysis and Methods

Virtual Room A:
976 5819 4637

Chair: Udo J. Birk — University of Applied Sciences of the Grisons, Switzerland

15:00-15:15 V027

Keypoint-based Static Object Removal from Photographs

Alexandr Volkov, Valeria Efimova, Viacheslav Shalamov and Andrey Filchenkov
ITMO University, Russia

When taking photos of historical buildings and landmarks, one can often encounter more modern and less attractive objects blocking the view (poles, electricity cables, road signs, and others). Photographers strive to take the unobstructed shot, without random objects popping up in the image. However, they avoid deleting unwanted objects with inpainting or exemplar-based methods, because it can introduce notable artefacts. We propose a new algorithm for removing large static objects such as road signs, sitting people, and parked cars from a photograph keeping its originality and unique details. We do not use artificial patterns to fill covered regions, only two shots taken from different angles. Usually, removing long thin objects such as road sign poles cause image deformations, but our approach avoids this. For comparison, we created a dataset from “Caltech Buildings” and “Kaggle Architecture” Datasets and added road signs, cars, and other objects to photos. We compared our approach with state-of-the-art methods such as Deep Image Prior, Gated Convolution, and the Region Filling by block sampling. The real photographs of historical building demonstrate the effectiveness of our algorithm. Code and example images are available on GitHub.

15:15-15:30 V033

Improvement of U-Net architecture for image binarization with activation functions replacement

Alexander Gayer, Alexander Sheshkus, Dmitri P. Nikolaev, Vladimir V. Arlazarov
NUST “MISIS”, Moscow, Russia

In this work we study the effect of activation functions in a neural network. We consider how activation functions with different properties and their combination affect the final quality of the model. Due to optimization and speed performance issues with most of bounded functions that are represented by sigmoids, we propose the generalized version of SoftSign function - ratio function (rf). Its shape greatly depends on introduced degree parameter, which in theory leads to new interesting property - contraction to zero. For evaluation, we chose image binarization problem: based on U-Net architecture of DIBCO-2017 winners, we conducted all experiments with replacing activation functions only. Our research has led us to the state-of-the-art results in binarization quality on DIBCO-2017 test dataset. U-Net with modified activation functions significantly outperforms all existing solutions in all metrics.

15:30-15:45 V037

On the effectiveness of Adversarial Unsupervised Domain Adaptation for Iris Presentation Attack Detection in Mobile Devices

Yomna Safaa-El-Din, Mohamed Moustafa and Hani Mahdi
Ain Shams University – Cairo, Egypt

The growing usage of smart mobile devices have made authentication with biometric data more convenient. On the other side, videos and photos of users are becoming more available online. This makes it easier for attackers to spoof the authentication systems which rely on face and eye-region data for instance. One major problem with current Presentation Attack Detection (PAD) systems is their lack of generalization to data captured by different sensors or in different environments. In this paper, we propose the use of unsupervised domain adaptation to solve this PAD problem, specifically the iris PAD. Our model is composed of symmetric classifiers and two per-class domain discriminators. Interaction between class probabilities and domain classification is utilized to jointly adversarialy train a mobile-oriented feature extraction network, capable of generating domain-invariant features. The approach is evaluated on three benchmark iris PAD datasets. Results show up to 40% improvement in cross-dataset Average Classification Error Rate (ACER) proving the effectiveness of the approach in increasing the robustness and generalization of biometric PAD systems.

15:45-16:00 V038

The method of search for falsifications in copies of contractual documents based on N-grams

Oleg Slavin, Elena Andreeva, Vladimir V. Arlazarov
Smart Engines Service Ltd., Moscow, Russian Federation

Thursday, 5 Nov. 2020

This article is focused on methods of search for falsifications in scanned copies of business documents. This task arises from a comparison of two copies of business documents signed by two parties. The comparison should be performed to detect possible changes made by one of the parties. This problem is relevant, for instance, in the banking sector when signing agreements on paper. The method of partial search for matching flexible documents, where text attributes may be changed, and unintentional modifications of non-essential words may be made is considered. The method of comparison of two scanned images based on the recognition and analysis of N-grams word sequences is proposed. The proposed method has been tested on private dataset. The proposed method has demonstrated high quality and reliability of the search for differences in two samples of one agreement-type document.

16:00-16:15 V049

LRA-Net: Local Region Attention Network for 3D Point Cloud Completion

Hang Wu, Yubin Miao
Shanghai Jiao Tong University, China

Incomplete Point clouds obtained from one-side scanning always result in structural loss in 3D shape representations, thus many learning-based methods are proposed to restore complete point clouds from partial ones. However, most of them only utilize global features of inputs to generate outputs, which might lose details. In this paper, a new method that utilizes both global and local features is proposed. First, Local features are extracted from inputs and analyzed under the conditions interpreted by global features. Second, conditional local feature vectors are deeply fused with each other via graph convolution and self-attention. Third, deeply-fused features are decoded for generating coarse point clouds. Last, global features extracted from inputs and coarse outputs are combined to generate fine outputs with high-density. Our network is trained and tested on eight categories of objects in ModelNet. The results show that our network is able to overcome instability in local feature awareness, restore complete point clouds with more details and smoother shapes, and outperform most of those existing methods both intuitively and quantitatively. Our source codes will be available at: https://github.com/wuhang100/LRA-Net.

16:15-16:30 V087

Slope detection criterion robust to sparse 2D data

Dmitry Bocharov, Alexey Kroshnin, Dmitry Nikolaev
Institute for Information Transmission Problems, RAS, Russia

The study is referred to a task of 2D data slope estimation. We consider the integral projections analysis technique and a common criterion of sum of squared values (SSV) for optimal angle detection. This criterion is dependent on the density of input data and for very sparse data its efficiency significantly decreases. We propose the alternative criteria – the sum of the inversed lengths (SIL) that preserves SSV characteristics for dense data but that is much more robust for sparse input. The experiments conducted on simulated and real datasets demonstrate better quality of slope detection using the proposed criterion.

Thursday, 5 Nov. 2020

Session 11
Computer and Information Security

Virtual Room B:
680 4475 3785

Chair:

15:00-15:15 SE012

A Cost Model for Decoder Decision Trees

Ms. Lillian Tadros
Technical University of Dortmund, Germany

Instruction decoders are an innate part of the System-on-Chip design flow and major constituents of instruction set simulators and processor toolchains. The growing complexity of instruction sets has rendered the manual implementation of decoders a daunting task and one which seldom yields acceptable performance. Tools that autonomously generate decoder decision trees based on high level instruction descriptions attempt to address this situation. Due to the large number of possible solutions, such tools typically support only a vague notion of tree cost and instead make heavy use of heuristics in the generation process, possibly excluding the best candidates. This paper analyzes the performance of decoder decision trees with distinct parameters regarding shape, size, and implementation. Based on the experimental evaluation, a model of tree cost is developed. Our analysis takes the frequency of occurrence of instructions into account. Our model of tree cost is not restricted to the use-case of decoding instructions and is parametrized so as to be applicable to decision trees of given properties.

15:15-15:30 SE021

A Generic Framework for Capturing Reliability in Cyber-Physical Systems

Mr. Nazakat Ali, Manzoor Hussain, YoungJae Kim and Jang-Eui Hong
Chungbuk National University, Republic of South Korea

Research Purpose: Cyber-Physical Systems solve complex problems through their tight integration between the physical and computational components. Therefore, the reliability of a complex system is the most critical requirement for the cyber-physical system because an unreliable system often leads to service disruption, property damage, financial loses and sometimes lead to fatality. In order to develop more reliable CPS, this paper proposes a generic framework for reliability modeling and analysis for our ongoing work on cyber-physical systems. Research Method: This paper, at first defines an architecture for general CPS which is comprised of three layers; environment layer, communication layer, and computational layer. Secondly, we formalize a reliability model for the architectural components, and then propose a framework for the reliability of CPS with the consideration of how to capture the reliability. Results: Based on the research method, we demonstrate the proposed framework with an illustrative example by using different reliability values from “offshore and onshore reliability data” library. We confirmed that the reliability model covers almost all possible reliabilities required to general cyber-physical systems.

15:30-15:45 SE025

LDA Categorization of Security Bug Reports in Chromium Projects

Mr. Wajdi Mohammed Aljedaani, Yasir Javed, Mamdouh Alenezi
Al-Kharj College of Technology, Saudi Arabia

Security bug reports (SBR) depict potential security vulnerabilities in software systems. Bug tracking systems (BTS) usually contain huge numbers of bug reports including security-related ones. Malicious attackers could exploit these SBRs. Henceforth, it is very critical to pinpoint SBRs swiftly and correctly. In this work, we studied the security bug reports of the Chromium project. We looked into three main aspects of these bug reports, namely: frequencies of reporting them, how quickly they get fixed and is LDA effective in grouping these reports to known vulnerabilities types. We report our findings in these aspects.

15:45-16:00 SE026

Supporting Contact Tracing by Privacy-Friendly Registration at Catering Facilities

Dr. Michiel Willocx, Dave Singelée, Jorn Lapon, Vincent Naessens
Imec-DistriNet, Belgium

The corona virus has hit the world with an unprecedented global pandemic, forcing many countries around the world in a lockdown with social distancing measures. Along with a wide range of hygienical measures such as wearing masks and washing hands, contact tracing solutions aim at reacting quickly to new infections. Two approaches are widely employed in an attempt to control the spreading of the corona virus, namely permanent contact tracing and hospitality unit registration.

Thursday, 5 Nov. 2020

For permanent contact tracing, privacy-friendly solutions such as DP-3T [1] are available. However, for the mandatory hospitality unit registration, current (digital) solutions remain rather primitive and do not offer satisfying privacy properties. This work proposes a privacy-preserving, digital solution for the mandatory registration in bars and restaurants while satisfying the needs of governmental institutes, the customer and the bar owners.

16:00-16:15 S025

CRATOS: Cognition of Reliable Algorithm for Time-series Optimal Solution

Dr. Ziling Wu, Ping Liu, Zheng Hu, Bocheng Li, Jun Wang
Sun Yat-sen University, China

Anomaly detection of time series plays an important role in reliability systems engineering. However, in practical application, there is no precisely defined boundary between normal and anomalous behaviors in different application scenarios. Therefore, different anomaly detection algorithms and processes ought to be adopted for time series in different situation. Although such strategy improve the accuracy of anomaly detection, it takes a lot of time for practitioners to configure various algorithms to millions of series, which greatly increases the development and maintenance cost of anomaly detection processes. In this paper, we propose CRATOS which is a self-adapt algorithms that extract features from time series, and then cluster series with similar features into one group. For each group we utilize evolutionary algorithm to search the best anomaly detection methods and processes. Our methods can significantly reduce the cost of development and maintenance of anomaly detection. According to experiments, our clustering methods achieves the state-of-art results. The accuracy of the anomaly detection algorithms in this paper is 85.1%.

16:15-16:30 S002

A Survey on Information Diffusion in Online Social Networks

Dr. YANG Yujie
Wuhan University, China

Nowadays, social networks have become a critical data dissemination platform with the drastic proliferation of social networks and the growing recipient of data. In this paper, recent successful works of information diffusion in online social networks are introduced. We summarize several classic information diffusion models like an explanatory model: the SI Model, the SIS Model, the SIRS Model; predictive Model: Independent Cascade Model, Linear Threshold Model. Then, we discuss some applications of those information diffusion model in different social networks.

16:30-16:45 SE028

An Approach based on Model Driven Engineering to Support the Development of Web of Things

Ms. Rayanne Silva de Oliveira, Denivaldo Lopes
Federal University of MaranhãoUFMA, Brazil

Web of Things (WoT) is one of the most emerging technologies because it provides a solution to support protocols interoperability and applications heterogeneity that can be found in Internet of Things (IoT) based proposals. In order to facilitate the development of WoT solutions and communication with IoT proposals, thus preventing them from becoming out of date and allowing them to benefit from WoT based solutions, a framework based on Model Driven Engineering (MDE) technologies is proposed in this paper. MDE Technologies are responsible for automating the framework's development processes, reducing development time and human errors. Node-WoT is used, which is a library responsible for creating Thing Description and supporting WoT proposals. WebSockets, which is the communication link between WoT proposals and IoT devices, is also used. The necessary methods for framework implementation are described, and finally an illustrative example is shown, which consists of creating the Thing Description of an LED that communicates with an IoT device through WebSockets that represents the physical part of the device, thus proving its applicability in supporting the development of WoT and in increasing the interoperability of IoT based proposals.

Thursday, 5 Nov. 2020

16:45-17:00 SE044

An Empirical Investigation of Spikes in Agile Software Development

Mr. Hussein Al Hashimi, Abdullah Altaleb, Andrew Gravell
University of Southampton, UK

The use of spikes in agile software development (ASD) can enable organizations to produce quality software by employing the required technical expertise, planning the entire development cycle and ensuring that the client's requirements are adhered to. This study aims to examine the use of spikes in ASD. It explores the role, efficiency and efficacy of spikes in various software development domains through the different agile methods. An exploratory research design is adopted to achieve this purpose, whereby mixed methods are used to collect concurrently both qualitative and quantitative data from the experts recruited to the study. Based on the study's findings, this paper shows spikes' impact on risk reduction and their role in the estimation process, as well as how their use is related to the team's understanding and the consistency and reliability of the story estimate. Through interviews, it establishes that the primary role of spikes is risk management through investigations to understand user stories and reveal any uncertainty. Both the efficiency and efficacy of spikes were found to be high. The findings further showed that spikes improve the quality of the end product.

Friday, 6 Nov. 2020

Session 12
Intelligence Recognition Technology and Algorithm

Virtual Room A:
976 5819 4637

Chair:

10:00-10:15 V005

A Novel Approach To Improve The Social Acceptance Of Autonomous Driving Vehicles By Recognizing The Emotions Of Passengers

Antonio Costantino Marceddu, Jacopo Sini, Massimo Violante, Bartolomeo Montrucchio
Polytechnic of Turin, Italy

For some years now, the idea of a fully autonomous driving car has been monopolizing the attention of the entire automotive sector. The main motivation behind all this interest is that autonomous driving cars could potentially improve road safety simply by subtracting human error, which causes thousands of road fatalities worldwide every year. One of the biggest challenges, for which researchers are working hard to give answers, concerns the search of solutions to improve the people's response to this vehicles once they are put on the road: an incorrect answer can lead to their unsuccess, with a great monetary loss for all the companies that have invested in these technologies. Recently, we worked on a project that went in this direction and which regarded the possibility to change the driving style of the autonomous driving cars based on the passengers' facial expressions. This is particularly useful in the event that they experience fear: the car may react to these feelings by adopting a careful driving style and stopping if these feelings do not improve within a certain period. In this paper, we want to discuss about the improvements we have made to this project.

10:15-10:30 V015

Steel pipe counting system based on image recognition

Yuxin Jia, Jie Xu, Qujiang Lei , Xiuhao Li, Guangchao Gui, Weijun Wang
Guangzhou Institute of Advanced Technology, Chinese Academy of Sciences, China

Traditional steel pipe counting can only be performed manually, which has the problems of large workload, error-prone and low efficiency. This paper proposes a steel pipe counting system based on image recognition. First of all, the steel pipe image is collected from the camera, and then the digital image processing technology is used to perform image enhancement, edge detection and morphological operation and other preprocessing, then the steel pipe is identified and counted by the Hough circle transformation, and finally the steel pipe automatic is developed using VC++ software Technical software. Experiments show that the automatic counting accuracy and efficiency of steel pipes in this system are high.

10:30-10:45 V018

The Impact of Pre-Processing Algorithms in Facial Expression Recognition

Daniel Canedo, António J. R. Neves
University of Aveiro, Portugal

This paper proposes several pre-processing algorithms to improve facial expression recognition based on Convolutional Neural Networks (CNNs) models. The proposed CNN model was trained on the Extended Cohn-Kanade dataset (CK+) after applying the pre-processing stages and achieved competitive results (93.90% recognition accuracy) despite its simple and light architecture. Using this CNN model, a study on the impact of each pre-processing algorithm when extracting facial features is presented. In the end, it is understood that pre-processing algorithms help CNNs to extract the most relevant features for each facial expression more effectively, reducing the overfitting and increasing the recognition accuracy. Attention maps before and after the pre-processing step are shown in order to visualize its impact when the proposed CNN model makes a prediction.

10:45-11:00 V052

Precise localization of synchronization patterns for Aztec Code matrix extraction

Aleksey D. Bursikov, Pavel V. Bezmaternykh and Vitaliy M. Kliatskine
Moscow Institute for Physics and Technology, Russia

Two-dimensional barcode reading is still a challenging problem in the computer vision domain. To simplify this reading process specially designed synchronization patterns are introduced into the barcode structure. Their precise localization allows to use block-based matrix extraction strategy, which is essential for processing images captured in an uncontrolled environment. This paper describes an algorithm for precise localization of such synchronization patterns for the Aztec Code symbology. The proposed algorithm is based on uniform linear motion model usage supported by

Friday, 6 Nov. 2020

Kalman filtering to increase its estimation accuracy. Experimental results showed that the replacement of basic ISO-based algorithm with the proposed algorithm increased reading quality from 39% to 98% on the dataset which consists of 1628 real-life document images captured with mobile device cameras in uncontrolled conditions.

11:00-11:15 V073

Application of shared backbone DNNs in ADAS perception systems

Mikhail G. Lobanov, Dmitry L. Sholomov
Cognitive Robotics Ltd., Russia

Advanced Driver Assistance System (ADAS) is a very important part of an up to date vehicle. For achieving high-level objectives in such ADAS functionality like LKA (lane keeping assistance), LDW (lane departure warning system), FCW (forward collision warning) the quality of the algorithms under the hood must be extremely high. In the last few years, it is common that these algorithms are based on DNNs (deep neural networks) applied to the tasks of semantic and instance segmentation, 2D/3D object detection and visual object tracking. Recent state-of-the-art DNN models as usual solve only one single task from the listed above and running several neural networks is rather computationally expensive and even impossible due to the lack of the GPU memory. One of the approaches used to overcome such a problem is a shared backbone (also called feature extractor or encoder). The backbone consumes most of the computing resources thus the model with a shared backbone achieves better inference performance. Unfortunately, the training procedure for a shared backbone model has several difficulties. The first one is the lack of datasets with all the required and uniform annotation types. The second problem is a more sophisticated backpropagation procedure. In this paper, we consider several methods for multi-task neural network training and present the results of such training procedures on several public datasets with dissimilar annotation types. The shared backbone is applied to the following three tasks performed simultaneously on the road scene: semantic segmentation, 2D object detection and 3D object detection. While the performance of the DNNs with shared backbone increased significantly, we obtained the quality evaluation results, which are quite close to the original separate state-of-the-art DNNs and even outperforms them in some evaluation indices.

11:15-11:30 V103

Fingerspelling Recognition Using Synthetic Images and Deep Transfer Learning

Nguyen Tu Nam, Shinji Sako, Bogdan Kwolek
Nagoya Institute of Technology, Japan

Although gesture recognition has been intensely studied for decades, it is still a challenging research topic due to difficulties posed by background complexity, occlusion, viewpoint, lighting changes, the deformable and articulated nature of hands, etc. Numerous studies have shown that extending the training dataset with real images about synthetic images improves the recognition accuracy. However, little work is devoted to demonstrate what improvements in recognition can be achieved thanks to transferring the style onto synthetically generated images from the real gestures. In this paper, we propose a novel method for Japanese fingerspelling recognition using both real and synthetic images generated on the basis of a 3D hand model. We propose to employ a neural style transfer to include information from real images onto synthetically generated dataset. We demonstrate experimentally that neural style transfer and discriminative layer training applied to training deep neural models allow obtaining considerable gains in the recognition accuracy.

Friday, 6 Nov. 2020

Session 13
Digital Image Processing and Application

Virtual Room B:
680 4475 3785

Chair:

10:00-10:15 S003

Deep Image Compositing

Shivangi Aneja, Dr. Soham Mazumder
Technical University of Munich, Germany

In image editing, the most common task is pasting object from one image to the other and then eventually adjusting the manifestation of the foreground object with the background object. This task is called image compositing. But image compositing is a challenging problem which requires professional editing skills and considerable amount of time. Not only these professionals are expensive to hire, but the tools (like Adobe Photoshop) used for doing such tasks are also expensive to purchase making the overall task of image compositing difficult for people without this skillset. In this work we aim to cater to this problem by making composite images look realistic. To achieve this, we are using Generative Adversarial Networks (GANS). By training the network with diverse range of filters applied to the images and special loss functions, the model is able to decode the color histogram of foreground and background part of the image and also learns to blend the foreground object with the background. The hue and saturation values of the image plays an important role as discussed in this paper. To the best of our knowledge, this is the first work that uses GANs for that task of image compositing. Currently, there is no benchmark dataset available for image compositing. So we created the dataset and will also make the dataset publicly available for benchmarking. Experimental results on this dataset show that our method outperforms all current state-of-the-art methods.

10:15-10:30 S101

Dog and Cat classification with deep residual network

Ms. Yao Yuchen
Software School of North University China, China

I With the development of artificial intelligence, the deep neural network(DNN) has achieved excellent results in image processing domain such as image classification[1] and objct detection[2].

The convolution neural networks(CNN) [3] is a Representative algorithm of DNN which have the representation learning ability . According to its convolutional structure, input information is extracted with translation invariance.Based on the widely used CNN ,there are many efficient models.For image classification there are Lenet-5[4],VGG[5], Resnet[6] and so on.For object detection ,the yolo series[7] is well-known.Also few well known datasets are proposed to measure their performance such as ImageNet and Cifar-10[8]. These data sets are dedicated to the classification of multiple objects in natural scenes.Nowadays ,pets play an increasingly important role in our life,so we built a cat and dog dataset, each of which categories with 12500 samples which is larger then 1260 in Imagenet.For our dataset,we trained an image classification model.We focus on the performance of distinguish dog and cat In different scenes, lighting and noise.Our method achieved an accuracy of 92.7 percent and remained robust under adversarial attack.

10:30-10:45 S018

Generation of Datasets for Semantic Segmentation from 3D Scanned Data to Train a Classifier for Visual Navigation

Dr. Hayato Komatsuzaki, Raimu Yokota, Shogo Sakata, Miho Adachi, Ryusuke Miyamoto
Meiji University, Japan

The authors attempt to actualize robot navigation using only visual information and propose a novel scheme based on the results of semantic segmentation for road-following in urban and indoor scenes. It is shown that the proposed vision-based navigation performs well when accurate results of semantic segmentation are obtained. Generally, to construct a good classier, an appropriate dataset should be prepared. However, creating a dataset for semantic segmentation with pixel-wise class labels to all pixels involves a significant amount of human effort. Hence, we proposes a novel approach to generate datasets for semantic segmentation that is suitable for the visual navigation of a robot. The proposed approach provides two-dimensional images with pixel-wise class labels from three-dimensional scanned data. Experimental results show that the generated dataset can be used for training semantic segmentation using ICNet.

10:45-11:00 S013

Attention Guided Multi-Scale Regression for Scene Text Detection

Dr. Zhiwei Zheng
Huazhong University of Science and Technology, China

Friday, 6 Nov. 2020

With the development of deep learning, scene text detection has made significant progress. A large number of neural network models have been applied to this task, one of which is a fully convolutional network (FCN) model named An Efficient and Accurate Scene Text Detector (EAST). However, it usually falls short when detecting long text regions like text lines lying across the images. Therefore, we propose a new network with Attention Module and Image Pyramid Module, based on EAST but largely ameliorated, to overcome the obstacles. Experiments conducted on several public benchmarks demonstrate that our method outperforms the original EAST model.

11:00-11:15 S021

Breast Lesion Detection from Mammograms Using Deep Convolutional Neural Networks

Dr. Gloria Gonella, Marco Paracchini, Elisabetta Binaghi, Marco Marcon
University of Insubria, Italy

Mammography has a central role in screening and diagnosis of breast lesions, allowing early detection of the pathology and reduction of fatal cases. Deep Convolutional Neural Networks have shown a great potentiality to address the issue of early detection of breast cancer with an acceptable level of accuracy and reproducibility. In the present paper, we illustrate the development of a deep learning study aimed to process and classify lesions in mammograms with the use of slender neural networks not yet used in literature. For this reason, a traditional convolution network was compared with a novel one obtained making use of much more efficient depth wise separable convolution layers. Preliminary numerical results are detailed and future plans outlined.

11:15-11:30 S022

Development of a Unified and Dynamic Geometric Framework for Modelling Plant Leaf Spots

Dr. Duaa Alshadli, Kambiz Borna, Cesar Lador
Unitec Institute of Technology, New Zealand

The use of computer vision and image processing techniques have proven to be effective in detecting and classifying plant diseases from symptoms, such as leaf spots or mosaic leaf patterns caused by pathogens. To identify these symptoms, these techniques typically utilize a static geometry specified by a human expert via pixels or image objects. Thus the results rely on generic parameters defined by the user before or after classification. In this paper, a dynamic geometry is proposed that can be applied to identify plant disease symptoms without setting any geometric parameters. The offered method consists of two primary phases established based on the notion of Vector Agents (VAs): construction of a unified geometry, and creation of a dynamic geometry. In the construction step, the method utilises a set of geometric rules to link the raster space to the vector space during the simulation process. These rules specify the growth direction of a leaf in the simulation space and the maximum length of each edge specified based on the image spatial resolution and size. The creation step includes four main geometric operators: 1) vertex displacement, 2) half-edge joining, 3) converging vertex displacement, and 4) edge remove. This enables the leaves to automatically change their geometry in the simulation space without setting any geometric parameters. This structure allows a classifier to model a wide range of leaf shapes in the image space. The proposed geometry method was tested to model different leaves and the results demonstrate its effectiveness in generating different leaf shapes.

11:30-11:45 S008

A Fabric Image Segmentation Pipeline to Find the Yarn Weaving Pattern from a Single Photograph

Dr. Cihan Bal, Mehmet K. Baran
Marmara University, Turkey

Rendering high quality woven fabric images need special rendering techniques that differ from standard approaches. Since a woven fabric consists of vertically and horizontally crossing elliptical yarns and yarns, consist of cylindrical microfibers, a suitable, revised rendering approach is needed. Therefore, a rendering model for woven fabric images is proposed “Microcylinder Appearance Model”. The model needs some parameters to produce a realistic rendering of different fabric types such as silk, linen, velvet, and polyester satin. In the original article, authors derived these parameters via eye inspection. Our ultimate aim is to automatize this process. In this article, we propose a pipeline that does the first step of this automatization, i.e. segmenting the captured image via image processing techniques and finding the yarn crossing points. The segmented image can further be utilized to extract parameters of cylindrical shading models in future works.

Friday, 6 Nov. 2020

Session 14
Digital Image Processing and Application

Virtual Room A:
976 5819 4637

Chair: Alexander Bernstein — Skolkovo Institute of Science and Technology, Russia

13:00-13:15 V008

Event Correlation for Deception Detection in Long Video

Liqian Gao, Jianbang Qin, Chunhui Du, Wei Guo
Shanghai Jiao Tong University, China

Automated Deception Detection (ADD) is a challenging task and still under study as a visual analysis task. Based on the idea that human micro-expressions and body movements could be used as clues for ADD, many works have proposed some action recognition models for extracting face and body spatiotemporal features. However, these features are not sufficient evidence for deception; moreover, micro-expressions are difficult to detect and real-life deception samples are hard to collect, thus ADD still has many challenges. In this paper, we present a global two-stream network (GTSN), which not only extracts face and body features, but also utilizes the correlation between the deceptions. GTSN can improve the accuracy of deception detection by adding historical information based on the correlation between the deceptions. We build a dataset named Deception-Truthful (DT) for evaluating the performance of our proposed model. Experimental results demonstrate that our GTSN model outperforms other action recognition models used for ADD. Further, the proposed GTSN model also performs well on the real trial videos widely used in ADD.

13:15-13:30 V009

Weak Anomaly-reinforced Autoencoder for Unsupervised Anomaly Detection

Xinqiang Chen, Lumei Su, Guansen Deng, Mingyong Huang, Jiajun Wu and Yanqing Peng
Xiamen University of Technology, China

At present, most unsupervised abnormal behavior detection method only relies on powerful behavior detection classifiers, does not make full use of prior knowledge. This method often has the problem of a huge amount of calculation and affecting the detection speed. In view of the above problems, this paper proposes a weak anomaly-reinforced autoencoder for unsupervised anomaly detection method, using U-Net to reconstruct video frames and generative adversarial network to learn the correlation between image entropy and abnormal behavior. Comprehensive experiments on the avenue data set and UCSD data sets verify the effectiveness of our method to detect abnormal events.

13:30-13:45 V014

Automated Synthetic Datasets Construction for Part Semantic Segmentation of Non-Cooperative Satellites

Hongkai Ding, Jianjun Yi, Zhuoran Wang, Yajun Zhang, Hailei Wu, Shuqing Cao
East China University of Science and Technology, Shanghai, China

Part semantic segmentation based on deep learning provides a new insight for accurate vision understanding of non-cooperative satellite as well as for further on-orbit servicing tasks like inspection, repair, and close-proximity robotic manipulation. However, carrying out such researches requires a tremendous amount of data, which is extremely hard and expensive in space. Moreover, the manual annotation for fine-grained tasks like segmentation will cost a lot of labor. Thus, in this paper, we present an efficient method of automated synthetic datasets construction for part-level segmentation of non-cooperative satellite, which is capable of generating thousands of multi-source data (RGB image and point cloud) and the corresponding high-quality annotation. Specifically, the Fibonacci lattice is used for multiple viewpoints sampling of the virtual camera to capture RGB-D images. A trick of segmentation of the customized image in HSV color space is applied to get labels automatically. Furthermore, we employ several data augmentation techniques to expand and diversify the datasets, which improves the generalization of the algorithm. Finally, we carry out the case study using the pointnet++ network based on our generated point cloud data, to validate the feasibility and effectiveness of our method.

13:45-14:00 V040

An Application Of Geometric Aspects Of Variational Autoencoder Model To Forgery Detection Of Scanned Documents

Igor Janiszewski, Dmitry Slugin and Elena Andreeva
Federal Research Center Computer Science and Control of Russian Academy of Sciences, Russia

Friday, 6 Nov. 2020

The paper proposes an approach for matching of digitized copies of business documents. This task arises when comparing two versions of the same document – genuine and forgery – to find possible modifications, for example in the banking sector during the conclusion of contracts in paper form to avoid possible fraud. The matching method of two documents based on comparison images of text lines using Variational Autoencoder (VAE) trained on genuine images and calculation Fisher information metric to find modifications. Experiments were conducted on the public Payslips dataset (in French). The results show the high quality and reliability of finding document forgeries and are compared to the results of the method which applies OCR and image matching.

14:00-14:15 V067

Fader Networks for domain adaptation on fMRI: ABIDE-II study

Marina Pominova, Ekaterina Kondrateva, Maksim Sharaev, Alexander Bernstein and Evgeny Burnaev
Skolkovo University of Science and Technology, Russia

ABIDE is the largest open-source autism spectrum disorder database with both fMRI data and full phenotype description. These data were extensively studied based on functional connectivity analysis as well as with deep learning on raw data, with top models accuracy close to 75% for separate scanning sites. Yet there is still a problem of models transferability between different scanning sites within ABIDE. In the current paper, we for the first time perform domain adaptation for brain pathology classification problem on raw neuroimaging data. We use 3D convolutional autoencoders to build the domain irrelevant latent space image representation and demonstrate this method to outperform existing approaches on ABIDE data.

14:15-14:30 V085

Occlusion Aware Unsupervised Learning of Optical Flow from Video

Jianfeng Li, Junqiao Zhao, Tiantian Feng
Tongji University, China

In this paper, we proposed an unsupervised learning method for estimating the optical flow between video frames, especially to solve the occlusion problem. Occlusion is caused by the movement of an object or the movement of the camera, defined as when certain pixels are visible in one video frame but not in adjacent frames. Due to the lack of pixel correspondence between frames in the occluded area, incorrect photometric loss calculation can mislead the optical flow training process. In the video sequence, we found that the occlusion in the forward (t→t+1) and backward (t→t-1) frame pairs are usually complementary. That is, pixels that are occluded in subsequent frames are often not occluded in the previous frame and vice versa. Therefore, by using this complementarity, a new weighted loss is proposed to solve the occlusion problem. Our method achieves competitive optical flow accuracy compared to the baseline and some supervised methods on KITTI and Sintel benchmarks.

Friday, 6 Nov. 2020

Session 15
Computer Vision and Visualization

Virtual Room B:
680 4475 3785

Chair:

13:00-13:15 V039

How Good MVSNets Are at Depth Fusion

Oleg Voynov, Alexander Safin, Savva Ignatyev and Evgeny Burnaev
Skolkovo Institute of Science and Technology, Russia

We study the effects of the additional input to deep multi-view stereo methods in the form of low-quality sensor depth. We modify two state-of-the-art deep multi-view stereo methods for using with the input depth. We show that the additional input depth may improve the quality of deep multi-view stereo.

13:15-13:30 V074-A

Combining Compressed Sensing and Deep Learning to create a more efficient Machine Vision setup

Alexander Birk, Karsten Frenner, Wolfgang Osten
University of Stuttgart, Germany

Compressed Sensing and Single Pixel Cameras, which combine structured illumination and a single photodiode to acquire images, have been an area of significant interest in recent years. As one result of these efforts, it was shown that leveraging Machine Learning techniques in measurement and reconstruction could significantly improve on the results of classic Compressed Sensing methods. At the same time, increasing automation necessitates advances in Machine Vision. Bringing these together, we present a novel Machine Vision concept based around measuring as few data points as possible. We achieve this by taking the classic Compressed Sensing setup and combining it with Deep Learning based measurement optimization and information reconstruction. Furthermore, we do not reconstruct an image, but instead directly deliver the data subsequent algorithms in autonomous systems such as path planning require, saving measurements and total processing power needed. We describe how to build a sufficiently large dataset of simulated images to train the Neural Networks with and how to enforce a degree of robustness in the resulting model. We train and verify favorable results and that those can be optimized further by switching to a recurrent architecture.

the proposed method on generic random objects like ellipses and rectangles with 5 degrees of freedom in placement and pose. Our results show that a relatively small and simple multilayer perceptron architecture can already deliver

13:30-13:45 V066

Domain Shift in Computer Vision models for MRI data analysis: An Overview

Ekaterina Kondrateva, Marina Pominova, Elena Popova, Maksim Sharaev, Alexander Bernstein and Evgeny Burnaev
Skolkovo Institute of Science and Technology, Russia

Machine learning and computer vision methods are showing good performance in medical imagery analysis. Yet only a few applications are now in clinical use and one of the reasons for that is poor transferability of the models to data from different sources or acquisition domains. Development of new methods and algorithms for the transfer of training and adaptation of the domain in multi-modal medical imaging data is crucial for the development of accurate models and their use in clinics. In present work, we overview methods used to tackle the domain shift problem in machine learning and computer vision. The algorithms discussed in this survey include advanced data processing, model architecture enhancing and featured training, as well as predicting in domain invariant latent space. The application of the autoencoding neural networks and their domain-invariant variations are heavily discussed in a survey. We observe the latest methods applied to the magnetic resonance imaging (MRI) data analysis and conclude on their performance as well as propose directions for further research.

13:45-14:00 V076

DOME-T: Adversarial computer vision attack on deep learning models based on Tchebichef image moments

T. Maliamanis and G.A. Papakostas
International Hellenic University, Greece

In this paper, a novel black box adversarial computer vision attack is proposed. The introduced attack is based on removing from images some components described by their Tchebichef discrete orthogonal moments, rather than to perturb them. The contribution of this work is focused on the addition of one more clue, supporting the critical hypothesis that computer vision systems fail because they support their decisions not only in robust features but also in others non-robust ones.

Friday, 6 Nov. 2020

In this, context non-robust image features described in terms of Tchebichef moments are excluded from the original images and the approximated reconstructed versions of them are used as adversarial examples in order to attack some popular deep learning models. The experiments justify the effectiveness of the proposed adversarial attack in terms of imperceptibility and recognition error rate of the deep learning classifiers. It is worth noting that the top-1 accuracy of the attacked models was degraded by a factor between 9.48%-70.89% for adversarial images of 65dB to 57dB PSNR values. The corresponding degradation of the top-5 models' accuracy was between 6.9% and 55.14% for the same quality images. Moreover, the proposed attack seems to have more strength than the Fast Gradient Sign Method (FGSM) attacking method traditionally applying in most cases. These results reveal that the proposed attack is able to exploit the vulnerability of the deep learning models' towards degrading their generalization abilities.

14:00-14:15 V007

Scalable Multi-View Stereo using CMA-ES and Distance Transform-based Depth Map Refinement

Nirmal S. Nair, Madhu S. Nair
University of Kerala, India

Recovering three-dimensional structure from images is a long-standing ill-posed inverse problem in computer vision. This paper presents a simple and highly scalable method to reconstruct dense 3D point cloud from multi-view images by estimating per-pixel depth using an evolutionary computation technique – CMA-ES. The proposed method uses ZNCC-based template matching to reconstruct fine details of textured regions and DAISY-based feature matching to reconstruct smooth surface of homogeneous regions. We handle the problem of reconstructing large homogeneous regions using distance transform-based adaptive median filtering. The proposed method is highly scalable since pixels are processed independently at all stages of reconstruction – depth map estimation, refinement, and fusion. This enables the proposed method to be parallelized at the pixel-level, unlike most existing methods that can only be parallelized at the image-level. Experimental results on Middlebury benchmark dataset demonstrate the robustness and efficacy of the proposed method in reconstructing textured as well as homogeneous regions.

14:15-14:30 V048

Real-time Vineyard Trunk Detection for a Grapes Harvesting Robot via Deep Learning

Eftichia Badeka, Theofanis Kalampokas, Eleni Vrochidou, Konstantinos Tziridis, George A. Papakostas, Theodore Pachidis, Vassilis G. Kaburlasos
International Hellenic University (IHU), Greece

Research and development in agricultural robots are continuously increasing. However, dynamically changing agricultural environments provide adverse conditions to robotics operability. In order to perform the agricultural tasks safely and accurately, reliable landmarks from the surrounding environment need to be identified. In this work, deep learning is employed for accurate and fast detection of high-level features of vineyards, the vine trunks. More specifically, Faster regions-convolutional neural network (Faster R-CNN), You Only Look Once version 3 (YOLOv3) and YOLOv5 are tested for real-time vine trunk detection. The models are trained with an in-house dataset designed for the needs of this study, containing 1927 annotated vine trunks in 899 different images. Comparative results indicate YOLOv5 as the configuration that allows the faster and most accurate vine trunk detection, achieving an overall Average Precision of 73.2% in 29.6 ms. The high precision combined with the fast runtime performance prove that the YOLOv5 detector is suitable for real-time vine trunk detection executed by an autonomous harvesting robot.

Friday, 6 Nov. 2020

Session 16
Computer Photography and Image Processing

Virtual Room A:
976 5819 4637

Chair:

15:00-15:15 V097

Bipolar Morphological U-Net for Document Binarization

Elena Limonova, Dmitry Nikolaev and Vladimir Arlazarov
Institute for Systems Analysis, FRC CSC RAS, Moscow, Russia

Deep neural networks are widely used in various AI systems. Many such systems rely on the edge computing concept and try to perform computations on end devices while still being energy and memory efficient. Therefore, substantial time and memory requirements are imposed on neural networks. One way to improve neural network efficiency is to simplify computations inside a neuron. A bipolar morphological neuron uses only addition, subtraction, and maximum operations inside the neuron and exponent and logarithm as activation functions for the network layers. These operations allow fast and compact gate implementation for FPGA and ASIC. In the paper, we consider the usage of bipolar morphological (BM) networks for document binarization. We examine the DIBCO 2017 binarization challenge and train the bipolar morphological convolutional neural network of U-Net architecture. Despite some accuracy decrease for a model with all BM convolutional layers, one can flexibly control the accuracy by using the partially converted model. It should be noted that even the fully BM model is suitable for solving the problem in practice.

15:15-15:30 V079

An instance segmentation framework for in-situ plankton taxa assessment

Aya Saad, Sondre Bergum and Annette Stahl
The Norwegian University of Science and Technology, Norway

In this paper, we propose a deep learning instance segmentation framework for particle extraction of microscopic images that aims at calculating planktonic species distribution and concentration in-situ. The framework comprises three essential functional tasks on in-situ time-series images collected from an autonomous underwater vehicle: 1) manual labeling of the captured images, 2) object localization, segmentation, and identification, and 3) class distribution and planktonic organisms concentration calculation. Our proposed framework is based on the mask R-CNN architecture provided by the Detectron2 library developed by Facebook Artificial Intelligence Research (FAIR) for instance segmentation. Due to its modular design, we compare the performance of different networks by alternating the backbone sub-network in order to choose the most suitable architecture for the task of instance and semantic segmentation. We compile a custom annotated dataset from planktonic time-series images and train the different models over this dataset to perform the instance semantic segmentation. Evaluation results of the proposed framework, utilizing the best performing deep learning architecture along with the new annotated dataset, show better performance in terms of speed and accuracy of both in-situ segmentation and classification compared to traditional segmentation methods. In addition, we observe a significant improvement in the object classification quality when we train the model over our newly annotated dataset instead of training it over the dataset generated from the traditional methods. The inferred data from our novel instance segmentation framework, which provides the particle class distribution and concentration, can then be used to assist in constructing a dynamic probability density map of planktonic communities dispersion and abundance.

15:30-15:45 V080

Thermal Image Processing for Feature Extraction from Encapsulated Phase Change Materials

Brian Whinery, Yuri Gulak, Vedang Chauhan, Jingzhou Zhao, Jingru Benner, Feng Ye
Western New England University, US

Encapsulated inorganic particles with high melting points (>300 °C) are desired as high-temperature Phase Change Materials (PCMs) for next-generation Latent Heat Thermal Energy Storage (LHTES) systems. One of the many challenges during the development of PCMs is to achieve a high throughput that in turn depends on accurately modeling the relation between process parameters and geometric & thermal properties of the PCMs particle. During the production of the PCMs, a high-speed infrared camera is used to acquire images of the encapsulated material under controlled illumination conditions. This research article focuses on the development of image processing techniques for both geometric and thermal feature extraction during the development of the PCMs. A user-friendly GUI has been designed in MATLAB and preliminary experimental results have demonstrated that the method is fast, accurate and reliable for a high throughput production. The extracted features will be used to develop Machine Learning (ML) models to predict the geometric and thermal properties of the PCM based on the process parameter settings. The ML model will accelerate the search for the optimized process settings to boost the throughput of the production.

Friday, 6 Nov. 2020

15:45-16:00 V081

Application of Fractional Bio-inspired Filter for Salient Color Detection

Juan Anaya-Jaimes, Angie Garcıá -Castro, J.A. Tenreiro-Machado and R.E. Gutiérrez-Carvajal
Universidad Militar Nueva Granada, Colombia .

Computational modeling of visual attention has been a research field focused on emulating the behavior of biological visual systems in a given scenario, by using mechanisms developed for fixation prediction or salient region detection. In the literature, different approaches have been presented to emulate the interactions that occur in the early vision system of biological structures. However, mathematical modeling of these systems applying theories related to fractional operators could outperform the existing models. In this paper, we present a fractional bio-inspired filter for salient color detection in natural scenarios, based on the behavior and distribution of the cone photoreceptors cells in the retina. The filter was compared with two classic saliency algorithms over a natural color image dataset in terms of saliency prediction and processing time, using a Similarity (SIM) score and runtime performance, respectively. Our approach reach the second best result in therms of saliency prediction with a 48,9% of SIM with ground truth fixations maps and the fastest time response, with an average time of 0.12 s when processing a high resolution image, being 25% faster than Itti et al. algorithm, one of the most applied in robotic vision tasks.

16:00-16:15 V093

DNNs for Multi-Map Semantic Segmentation

Mr. Pavel Kurnikov and Dmitry Sholomov
Cognitive Robotics Ltd., Russia

Modern vehicles include a vast number of intellectual functions such as lane-keeping assist (LKA), vehicle, pedestrian and obstacle recognition (FCW, PPS) which are implemented in the advanced driver assistance system (ADAS). These functions allow a vehicle to localize itself correctly within the road lane and to increase the overall system safety. It is also critical for vehicle motion and planning the target trajectory. Previously, algorithms implementing ADAS functions were based on classical computer vision approaches (e.g. edge detection, morphology, Hough transform), which did well only on a rather simple road scene. Modern state-of-the-art systems are based on semantic segmentation networks, it is the unquestionable trend. With a more thorough study of the road scene segmentation issues we face the problem that the existing benchmark suites such as MOTS, KITTI as well as recent DNNs for the road/lane semantic segmentation employ only mutually exclusive classes i.e. in this case, a pixel can belong to a single class only. But if we recognize the road scene, a pixel can easily refer to several classes, e.g. to ego-lane and crosswalk. The classical approach with mutually exclusive classes will give preference to only one class in this case and we will get an ego-lane consisting of two components. As a result, it may be difficult to restore the ego-lane at the stage of post-processing, see Figure 1. To overcome this problem, in the paper we propose the approach with multiple segmentation maps as an output of the DNN architecture, as well as a " multi-map loss function. In this case, each pixel is referred to several classes at the same time (depending on the number of layers) and we don’t have the restriction to use mutually exclusive classes. The DNN classifier for each segmentation map has a separate activation branch and loss function.

Friday, 6 Nov. 2020

Session 17
Intelligent Image Analysis and Key Technologies

Virtual Room B:
680 4475 3785

Chair: Andrey Kuznetsov — Samara National Research University, Russia

15:00-15:15 V098

Fast and Accurate Mobile-Aided Screening System of Moderate Diabetic Retinopathy

Yaroub Elloumi, Manef Ben Mbarek, Rahma Boukadida, Mohamed Akil, Mohamed Hedi Bedoui
University of Monastir, Tunisia

The Diabetic Retinopathy (DR) is a worldwide eye disease that causes visual damages and can leads to blindness. Therefore, the detection of the DR in the early stages is highly recommended. However, a delay is registered for ensuring early DR diagnosis which caused by the low-rate of the ophthalmologists, the deficiency of diagnosis equipment and the lack of mobility of elderly patients.
In this paper, the main objective is to provide a mobile-aided screening system of moderate DR. Within this aim, we propose a classifier-based method which is based on detecting the Hard Exudate (HE) lesions that occur in moderate DR stage. A set of features are extracted to ensure an accurate and robust detection with respect to modest quality of fundus images. Moreover, the detection is provided in a low complexity processing to be suitable for mobile device. The aimed system corresponds to the implementation of the method on a smartphone associated to an optical lens for capturing fundus image. The system reached satisfactory screening performance where an accuracy of 98.36%, a sensitivity of 100% and specificity of 96.45% are registered using the DIARETDB1 fundus image databases. Moreover, the screening is performed in an average execution time of 2.68 seconds.

15:15-15:30 V058

Shape-aware Generative Adversarial Networks for Attribute Transfer

Lei Luo, William Hsu, Shangxian Wang
Kansas State University, United States

Generative adversarial networks (GANs) have been successfully applied to transfer visual attributes in many domains, including that of human face images. This success is partly attributable to the facts that human faces have similar shapes and the positions of eyes, noses, and mouths are fixed among different people. Attribute transfer is more challenging when the source and target domain share different shapes. In this paper, we introduce a shape-aware GAN model that is able to preserve shape when transferring attributes, and propose its application to some real-world domains. Compared to other state-of-art GANs-based image-to-image translation models, the model we propose is able to generate more visually appealing results while maintaining the quality of results from transfer learning.

15:30-15:45 V011

Robust Technique for Representative Volume Element Identification in Noisy Microtomography Images of Porous Materials Based on Pores Morphology and Their Spatial Distribution

Grigoriev M., Khafizov A., Kokhan V., Asadchikov V.
FSRC "Crystallography and photonics" RAS, Russia

Microtomography is a powerful method of materials investigation. It enables to obtain physical properties of porous media non-destructively that is useful in studies. One of the application ways is a calculation of porosity, pore sizes, surface area, and other parameters of metal-ceramic (cermet) membranes which are widely spread in the filtration industry. The microtomography approach is efficient because all of those parameters are calculated simultaneously in contrast to the conventional techniques. Nevertheless, the calculations on Micro-CT reconstructed images appear to be time-consuming, consequently representative volume element should be chosen to speed them up. This research sheds light on representative elementary volume identification without consideration of any physical parameters such as porosity, etc. Thus, the volume element could be found even in noised and grayscale images. The proposed method is flexible and does not overestimate the volume size in the case of anisotropic samples. The obtained volume element could be used for computations of the domain's physical characteristics if the image is filtered and binarized, or for selections of optimal filtering parameters for denoising procedure.

15:45-16:00 V064

Memory consumption reduction for identity document classification with local and global features combination

Natalya Skoryukina, Vladimir Arlazarov, Artemiy Milovzorov
Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Russia

Friday, 6 Nov. 2020

In this paper we explore possibilities of memory cost reduction without significant loss of classification accuracy in connection with the problem of the ID document type recognition on mobile devices. The studied classic approach is based on representing images using constellation of feature points and descriptors. The distortion parameters are estimated by applying RANSAC. Experimental data details the approach limitations (memory, speed and accuracy) in dependence of the descriptor type. In order to maintain accuracy when using low dimensional descriptors we suggest to modify the basic approach using additional features characteristic of the document such as straight lines and quadrangles. In addition, an early filtration of the samples and the hypotheses used in RANSAC. It was shown that the proposed modifications have a positive contribution for all types of descriptors considered. The suggested algorithm was tested using the open dataset MIDV-500. The modified approach allows to achieve an accuracy improvement and significant speed up of distortion parameters estimation in RANSAC. It was shown that using compact descriptors in conjunction with the presented method allows reduce required memory cost by more than 7 times with near-zero (0.2%) loss of accuracy, and more than 14 times with the loss of accuracy is about 18%.

16:00-16:15 V063

Language of Gleam: Impressionism Artwork Automatic Caption Generation for People with Visual Impairments

Dongmyeong Lee, Hyegyeong Hwang, Muhammad Shahid Jabbar, Jun-Dong Cho
Sungkyunkwan University, South Korea

User Experience Design (UX Design) comes from focusing on how products, in reality, affect the user's experience. In particular, the design of multi-modal interfaces for blind people facilitates the flexible and natural product or service capacity and improves blind people's interaction by overcoming the various existing constraints associated with any particular interaction. There have been various attempts to help visually impaired people appreciation of visual artwork, including multi-modal associations. However, these methods can only provide general information in terms of edge and pattern recognition by the sense of touch and restrained by the availability and number of specially developed artworks. We propose a novel method explaining visual artworks through image caption generation using artificial intelligence (AI) to improve artwork accessibility. This method can objectively describe any impressionism artwork used as a standalone description of art interpretation for blind people or can aide tactile-based methods. Based on end-to-end learning with a deep neural network, an encoder-decoder architecture model is adopted, and comprehensive experiments perform to confirm the stability of generated image captioning for stylized MS-COCO datasets with impressionism.

16:15-16:30 V4007

Feathers dataset for Fine-Grained Visual Categorization

Alina Belko, Konstantin Dobratulin, Andrey Kuznetsov
Samara University, Russia

This paper introduces a novel dataset FeatherV1, containing 28,272 images of feathers categorized by 595 bird species. It was created to perform taxonomic identification of bird species by a single feather, which can be applied in amateur and professional ornithology. FeatherV1 is the first publicly available bird’s plumage dataset for machine learning, and it can raise interest for a new task in fine-grained visual recognition domain. The latest version of the dataset can be downloaded at https://github.com/feathers-dataset/feathersv1-dataset. We also present feathers classification task results. We selected several deep learning architectures (DenseNet based) for categorical crossentropy values comparison on the provided dataset.

Friday, Nov. 6, 2020
16:30-17:00
ZOOM ID: 976 5819 4637

Closing Ceremony

THANKS DELEGATES