POSTER SESSION BOOKLET

http://www.dmi.unict.it/icvss

University of Catania - University of Cambridge

International Computer Vision Summer School 2019

Computer Vision - Where are we?

Sicily, 7 - 13 July 2019

International Computer Vision Summer School

Computer Vision is the science and technology of making machines that see. It is concerned with the theory, design and implementation of algorithms that can automatically process visual data to recognize objects, track and recover their shape and spatial layout.

The International Computer Vision Summer School - ICVSS was established in 2007 to provide both an objective and clear overview and an in-depth analysis of the state-of-the-art research in Computer Vision. The courses are delivered by world renowned experts in the field, from both academia and industry, and cover both theoretical and practical aspects of real Computer Vision problems.

The school is organized every year by University of Cambridge (Computer Vision and Robotics Group) and University of Catania (Image Processing Lab). The general entry point for past and future ICVSS editions is:

http://www.dmi.unict.it/icvss

ICVSS Poster Session

The International Computer Vision Summer School is especially aimed to provide a stimulating space for young researchers and Ph.D. Students. Participants have the possibility to present the results of their research, and to interact with their scientific peers, in a friendly and constructive environment.

This booklet contains the abstract of the posters accepted to ICVSS 2019.

Best Presentation Prize A subset of the submitted posters will be selected by the school committee for short oral presentation. A best presentation prize will be given to the best presentations selected by the school committee.

Scholarship A scholarship will be awarded to the best PhD student attending the school. The decision is made by the School Committee at the time of the School, taking into account candidates' CV, poster and oral presentation.

Sicily, June 2019

Roberto Cipolla
Sebastiano Battiato
Giovanni Maria Farinella

List of Posters [1]

1. UNSUPERVISED DOMAIN ADAPTATION FOR TOF DATA DENOISING Agresti G., Schaefer H., Sartor P., Zanuttigh P.

2. VOLUMETRIC SEGMENTATION OF TUBULAR NETWORKS Arnavaz K., Feragen A., Nyeng P., Krause O.

3. THE SURPRISING EFFECTIVENESS OF FEW-IMAGE UNSUPERVISED FEATURE LEARNING Asano YM., Rupprecht C., Vedaldi A.

4. REAL-TIME 3D LIDAR FLOW FOR AUTONOMOUS VEHICLES Baur Stefan A., Moosmann Frank, Wirges Sascha, Rist Christoph

5. AUTOMATED KNOT-TYING EVALUATION WITH POSE DATA Bhattacharya Jhilik, Lacey Gerard

6. GOOD NEWS EVERYONE! CONTEXT DRIVEN ENTITY-AWARE CAPTIONING FOR NEWS IMAGES. Biten Ali Furkan, Gómez Lluis, Rusinol Marcal, Karatzas Dimosthenis.

7. MAXIMALLY COMPACT AND SEPARATED FEATURES WITH REGULAR POLYTOPE NETWORKS Bruni M., Pernici F., Baecchi C., Del Bimbo A.

8. SOURCE CAMERA IDENTIFICATION ON STABILIZED DIGITAL VIDEO Bruno A., Cattaneo G.

9. DECOUPLING FEATURE EXTRACTION FROM POLICY LEARNING : ASSESSING BENEFITS OF STATE REPRESENTATION LEARNING IN GOAL BASED ROBOTICS Raffin Antonin, Hill Ashley, Traoré René, Lesort Timothée, Díaz-Rodríguez Natalia, Filliat David

10. LEARNING TO EXPLORE 3D ENVIRONMENTS Ermolov A.

[1] Posters are ordered by surname of the speaker. Each poster is identified by a number. The page of a poster in this booklet corresponds with the ID of the poster.

11. EFFICIENT MERGING OF MAPS AND DETECTION OF CHANGES Flood G., Gillsjö D., Heyden A., Åström K.

12. WHAT WOULD YOU EXPECT? ANTICIPATING EGOCENTRIC ACTIONS WITH ROLLING-UNROLLING LSTMS AND MODALITY ATTENTION Furnari A., Farinella, G. M.

13. BEAUTY AND THE BEAST: OPTIMAL METHODS MEET LEARNING FOR DRONE RACING Kaufmann E., Gehrig M., Foehn P., Ranftl R., Dosovitskiy A., Koltun V., Scaramuzza D.

14. END-TO-END LEARNING OF REPRESENTATIONS FOR ASYNCHRONOUS EVENT BASED DATA Gehrig D., Loquercio A., Derpanis K. G., Scaramuzza D.

15. SIAMESE BALLISTICS NEURAL NETWORK Giudice O., Guarnera L., Paratore A. B., Farinella G. M. , Battiato S.

16. A NEW STUDY ON WOOD FIBER TEXTURE: DOCUMENTS AUTHENTICATION THROUGH LBP FINGERPRINT Guarnera F.

17. DETECTFUSION: DYNAMIC OBJECT-LEVEL SLAM IN REAL-TIME Hachiuma R., Pirchheim C., Schmalstieg D., Saito H.

18. SEMANTIC UNDERSTANDING OF FOGGY SCENES WITH PURELY SYNTHETIC DATA Hahner M., Dai D., Sakaridis C., Zaech J., Van Gool L.

19. VIDEO REPRESENTATION LEARNING BY DENSE PREDICTIVE CODING Han T., Xie W., Zisserman A.

20. HUMAN VISUAL ATTENTION IN IMAGE CAPTIONING Sen He,Nicolas Pugeault

21. DELF ATTENTION POOLED FEATURES Wahab F., Sulthanpete Govindappa K., Heße L.

22. MULTIVIEW ABSOLUTE POSE USING 3D - 2D PERSPECTIVE LINE CORRESPONDENCES AND VERTICAL DIRECTION Horanyi N.

23. ATLAS-BASED SKULL SEGMENTATION IN MRI USING STATISTICAL SHAPE MODELS Kahr P.

24. DOMAIN ADAPTATION FOR PRIVACY-PRESERVING PEDESTRIAN DETECTION IN THERMAL IMAGERY My Kieu, Andrew D. Bagdanov, Marco Bertini, Alberto D. Bimbo

25. PROGRESSIVE ATTENTION MEMORY NETWORK FOR MOVIE STORY QUESTION ANSWERING Kim J., Ma M., Kim K., Kim S., Yoo C. D.

26. SUBMODULAR BATCH SELECTION FOR TRAINING DEEP NEURAL NETWORKS Joseph K.J., Vamshi T.R., Krishnakanth S., Balasubramanian V.N.

27. EMBODIED VISION-AND-LANGUAGE NAVIGATION WITH DYNAMIC CONVOLUTIONAL FILTERS Landi F., Baraldi L., Corsini M., Cucchiara R.

28. 3D APPEARANCE SUPER-RESOLUTION WITH DEEP LEARNING Li Yawei , Tsiminaki Vagia, Timofte Radu, Pollefeys Marc, van Gool Luc

29. UNSUPERVISED MOVING OBJECT DETECTION VIA CONTEXTUAL INFORMATION SEPARATION Yang Y., Loquercio A., Scaramuzza D., Soatto S.

30. INTERPRETABILITY BEYOND CLASSIFICATION OUTPUT: SEMANTIC BOTTLENECK NETWORKS Losch M., Fritz M., Schiele B.

31. VIDEO OBJECT SEGMENTATION: PREMVOS VS BOLTVOS Luiten J., Voigtlaender P., Leibe B.

32. PATTERNS OF 3D FACIAL SEXUAL DIMORPHISM: A TRIPLET LOSS APPROACH Mahdi S., Nauwelaers N., Bouritsas G., Bokhnyak S., Bronstein M., Zafeiriou S., Boets B., Claes P.

33. GRADIENT BOOSTING ON DECISION TREES FOR MORTALITY PREDICTION IN TRANSCATHETER AORTIC VALVE IMPLANTATION Mamprin M., Zelis J.M., Tonino P.A.L., Zinger S., de With P.H.N.

34. DEEP MODEL-BASED 6D POSE REFINEMENT IN RGB Manhardt F., Kehl W., Navab N, Tombari F.

35. NEURAL NETWORKS FOR VIDEO INTRA PREDICTION Meyer M., Schneider J., Rohlfing C.

36. IMPLICIT REPRESENTATION FOR 3D SHAPE INFERENCE Michalkiewicz M. , Pontes J. , Baktashmotlagh M. , Eriksson A.

37. UNSUPERVISED DOMAIN ADAPTATION FOR URBAN SCENES SEGMENTATION Biasetton M., Michieli U., Agresti G., Zanuttigh P.

38. MARKERLESS HUMAN MOTION ANALYSIS ON RGB VIDEOS Moro M., Odone F., Casadio M.

39. DEEP METRIC LEARNING FOR DENSE CORRESPONDENCES Nickel M., Dhamo H., Tombari F., Navab N.

40. BEYOND ACTION RECOGNITION Nicora E., Goyal G., Noceti N., Odone F.

41. FIRST PERSON VISION FOR VISITORS BEHAVIOURAL UNDERSTANDING AND FRUITION AUGMENTATION IN CULTURAL SITES Orlando S. A., Furnari A., Gallo G., Farinella G. M.

42. GAUSSIAN BELIEF PROPAGATION FOR ESTIMATION IN SLAM Ortiz J., Davison A.

43. SMPLR: LEARNING A COMPACT, EFFICIENT, AND ACCURATE MODEL OF 3D HUMAN BODY SHAPE Osman Ahmed, Black J Michael

44. IMPACT OF ADVERSARIAL EXAMPLES ON DEEP LEARNING BIOMEDICAL SEGMENTATION MODELS Ozbulak U., Van Messem A., De Neve W.

45. COLORECTAL POLYPS DETECTION: WHERE ARE WE? Pappalardo G., Cherubini A., Farinella G.M.

46. DEPTH-BASED HUMAN ANALYSIS Pini S., Borghi G., Vezzani R., Cucchiara R.

47. DEEP NEURAL NETWORK BASED COOPERATIVE VISUAL TRACKING THROUGH MULTIPLE MICRO AERIAL VEHICLES Price E., Lawless G., Ludwig R., Martinović I., Bülthoff H. H., Black M. J., Ahmad A.

48. UNICT-VEDI: DATASET AND CHALLENGES FOR VISITORS BEHAVIORAL UNDERSTANDING FROM EGOCENTRIC VISION Ragusa F., Furnari A., Battiato S., Signorello G., Farinella G.M.

49. OUTLIER DETECTION AND ROBUST FITTING WITH STATISTICAL SHAPE MODELS Rahbani D., Morel-Forster A., Madsen D., Lüthi M., Vetter T.

50. ROBOT NAVIGATION IN REAL INDOOR ENVIRONMENTS USING REINFORCEMENT LEARNING Rosano M., Furnari A., Farinella G. M.

51. UNSUPERVISED DOMAIN ADAPTATION USING FEATURE-WHITENING AND CONSENSUS LOSS Roy S., Siarohin A., Sangineto E., Bulò SR., Sebe N. and Ricci E.

52. CRITICAL TRAINING SET SIZE AND DROPOUT KEEP PROBABILITY FOR LENET4 CONVOLUTIONAL NEURAL NETWORK TRAINED ON MNIST. IS IT PERCOLATION? Sala V.

53. AUDIO-VISUAL MODEL DISTILLATION USING ACOUSTIC IMAGES Pérez A.F., Sanguineti V., Morerio P., Murino V.

54. PRIVACY PROTECTION IN STREET-VIEW PANORAMAS USING DEPTH AND MULTI-VIEW IMAGERY Uittenbogaard R., Sebastian C., Vijverberg J., Boom B., Gavrila D.M., de With P.H.N

55. DEEP BUILT-STRUCTURE COUNTING IN SATELLITE IMAGERY USING ATTENTION BASED RE-WEIGHTING Shakeel Anza., Sultani Waqas., Ali Mohsen.

56. LOCAL FEATURES AND VISUAL WORDS EMERGE IN ACTIVATION Oriane S., Yannis A., Ondra C.

57. FACE DETECTION IN THE OPERATING ROOM: A COMPARISON OF STATE-OF-THE-ART METHODS AND A SELF-SUPERVISED APPROACH Issenhuth Thibaut, Srivastav Vinkle, Gangi Afshin, Padoy Nicolas

58. DEEP LEARNING FOR GASTRIC PATHOLOGY DETECTION IN ENDOSCOPIC IMAGES Stefanidi A.

59. ASSISTED TARGET DETECTION IN SEARCH AND RESCUE Taheri-Shirazi, M., Ehinger, K. A., & Elder, J. H.

60. VRPROP-NET: TRANSPORTING REAL OBJECTS INTO VIRTUAL AND AUGMENTED ENVIRONMENTS Taylor C., McNicholas R., Cosker D.

61. ART2REAL: UNFOLDING THE REALITY OF ARTWORKS VIA SEMANTICALLY AWARE IMAGE-TO-IMAGE TRANSLATION Tomei M., Cornia M., Baraldi L., Cucchiara R.

62. AUTOMATED TISSUE SEGMENTATION OF MICRO-CT IMAGES BY MACHINE LEARNING AND ITS APPLICATION TO COMPARATIVE MORPHOLOGY Toulkeridou E., Baum D., Doya K., Economo E.

63. COMPUTER VISION-BASED SYSTEM FOR ADAS Trenta Francesca

64. SYNTHESIS OF REALISTIC HIGH-RESOLUTION LIDAR DATA Triess L., Peter D.

65. QT BASED ROBOTIC VISION SOFTWARE: SMART VISION Temürlenk Y., Taşdelen A., Kaya A., Tükel D.B

66. ISOLATED SIGN RECOGNITION WITH A SIAMESE NEURAL NETWORK OF RGB AND DEPTH STREAMS Tur A. O., Yalim Keles H.

67. DECIDING WHAT LAYERS TO SHARE IN MULTI-TASK NETS Vandenhende S., Georgoulis S., De Brabandere B., Van Gool L.

68. SCENE UNDERSTANDING FOR SELF-DRIVING CARS Van Gansbeke W., De Brabandere B., Neven D., Van Gool L.

69. MOTS: MULTI-OBJECT TRACKING AND SEGMENTATION Voigtlaender P., Krause M., Osep A., Luiten J., Gnana B., Geiger A., Leibe B.

70. MACHINE LEARNING FOR 3D RECONSTRUCTION Weilharter R., Fraundorfer F.

71. 6DOF VEHICLE POSE ESTIMATION FOR MONOCULAR RGB IMAGES WU DI

72. PHOTO-GEOMETRIC AUTOENCODING TO LEARN 3D OBJECTS FROM UNLABELLED IMAGES Wu S., Rupprecht C., Vedaldi A.

73. SCSLAM - ON-THE-FLY SCENE COMPLETION WITH RGBD SLAM Wu S., Tateno K., Tombari F., Navab N.

74. DEEP LEARNING FOR MATERIAL CLASSIFICATION BASED ON LOCAL AND GLOBAL FEATURES Xu S., Tremeau A., Muselet D.

75. AUTOMATIC PAIN INTENSITY ESTIMATION FROM FACIAL VIDEOS Yang R.J., Hong X.P., Peng J.Y., Zhao G.Y.

76. STRUCTUREVO: VISUAL ODOMETRY IN STRUCTURED ENVIRONMENTS Yanyan Li, Nikolas Brasch, Federico Tombari, Nassir Navab

77. GAN-BASED TEXTURE INTERPOLATION AND GENERATED IMAGE ATTRIBUTION Yu N., Barnes C., Shechtman E., Amirghodsi S., Lukáč M., Davis L., Fritz M.

78. FONT ADAPTIVE TEXT RECOGNITION NETWORK ZHANG C, GUPTA A, ZISSERMAN A

UNSUPERVISED DOMAIN ADAPTATION FOR TOF DATA DENOISING

Agresti G., Schaefer H., Sartor P., Zanuttigh P.

Abstract: ToF depth data is affected by a high level of noise and by Multi-Path Interference corruption. Machine learning has seldom been applied for ToF denoising due to the limited availability of labeled real data. Here, a CNN is trained on synthetic data in a supervised way and an adversarial learning strategy is used to perform unsupervised domain adaptation to real world data, avoiding the domain shift issue. Experimental results demonstrate that our method outperforms state-of-the-art techniques.

Contact: gianluca.agresti@phd.unipd.it

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

VOLUMETRIC SEGMENTATION OF TUBULAR NETWORKS

Arnavaz K., Feragen A., Nyeng P., Krause O.

Abstract: The transportation system in a mature pancreas has a branched tree-like structure, which is derived from a network of loops that must rearrange and close during embryonic stage. Recent biological research suggests that there might be a connection between appearance and closure of network loops and creation of insulin-producing cells. Thus to make the analysis of the structure easier, we first need to do image segmentation.

Contact: kasra@di.ku.dk

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

THE SURPRISING EFFECTIVENESS OF FEW-IMAGE UNSUPERVISED FEATURE LEARNING

Asano YM., Rupprecht C., Vedaldi A.

Abstract: State-of-the-art methods for unsupervised representation learning can train well the first few layers of standard convolutional neural networks without labels. However, these approaches are applied to millions of images, scalability being advertised as their major advantage since unlabelled data is cheap to collect. In this paper we question this practice and ask whether so many images are actually needed to learn the layers for which unsupervised learning works best. Our main result is that a few or even a single image together with strong data augmentation are sufficient to nearly saturate performance. Specifically, we provide an analysis for three different self-supervised feature learning methods (BiGAN, RotNet, DeepCluster) vs number of training images (1, 10, 1k) and show that we can top the accuracy for the first two convolutional layers of common networks using just a single unlabelled training image and obtain competitive results for other layers. We further study and visualize the learned representation as a function of which (single) image is used for training. https://arxiv.org/abs/1904.13132

Contact: yuki@robots.ox.ac.uk

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

REAL-TIME 3D LIDAR FLOW FOR AUTONOMOUS VEHICLES

Baur Stefan A., Moosmann Frank, Wirges Sascha, Rist Christoph

Abstract: Autonomous vehicles require an accurate understanding of the motion of their surroundings. We propose a novel method to estimate point-wise 3D motion vectors from LiDAR point clouds using fully convolutional networks. This motion information can also be used to efficiently estimate odometry. Our approach is capable of predicting dense 3D LiDAR flow in real-time, achieving significant speed ups over the current state of the art.

Contact: stefan_andreas.baur@daimler.com

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

AUTOMATED KNOT-TYING EVALUATION WITH POSE DATA

Bhattacharya Jhilik, Lacey Gerard

Abstract: One of the primary skills required of surgeons is high quality knot tying. The development of these skills takes time and human resources as it requires one-on-one expert supervision. Our motivation is to provide good quality automatic evaluation in order to reduce the resource requirements and provide continuous evaluations to support a deliberate practice training paradigm. Our aim is to develop automated knot tying evaluation exploiting hand pose information.

Contact: jbhattac@tcd.ie

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

GOOD NEWS EVERYONE! CONTEXT DRIVEN ENTITY-AWARE CAPTIONING FOR NEWS IMAGES.

Biten Ali Furkan, Gómez Lluis, Rusinol Marcal, Karatzas Dimosthenis.

Abstract: Current image captioning systems perform at a merely descriptive level, essentially enumerating the objects in the scene and their relations. Humans, on the contrary, interpret images by integrating several sources of prior knowledge of the world. In this work, we aim to take a step closer to producing captions that offer a plausible interpretation of the scene, by integrating such contextual information into the captioning pipeline.

Contact: abiten@cvc.uab.es

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

MAXIMALLY COMPACT AND SEPARATED FEATURES WITH REGULAR POLYTOPE NETWORKS

Bruni M., Pernici F., Baecchi C., Del Bimbo A.

Abstract: CNNs are widely used classification models for several vision tasks. Typically, a learnable transformation is placed at the end of such models returning class scores. This learnable transformation has a fundamental role in determining the network internal feature representation. In this work we show how to extract from CNNs features with the properties of maximum inter-class separability and intra-class compactness by setting the parameters of the classifier transformation as not trainable.

Contact: matteo.bruni@unifi.it

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

SOURCE CAMERA IDENTIFICATION ON STABILIZED DIGITAL VIDEO

Bruno A., Cattaneo G.

Abstract: The use of PRNU-based techniques for Source Camera Identification is a de-facto standard for photo. Digital videos presents an increasing difficulties caused in first place by the effects of digital video stabilization process. The solution proposed in literature try to solve the problem using a brute force attacks to identify the inverse transformation to nullify the effects of the stabilization are heavy computational expensive. In this work we analyze the solution of Taspinar et al. and propose a more computationally efficient solution to the problem.

Contact: andbruno@unisa.it

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

DECOUPLING FEATURE EXTRACTION FROM POLICY LEARNING : ASSESSING BENEFITS OF STATE REPRESENTATION LEARNING IN GOAL BASED ROBOTICS

Raffin Antonin, Hill Ashley, Traoré René, Lesort Timothée, Díaz-Rodríguez Natalia, Filliat David

Abstract: Scaling end-to-end reinforcement learning to control real robots from vision presents a series of challenges, in particular in terms of sample efficiency. Against end-to-end learning, state representation learning can help learn a compact, efficient and relevant representation of states that speeds up policy learning, reducing the number of samples needed, and that is easier to interpret. We evaluate several state representation learning methods on goal based robotics tasks and propose a new unsupervised model that stacks representations and combines strengths of several of these approaches. This method encodes all the relevant features, performs on par or better than end-to-end learning with better sample efficiency, and is robust to hyper-parameters change.

Contact: natalia.diaz@ensta-paristech.fr

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

LEARNING TO EXPLORE 3D ENVIRONMENTS

Ermolov A.

Abstract: We investigate the efficiency of learning-based methods for navigation inside a 3d environment. The goal of the agent is to explore the available space, recognize the structure of the rooms, learn required locomotion and logical skills. We present a modification of two state-of-the-art reinforcement learning algorithms. We evaluate our approach in a visually rich environment Obstacle Tower, processing high-dimensional input (screen images) and output (54 possible actions).

Contact: aleksandr.ermolov@unitn.it

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

EFFICIENT MERGING OF MAPS AND DETECTION OF CHANGES

Flood G., Gillsjö D., Heyden A., Åström K.

Abstract: We have developed a method for merging maps – based on a small memory footprint representation of individual maps - in a way that is computationally efficient. We also demonstrate how the same framework can be used to detect changes in the map. The methods are tested on both sensor data from radio sensors and from cameras.

Contact: gabrielle.flood@math.lth.se

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

WHAT WOULD YOU EXPECT? ANTICIPATING EGOCENTRIC ACTIONS WITH ROLLING-UNROLLING LSTMS AND MODALITY ATTENTION

Furnari A., Farinella, G. M.

Abstract: Egocentric action anticipation consists in understanding which objects the camera wearer will interact with in the near future and which actions they will perform. We tackle the problem proposing an architecture able to anticipate actions at multiple temporal scales using two LSTMs to 1) summarize the past, and 2) formulate predictions about the future. The input video is processed considering three complimentary modalities: appearance (RGB), motion (optical flow) and objects (object-based features). Modality-specific predictions are fused using a novel Modality ATTention (MATT) mechanism which learns to weigh modalities in an adaptive fashion. Extensive evaluations on two large-scale benchmark datasets show that our method outperforms prior art by up to +7% on the challenging EPIC-KITCHENS dataset including more than 2500 actions, and generalizes to EGTEA Gaze+. Our approach is also shown to generalize to the tasks of early action recognition and action recognition. Our method is ranked first in the leaderboard of the EPIC-Kitchens egocentric action anticipation challenge. More details available here: http://iplab.dmi.unict.it/rulstm/.

Contact: furnari@dmi.unict.it

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

BEAUTY AND THE BEAST: OPTIMAL METHODS MEET LEARNING FOR DRONE RACING

Kaufmann E., Gehrig M., Foehn P., Ranftl R., Dosovitskiy A., Koltun V., Scaramuzza D.

Abstract: Autonomous micro aerial vehicles still struggle with fast and agile maneuvers, dynamic environments, imperfect sensing, and state estimation drift. Autonomous drone racing brings these challenges to the fore. Human pilots can fly a previously unseen track after a handful of practice runs. In contrast, state-of-the-art autonomous navigation algorithms require either a precise metric map of the environment.

To bridge this gap, we propose an approach that can fly a new track in a previously unseen environment without a precise map. This is achieved by predicting gate poses together with their uncertainties such that accurate gate locations can be estimated over time. These estimates are used to plan trajectories tracked by a model predictive control scheme for high performance flight. The presented approach was used to win the IROS 2018 Autonomous Drone Race Competition, outracing the second-placing team by a factor of two.

Contact: mgehrig@ifi.uzh.ch

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

END-TO-END LEARNING OF REPRESENTATIONS FOR ASYNCHRONOUS EVENT BASED DATA

Gehrig D., Loquercio A., Derpanis K. G., Scaramuzza D.

Abstract: Event cameras have appealing advantages over frame-based cameras for computer vision, including high temporal resolution, high dynamic range, and no motion blur. However, due to the sparse, asynchronous nature of the event signal, pattern recognition algorithms typically aggregate events into a grid-based representation and process it by a standard vision pipeline, e.g., CNNs. We introduce a general framework to convert event streams into grid-based representations through a sequence of differentiable operations. Our framework (i) allows the joint learning of representation and task dedicated network in an end to end manner, and (ii) unifies the majority of extant representations and identifies novel ones. We show that our approach to learning the event representation end-to-end yields an improvement of approximately 12% on optical flow estimation and object recognition over state-of-the-art.

Contact: dgehrig@ifi.uzh.ch

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

SIAMESE BALLISTICS NEURAL NETWORK

Giudice O., Guarnera L., Paratore A. B., Farinella G. M. , Battiato S.

Abstract: Firearm identification is crucial in many investigative scenario. The crime scene often contains traces left by firearms in terms of bullets and cartridges. Traces analysis is a fundamental step in the Forensics Ballistics Analysis Process to identify which firearm fired a specific cartridge. In this paper we present a fully automated technique to compare cartridges represented as a set of 3D point-clouds. The overall approach is based on Siamese Neural Network learning paradigm that we use to build a suitable embedding space where the 3D point-cloud of the cartridges are compared. The proposed approach has been assessed by considering the NBTRD dataset. Obtained results support the exploitation of the proposed technique in ballistic analysis.

Contact: guarneraluca92@gmail.com

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

A NEW STUDY ON WOOD FIBER TEXTURE: DOCUMENTS AUTHENTICATION THROUGH LBP FINGERPRINT

Guarnera F.

Abstract: The authentication of printed material based on textures is a critical and challenging problem for many security agencies in many contexts: valuable documents, banknotes, tickets or rare collectible cards are often targets for forgery. This motivates the study of low cost, fast and reliable approaches for documents authenticity analysis. In this paper, we present a new approach based on the extraction of translucent patterns from paper sheet by means of a specific built framework. A fingerprint is obtained by computing a Local Binary Pattern descriptor on the digital image. To validate the robustness of the proposed method for authentication analysis, we introduce a nove dataset and perform retrieval tests under both, ideal and noisy conditions. Experimental results prove the validity of the proposed strategy

Contact: guarnera.francesco@gmail.com

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

DETECTFUSION: DYNAMIC OBJECT-LEVEL SLAM IN REAL-TIME

Hachiuma R., Pirchheim C., Schmalstieg D., Saito H.

Abstract: We present DetectFusion, an RGB-D SLAM system that runs in real time and can robustly handle semantically known and unknown objects that move independently. Our method detects, segments and assigns semantic class labels to known objects in the scene. For improving the accuracy of SLAM, we also segment semantically unknown objects. Our method outperforms previous work in terms of localization and reconstruction accuracy while achieving 20 FPS even if the objects are segmented every frame.

Contact: ryo-hachiuma@hvrl.ics.keio.ac.jp

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

SEMANTIC UNDERSTANDING OF FOGGY SCENES WITH PURELY SYNTHETIC DATA

Hahner M., Dai D., Sakaridis C., Zaech J., Van Gool L.

Abstract: This work addresses the problem of semantic scene understanding under foggy road conditions. We propose a novel method, which uses purely synthetic data to improve the performance on unseen real- world foggy scenes captured in the streets of Zurich and its surroundings. Our results highlight the potential and power of photorealistic synthetic images for training and especially fine-tuning deep neural nets.

Contact: martin.hahner@vision.ee.ethz.ch

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

VIDEO REPRESENTATION LEARNING BY DENSE PREDICTIVE CODING

Han T., Xie W., Zisserman A.

Abstract: The objective of the work is self-supervised learning of spatiotemporal embedding from video, suitable for human action recognition. We make three contributions: First, we introduce the Dense Predictive Coding (DPC) framework for self-supervised video representation learning. Second, we propose a curriculum training scheme that encourages the model to learn semantics. Third, we evaluate the approach by finetuning on action recognition benchmarks and achieve SOTA performance on UCF101 and HMDB51.

Contact: htd@robots.ox.ac.uk

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

HUMAN VISUAL ATTENTION IN IMAGE CAPTIONING

Sen He,Nicolas Pugeault

Abstract: This work aims to study how human use their visual attention mechanism to do image captioning task. We collected a dataset which synchronously recorded human eye-fixations and scene descriptions (in verbal form), and use the collected data to: (1) compare human attention during scene free-viewing with human attention during describing images; (2) analyse the relationship between eye-fixations and descriptions under the image captioning task; (3) compare human attention and machine attention in image captioning; (4) build a general image captioning architecture, using image saliency as prior knowledge, combined with the soft attention mechanism.

Contact: sh752@exeter.ac.uk

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

DELF ATTENTION POOLED FEATURES

Wahab F., Sulthanpete Govindappa K., Heße L.

Abstract: We study the problem of learning relevant feature descriptors for object retrieval. We mainly focus on the idea of attentive deep local features (DELF) traditionally used for extracting local features to address landmark retrieval [1, 2] or remote sensing [3]. We interpret the DELF pipeline as a top-down attention mechanism, which facilitates learning descriptors, while suppressing non-discriminative regions in the image.

Contact: lisa@nyris.io

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

MULTIVIEW ABSOLUTE POSE USING 3D - 2D PERSPECTIVE LINE CORRESPONDENCES AND VERTICAL DIRECTION

Horanyi N.

Abstract: We address the problem of estimating the absolute pose of a multiview calibrated perspective camera system from 3D - 2D line correspondences with the known vertical direction. We propose two solutions, both can be used as a minimal solver as well as a least squares solver. The first solution consists of a single linear system of equations, while the second yields a polynomial equation of degree three in one variable and one system of linear equations which can be efficiently solved in closed-form.

Contact: n.horanyi@pgr.bham.ac.uk

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

ATLAS-BASED SKULL SEGMENTATION IN MRI USING STATISTICAL SHAPE MODELS

Kahr P.

Abstract: We present an atlas-based approach to skull segmentation in MRI. In atlas registration, pre-segmented images are matched to a target image. A segmentation is obtained by combining the atlas labels. The novelty in our method is that the deformations used in the image registration step are learned from data. Thus, our method adapts to the shape of the atlas. We show that our method leads to more accurate results than a generic model of smooth deformations.

Contact: patrick.kahr@unibas.ch

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

DOMAIN ADAPTATION FOR PRIVACY-PRESERVING PEDESTRIAN DETECTION IN THERMAL IMAGERY

My Kieu, Andrew D. Bagdanov, Marco Bertini, Alberto D. Bimbo

Abstract: This research investigates two domain adaptation techniques for fine-tuning a YOLOv3 detector to perform accurate and robust pedestrian detection using thermal images. Our approaches are motivated by the fact that thermal imagery is privacy-preserving in that person identification is difficult or impossible. Results on the KAIST dataset show that our approaches perform comparably to the state-of-the-art and outperform the state-of-the-art on night-time pedestrian detection, even outperforming multimodal techniques using both thermal and visible spectrum imagery at test time.

Contact: my.kieu@unifi.it

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

PROGRESSIVE ATTENTION MEMORY NETWORK FOR MOVIE STORY QUESTION ANSWERING

Kim J., Ma M., Kim K., Kim S., Yoo C. D.

Abstract: We propose progressive attention memory network (PAMN) for movie story question answering. PAMN involves three main features: (1) progressive attention mechanism that progressively prune out irrelevant temporal parts in memory, (2) dynamic modality fusion that adaptively determines the contribution of each modality, and (3) belief correction answering scheme that successively corrects the prediction score on answer. PAMN shows SOTA results on MovieQA benchmark.

Contact: junyeong.kim@kaist.ac.kr

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

SUBMODULAR BATCH SELECTION FOR TRAINING DEEP NEURAL NETWORKS

Joseph K.J., Vamshi T.R., Krishnakanth S., Balasubramanian V.N.

Abstract: Abstract • Introduce a batch selection strategy for mini-batch gradient descent, based on submodular function maximization. • Formulate a novel submodular objective that captures the informativeness of each sample and diversity of the whole subset. • Design a scalable algorithm that gives quality solution to this NP-hard selection problem.

Contact: cs17m18p100001@iith.ac.in

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

EMBODIED VISION-AND-LANGUAGE NAVIGATION WITH DYNAMIC CONVOLUTIONAL FILTERS

Landi F., Baraldi L., Corsini M., Cucchiara R.

Abstract: In Vision-and-Language Navigation, an embodied agent needs to reach a target destination with the only guidance of a natural language instruction. We propose to exploit dynamic convolutional filters to ground the lingual description into the visual observation in an elegant and efficient way. Results show that our model that incorporates dynamic filters performs better than architectures with traditional convolution, being the new state of the art for embodied VLN in the low-level action space.

Contact: federico.landi@unimore.it

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

3D APPEARANCE SUPER-RESOLUTION WITH DEEP LEARNING

Li Yawei , Tsiminaki Vagia, Timofte Radu, Pollefeys Marc, van Gool Luc

Abstract: We tackle the problem of retrieving high-resolution (HR) texture maps of objects that are captured from multiple view points. In the multi-view case, model-based super-resolution (SR) methods have been recently proved to recover high quality texture maps. On the other hand, the advent of deep learning-based methods has already a significant impact on the problem of video and image SR. Yet, a deep learning-based approach to super-resolve the appearance of 3D objects is still missing. The main limitation of exploiting the power of deep learning techniques in the multi-view case is the lack of data. We introduce a 3D appearance SR (3DASR) dataset based on the existing ETH3D [42], SyB3R [31], MiddleBury, and our Collection of 3D scenes from TUM [21], Fountain [51] and Relief [53]. We provide the high- and low-resolution texture maps, the 3D geometric model, images and projection matrices. We exploit the power of 2D learning-based SR methods and design networks suitable for the 3D multi-view case. We incorporate the geometric information by introducing normal maps and further improve the learning process. Experimental results demonstrate that our proposed networks successfully incorporate the 3D geometric information and super-resolve the texture maps.

Contact: yawei.li@vision.ee.ethz.ch

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

UNSUPERVISED MOVING OBJECT DETECTION VIA CONTEXTUAL INFORMATION SEPARATION

Yang Y., Loquercio A., Scaramuzza D., Soatto S.

Abstract: We propose an adversarial contextual model for detecting moving objects in images. A deep neural network is trained to predict the optical flow in a region using information from everywhere else but that region (context),while another network attempts to make such context as uninformative as possible. The result is a model where hypotheses naturally compete with no need for explicit regularization or hyper-parameter tuning. Although our method requires no supervision whatsoever, it outperforms several methods that are pre-trained on large annotated datasets. Our model can be thought of as a generalization of classical variational generative region-based segmentation, but in a way that avoids explicit regularization or solution of partial differential equations at run-time.

Contact: loquercio@ifi.uzh.ch

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

INTERPRETABILITY BEYOND CLASSIFICATION OUTPUT: SEMANTIC BOTTLENECK NETWORKS

Losch M., Fritz M., Schiele B.

Abstract: Deep learning systems deliver high performance based on end-to-end training, yet they are hard to interpret. To address this issue, we propose Semantic Bottleneck Networks (SBN): deep networks with semantically interpretable intermediate layers that all downstream results are based on. We present a case study in which we show that SBNs can recover state of the art performance on street scene segmentation while the SB-activations can be used for the interpretation of failure cases.

Contact: mlosch@mpi-inf.mpg.de

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

VIDEO OBJECT SEGMENTATION: PREMVOS VS BOLTVOS

Luiten J., Voigtlaender P., Leibe B.

Abstract: We present two methods for Video Object Segmentation (VOS), PReMVOS (Proposal-generation, Refinement & Merging for VOS), and BoLTVOS (Box-Level Tracking for VOS). In PReMVOS object proposals are generated, and linked by optical flow and visual queues. BoLTVOS is a Siamese network that detects objects by comparing to the first frame. PReMVOS won the 2018 DAVIS & YOUTUBE-VOS challenges. BoLTVOS is 45x faster and uses only the given bounding-box. Combining these got 2nd in the 2019 DAVIS challenge.

Contact: luiten@vision.rwth-aachen.de

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

PATTERNS OF 3D FACIAL SEXUAL DIMORPHISM: A TRIPLET LOSS APPROACH

Mahdi S., Nauwelaers N., Bouritsas G., Bokhnyak S., Bronstein M., Zafeiriou S., Boets B., Claes P.

Abstract: Recent advances in deep learning and computer vision have made a significant impact on a variety of domains. However, there is a fundamental and often overlooked shortcoming of these techniques when applied to biological and medical image data. Hence, it is of interest to examine the biological plausibility and interpretability of the knowledge learned by deep networks. Particularly within this study, we aim to use 3D images to investigate the existence of multiple patterns of sexual dimorphism in human faces. The outcomes of this study can potentially result in a powerful alternative to the current linear biological shape analysis which provides only a single metric for sexual dimorphism. We first train an autoencoder for 3D facial images with a 32 dimensional latent space. Once obtained, the trained model will be fine-tuned for sex classification with the deep metric learner function called Triplet Loss (TL). The advantages of this combination are: (1) Triplet loss is trained with triplets of faces, thus, the number of training data can practically increase to n3. (2) Due to the metric measures included in the TL function, the latent space will reflect the similarity in sex groups, hence distances are meaningful measures for classification purposes. (3) The potentially multiple patterns of sexual dimorphism learned by the network can be visualized with the help of the decoder, which improves the interpretability of the learned model.

Contact: soha.mahdi94@gmail.com

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

GRADIENT BOOSTING ON DECISION TREES FOR MORTALITY PREDICTION IN TRANSCATHETER AORTIC VALVE IMPLANTATION

Mamprin M., Zelis J.M., Tonino P.A.L., Zinger S., de With P.H.N.

Abstract: Current prognostic risk scores in cardiac surgery, which are based on statistics, are not robust enough to identify patients who would benefit from TAVI procedure. This research aims to create a ML model to predict one-year mortality, by adopting a gradient boosting on decision trees algorithm in combination with a technique for model interpretations. A feature analysis and selection stage has been developed, enabling to identify the most important features for the prediction. We validated our model on 275 TAVI cases, reaching an AUC of 0.90 and outperforming current predictors.

Contact: m.mamprin@tue.nl

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

DEEP MODEL-BASED 6D POSE REFINEMENT IN RGB

Manhardt F., Kehl W., Navab N, Tombari F.

Abstract: We present a novel approach for model-based 6D pose refinement in color data. Building on the established idea of contour-based pose tracking, we teach a deep neural network to predict a translational and rotational update. At the core, we propose a new visual loss that drives the pose update by aligning object contours, thus avoiding the definition of any explicit appearance model. In contrast to previous work our method is correspondence-free, segmentation-free, can handle occlusion and is agnostic to geometrical symmetry as well as visual ambiguities.

Contact: fabian.manhardt@tum.de

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

NEURAL NETWORKS FOR VIDEO INTRA PREDICTION

Meyer M., Schneider J., Rohlfing C.

Abstract: In this work, the properties of different architectures for neural network-based intra prediction are evaluated. This includes an analysis of architecture type properties as well as proposing a way to integrate cross-component information for chroma prediction. Furthermore, a new way of signaling an additional intra prediction mode is investigated and different loss functions are tested. In total this improves the compression performance in terms of average BD-rate changes by -1.8%.

Contact: meyer@ient.rwth-aachen.de

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

IMPLICIT REPRESENTATION FOR 3D SHAPE INFERENCE

Michalkiewicz M. , Pontes J. , Baktashmotlagh M. , Eriksson A.

Abstract: How to effectively represent 3D shape in a learning framework still remains an open research problem. Recent works have been relying on volumetric, point cloud and mesh representations, but these approaches typically suffer from issues of computational complexity, unordered data, and lack of finer geometry. Here, we instead propose the use of implicit functions to represent shape for 3D shape inference. We show that this formulation will address many of the shortcomings of existing methods in an elegant way, while still compactly and accurately encoding highly detailed geometry.

Contact: mat.michalkiewicz@gmail.com

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

UNSUPERVISED DOMAIN ADAPTATION FOR URBAN SCENES SEGMENTATION

Biasetton M., Michieli U., Agresti G., Zanuttigh P.

Abstract: The semantic understanding of urban scenes is one of the key components for autonomous driving systems. Deep neural networks require huge amount of labeled data, which is difficult and expensive to acquire. A recent workaround is to exploit synthetic data but differences between real and synthetic scenes limit the performance. We propose an unsupervised domain adaptation strategy from a synthetic supervised training to real data. Experimental results demonstrate that the proposed approach is able to adapt a network trained on synthetic datasets to a real one.

Contact: umberto.michieli@dei.unipd.it

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

MARKERLESS HUMAN MOTION ANALYSIS ON RGB VIDEOS

Moro M., Odone F., Casadio M.

Abstract: In this work we investigate the possibility to use computer vision and deep learning techniques in order to study in a precise, reliable and unobtrusive way the human motion, in the context of neuroscience or motor control.

Contact: matteo.moro1994@libero.it

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

DEEP METRIC LEARNING FOR DENSE CORRESPONDENCES

Nickel M., Dhamo H., Tombari F., Navab N.

Abstract: We present an approach for learning dense correspondence fields in human bodies from depth images. Contrary to a previously published method that formulates learning as a classification problem, we draw inspiration from deep metric learning. Our approach inherently enforces smooth feature manifolds on the body, yielding robust correspondence matching even for unfamiliar poses. It allows to train on datasets without inter ground truth correspondence.

Contact: manuel.nickel@tum.de

Presentation Type: Poster

Date: Monday 8 July 2019

Time: 21:30

Poster Session: 1

BEYOND ACTION RECOGNITION

Nicora E., Goyal G., Noceti N., Odone F.

Abstract: Taking inspiration from children's developmental process in motion perception, we will focus our research on the design of a cognitive-inspired framework for motion representation and understanding able to transfer knowledge among semantic tasks of increasing complexity.

Contact: elena.nicora@dibris.unige.it

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

FIRST PERSON VISION FOR VISITORS BEHAVIOURAL UNDERSTANDING AND FRUITION AUGMENTATION IN CULTURAL SITES

Orlando S. A., Furnari A., Gallo G., Farinella G. M.

Abstract: Computer vision and machine learning techniques have been recently uses in cultural heritage domain to provide service useful for the managements of cultural sites and also to augment the experience of the visitors during their visit. Performing localization using first person vision provided from AR/VR goggles can help both for this purpose. Actual methods of visual localization require a huge amount of data to train a CNN for this purpose. We propose a tool able to collect virtual data useful for visual localization and object recognition, generating navigations in the virtualized cultural site and capturing frame per frame the ground truth camera position. Using triplet network and image-retrieval we performed 3 Degree of Freedom (3DoF) localization. Thanks to the capability of the tool to collect different channels of the camera, we also collect for each frame relative semantic information. We tested object detectors of the state-of-the-art to asses the usefulness of the tool. We drastically reduce the cost of dataset labelling, and the results obtained in the virtual domain are convincing and pushing us to explore the adaptation in real domain for future works.

Contact: santi.orlando@unict.it

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

GAUSSIAN BELIEF PROPAGATION FOR ESTIMATION IN SLAM

Ortiz J., Davison A.

Abstract: Inference problems in SLAM can be formulated as factor graphs, but have typically been solved using nonlinear optimisation. We apply Loopy Gaussian Belief Propagation (GBP) to the monocular SfM/SLAM problem using sparse features and show this approach can perform inference on extended 3D loopy sequences. We also propose an information theoretic thresholding to robustify the convergence properties of GBP by removing edges in the graph.

Contact: j.ortiz18@imperial.ac.uk

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

SMPLR: LEARNING A COMPACT, EFFICIENT, AND ACCURATE MODEL OF 3D HUMAN BODY SHAPE

Osman Ahmed, Black J Michael

Abstract: The SMPL body model is widely used for human pose estimation, animation, and commercial applications. While popular, SMPL suffers from several significant problems. First, SMPL has non-local corrective pose blend shapes that result in artifactual long-range shape deformations. This is related to the second problem, that SMPL has a huge number of parameters, yet is trained from a relatively small number of 3D scans, resulting in overfitting and poor generalization. To address these problems, we introduce SMPLR (Skinned Multi-person Learned sparse Regressor), a compact and simpler model that is a drop in replacement for SMPL. We introduce non-linear activation functions for the pose blendshapes and learn the subset of mesh vertices influenced by each body joint movement. Our formulation uses a significantly more compact pose parametrization based on quaternions, which results in a model only 11% of the SMPL pose blend shape parameters. We train SMPL and SMPLR on the same training set and show that SMPLR is more accurate. We go further and train the shape space of SMPLR on an additional 2845 male and 6434 female meshes and show that this larger dataset is necessary.

Contact: ahmed.osman@tuebingen.mpg.de

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

IMPACT OF ADVERSARIAL EXAMPLES ON DEEP LEARNING BIOMEDICAL SEGMENTATION MODELS

Ozbulak U., Van Messem A., De Neve W.

Abstract: Given that a large portion of medical imaging problems are effectively segmentation problems, we analyze the impact of adversarial examples on deep learning-based image segmentation models. We expose the vulnerability of these models to adversarial examples by proposing a novel algorithm that makes it possible to craft targeted adversarial examples that come with high intersection-over-union rates and with perturbation that is mostly invisible to the bare eye.

Contact: utku.ozbulak@ugent.be

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

COLORECTAL POLYPS DETECTION: WHERE ARE WE?

Pappalardo G., Cherubini A., Farinella G.M.

Abstract: We consider the problem of detecting polyps in images extracted from colonoscopy videos. To this aim we exploit deep learning to accurately localize the lesions on each frame of the colonoscopy video. We tested our approach on popular publicy available datasets as well as on a subset of our clinical trial.

Contact: giovannapappalardo91@gmail.com

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

DEPTH-BASED HUMAN ANALYSIS

Pini S., Borghi G., Vezzani R., Cucchiara R.

Abstract: Recently, research on depth devices and depth maps has emerged in the computer vision community. Our research addresses the face and human analysis exploiting this type of data. We focus on the automotive context, which presents challenging contexts and lighting conditions. To this end, we are working on face recognition and human pose estimation algorithms based on deep learning and depth images.

Contact: s.pini@unimore.it

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

DEEP NEURAL NETWORK BASED COOPERATIVE VISUAL TRACKING THROUGH MULTIPLE MICRO AERIAL VEHICLES

Price E., Lawless G., Ludwig R., Martinović I., Bülthoff H. H., Black M. J., Ahmad A.

Abstract: Our goal is: On-board, online, continuous and accurate vision based detection and tracking of humans using micro aerial vehicles (MAVs). Constraints: * Single monocular RGB Camera on each MAV. * Only on-board computation. * No inter-MAV image communication. * No marker or sensor on the human. Challenges: * State-of-the-art DNNs are prohibitively slow for flying hardware. * Measurement noise model is often not available for DNN-based detectors.

Contact: eric.price@tuebingen.mpg.de

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

UNICT-VEDI: DATASET AND CHALLENGES FOR VISITORS BEHAVIORAL UNDERSTANDING FROM EGOCENTRIC VISION

Ragusa F., Furnari A., Battiato S., Signorello G., Farinella G.M.

Abstract: We propose a new dataset composed by egocentric videos acquired in two cultural sites. The UNICT-VEDI Dataset contains multiple environments, over 200 Points of Interest and 70 visits acquired by real visitors. UNICT-VEDI Dataset is annotated with temporal labels, bounding box annotations and we release the surveys compiled by visitors at the end of the visit. In addiction we propose 4 challenges useful to understand visitors behavioral on this challenging dataset and a baseline for each challenge.

Contact: francescoragusa@outlook.com

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

OUTLIER DETECTION AND ROBUST FITTING WITH STATISTICAL SHAPE MODELS

Rahbani D., Morel-Forster A., Madsen D., Lüthi M., Vetter T.

Abstract: Statistical shape models are parametric models of shape variation. The main assumption behind them is correspondence: a one-to-one mapping between two shapes. This is not always true for shapes in medical images: pathologies create regions that cannot be mapped to the model. We develop a method to infer correspondences in the presence of pathologies and use it to improve shape reconstruction.

Contact: dana.rahbani@unibas.ch

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

ROBOT NAVIGATION IN REAL INDOOR ENVIRONMENTS USING REINFORCEMENT LEARNING

Rosano M., Furnari A., Farinella G. M.

Abstract: The ability to navigate autonomously in an indoor environment is a fundamental tasks for a robot. Several approaches have been proposed to address this open and challenging issue. In this work, we investigate different approaches to robot navigation based on reinforcement learning and image-based localization. Experiments are carried out a simulated environment obtained collecting images of a real environment in a regular grid.

Contact: marco.rosano@unict.it

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

UNSUPERVISED DOMAIN ADAPTATION USING FEATURE-WHITENING AND CONSENSUS LOSS

Roy S., Siarohin A., Sangineto E., Bulò SR., Sebe N. and Ricci E.

Abstract: A classifier trained on a dataset seldom works on other datasets due to domain shift. This problem is addressed by unsupervised domain adaptation (UDA) methods. In this work we propose a deep learning framework which unifies different paradigms in UDA. Specifically, we propose domain alignment layers which implement feature whitening for aligning source and target features. Additionally, we leverage the unlabelled target data by proposing the Min-Entropy Consensus loss which regularizes training

Contact: subhankar.roy@unitn.it

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

CRITICAL TRAINING SET SIZE AND DROPOUT KEEP PROBABILITY FOR LENET4 CONVOLUTIONAL NEURAL NETWORK TRAINED ON MNIST. IS IT PERCOLATION?

Sala V.

Abstract: In this work, LeNet4 deep convolutional neural network was modified by applying dropout on the fully connected layer and trained on subsets of variable size of the MNIST dataset. Two different training regimes were found, well separated by a jump in training and test error. Then, the effect of dropout keep probability was investigated showing its influence on the critical training set size. It was found that the model was able to fit the data only above a threshold for keep probability.

Contact: vittorio.sala@imagesspa.it

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

AUDIO-VISUAL MODEL DISTILLATION USING ACOUSTIC IMAGES

Pérez A.F., Sanguineti V., Morerio P., Murino V.

Abstract: We use a new multimodal dataset acquired by a sensor that provides RGB video, raw audio and acoustic images synchronized in time. Audio models are trained with distillation from both visual and acoustic image teachers [1]. We learn robust audio features towards variable environmental conditions more powerful than those learnt just using visual or single-microphone audio data.

Contact: valentina.sanguineti@iit.it

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

PRIVACY PROTECTION IN STREET-VIEW PANORAMAS USING DEPTH AND MULTI-VIEW IMAGERY

Uittenbogaard R., Sebastian C., Vijverberg J., Boom B., Gavrila D.M., de With P.H.N

Abstract: The current paradigm in privacy protection in street-view images is to detect and blur sensitive information. In this paper, we propose a framework that is an alternative to blurring, which automatically removes and in paints moving objects (e.g. pedestrians, vehicles) in street-view imagery. We propose a novel moving object segmentation algorithm exploiting consistencies in depth across multiple street-view images that are later combined with the results of a segmentation network. The detected moving objects are removed and inpainted with information from other views, to obtain a realistic output image such that the moving object is not visible anymore. We evaluate our results on a dataset of 1000 images to obtain a peak noise-to-signal ratio (PSNR) and L1 loss of 27.2 dB and 2.5, respectively. To ensure the subjective quality, To assess overall quality, we also report the results of a survey conducted on 35 professionals, asked to visually inspect the images whether object removal and inpainting had taken place.

Contact: c.sebastian@tue.nl

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

DEEP BUILT-STRUCTURE COUNTING IN SATELLITE IMAGERY USING ATTENTION BASED RE-WEIGHTING

Shakeel Anza., Sultani Waqas., Ali Mohsen.

Abstract: Here we tackle the challenging problem of counting built-structures in satellite imagery by proposing several deep learning based regression techniques. Attention-based re-weighting methods are used to combine features from different regions of an image; these features are fused together in FusionNet that weigh different features differently. To train and evaluate the methods, a new built-structure-count dataset is created. FusionNet has Mean Absolute Error of 3.61 over the testing data

Contact: anza.shakeel@durham.ac.uk

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

LOCAL FEATURES AND VISUAL WORDS EMERGE IN ACTIVATION

Oriane S., Yannis A., Ondra C.

Abstract: DSM is a novel method of deep spatial matching for image retrieval. Initial ranking is based on image descriptors globally pooled from convolutional neural network activations. The same sparse activation tensor is also approximated by a collection of local features. These local features are then matched to approximate the optimal alignment of the tensors. This happens without network modification or training. No local feature detection, no feature descriptors, no visual vocabulary are needed.

Contact: oriane.simeoni@inria.fr

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

FACE DETECTION IN THE OPERATING ROOM: A COMPARISON OF STATE-OF-THE-ART METHODS AND A SELF-SUPERVISED APPROACH

Issenhuth Thibaut, Srivastav Vinkle, Gangi Afshin, Padoy Nicolas

Abstract: Face detection is a needed component for the automatic analysis and assistance of human activities during surgical procedures. Efficient face detection algorithms can indeed help to detect and identify the persons present in the room and also be used to automatically anonymize the data. However, current algorithms trained on natural images do not generalize well to the operating room (OR) images. In this work, we provide a comparison of state-of-the-art face detectors on OR data and also present an approach to train a face detector for the OR by exploiting nonannotated OR images. We propose a comparison of six state-of-the-art face detectors on clinical data using multi-view OR faces, a dataset of OR images capturing real surgical activities. We then propose to use self-supervision, a domain adaptation method, for the task of face detection in the OR. The approach makes use of non-annotated images to fine-tune a state-of-the-art face detector for the OR without using any human supervision. The results show that the best model, namely the tiny face detector, yields an average precision of 0.556 at intersection over union of 0.5. Our self-supervised model using non-annotated clinical data outperforms this result by 9.2%.

Contact: srivastav@unistra.fr

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

DEEP LEARNING FOR GASTRIC PATHOLOGY DETECTION IN ENDOSCOPIC IMAGES

Stefanidi A.

Abstract: Computer-aided diagnosis of cancer based on endoscopic image analysis is a promising area in the field of computer vision and machine learning. This research presents the algorithm of pathology detection in endoscopic images of gastric lesions based on convolutional neural network. Training and testing of the algorithm was carried out on the NVIDIA DGX-1 supercomputer using endoscopic images from the test base, assembled together with the Yaroslavl Regional Cancer Hospital.

Contact: antonstefanidi@mail.ru

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

ASSISTED TARGET DETECTION IN SEARCH AND RESCUE

Taheri-Shirazi, M., Ehinger, K. A., & Elder, J. H.

Abstract: In Canada, search and rescue (SAR) for a downed aircraft in distress is challenging. The military still relies on the naked eye to conduct airborne search and rescue missions. Partial automation of the target detection would alleviate some of the operator's workload. A key limitation has been the lack of a large, realistic and ground-truthed search and rescue dataset. This project will take advantage of the new airborne SAR dataset created by the NRC Flight Research Laboratory.

Contact: maryam.taheri.shirazi@gmail.com

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

VRPROP-NET: TRANSPORTING REAL OBJECTS INTO VIRTUAL AND AUGMENTED ENVIRONMENTS

Taylor C., McNicholas R., Cosker D.

Abstract: To improve the immersion in virtual and augmented reality experiences, we propose an end to end pipeline for creating interactive virtual props from real-world physical objects. These props can be used as controllers or as a proxy for the real object. As part of our pipeline, we propose a new neural network - VRProp-Net - to accurately predict rigid and non-rigid deformation parameters from unlabelled RGB images.

Contact: cct43@bath.ac.uk

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

ART2REAL: UNFOLDING THE REALITY OF ARTWORKS VIA SEMANTICALLY AWARE IMAGE-TO-IMAGE TRANSLATION

Tomei M., Cornia M., Baraldi L., Cucchiara R.

Abstract: The applicability of computer vision to paintings has been rarely investigated due to the small amount of annotated artistic data. We propose a semantic aware architecture which can translate artworks to photo realistic visualizations by retrieving details from real photos through a similarity matching strategy. Experimental results show that the proposed technique leads to improved performance of pre-trained classification, detection and segmentation architectures applied to artworks.

Contact: matteo.tomei@unimore.it

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

AUTOMATED TISSUE SEGMENTATION OF MICRO-CT IMAGES BY MACHINE LEARNING AND ITS APPLICATION TO COMPARATIVE MORPHOLOGY

Toulkeridou E., Baum D., Doya K., Economo E.

Abstract: Three-dimensional (3D) scanning, e.g. micro-computed tomography (micro-CT), is increasingly used by organismal biologists, producing data-rich volumetric images. To date, the segmentation of anatomical structures remains a bottleneck in research, as it is commonly performed manually. As Machine Learning (ML) methods and especially Convolutional Neural Networks (CNNs) proved to be very efficient in semantic segmentation, they are promising candidates to overcome this hindrance.Here, a new tool is developed using CNNs that can automate the segmentation of micro-CT images of non-model organisms, using ants (Formicidae) as a model study taxon. As a result, the construction of segmented maps, and the morphological quantification and comparison of different insect species will become easier and more efficient, leading to better insight in evolution and genetics.

Contact: evropi.toulkeridou@oist.jp

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

COMPUTER VISION-BASED SYSTEM FOR ADAS

Trenta Francesca

Abstract: The increasing demand of algorithms to manage spatial data have led to the analysis of them in different domains where a great amount of data is available. The aim of this research is focused on automotive field whereas sequential data are represented by time-series. Several studies have revealed that further exploitation applying frameworks based on Computer Vision represent valid support in reaching effective results. The purpose is to develop an efficient system for the detection of driver drowsiness in order to reduce the number of fatigue-related car accidents.

Contact: francesca.trenta4479@gmail.com

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

SYNTHESIS OF REALISTIC HIGH-RESOLUTION LIDAR DATA

Triess L., Peter D.

Abstract: This poster gives an overview on my master thesis project that has been accepted as a paper at the 30th IEEE Intelligent Vehicles Symposium. The paper presents a novel CNN-based approach for synthesizing high-resolution LiDAR point clouds. Our approach generates semantically and perceptually realistic results with guidance from specialized loss-functions.

Contact: larissa.triess@daimler.com

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

QT BASED ROBOTIC VISION SOFTWARE: SMART VISION

Temürlenk Y., Taşdelen A., Kaya A., Tükel D.B

Abstract: Vision systems are widely used in robotic based manufacturing industry, mainly for object's position, orientation detection and inspection. The purpose of this project is to create a vision system that will mark the parts which are scattered on a plane or a conveyor, that match the reference part and provide the values containing the X, Y position and Θ angle according to the reference coordinate set as origin point. The Smart Vision* software developed as a flexible and user-friendly system.

Contact: dtukel@altinay.com

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

ISOLATED SIGN RECOGNITION WITH A SIAMESE NEURAL NETWORK OF RGB AND DEPTH STREAMS

Tur A. O., Yalim Keles H.

Abstract: Sign recognition is a challenging problem due to high variance of the signs among different signers and multiple modalities of the input information. In addition, the challenges that exist in the action classification problems in computer vision are similar in this domain too, such as variations in illumination and background. In this work, we propose a Siamese Neural Network (SNN) architecture that is used to extract features from the RGB and the depth streams of a sign frame in parallel. We use a pretrained model for the SNN without any finetuning to our training data. We then apply global feature pooling to the depth and color features that the SNN generates and feed the concatenation of the selected features to a recurrent neural network (RNN) to discriminate the signs. We trained our model parameters with the Montalbano dataset and achieve 93.19% test accuracy with ResNet-50 and 91.61% with VGG-16 Network Models.

Contact: aotur@ankara.edu.tr

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

DECIDING WHAT LAYERS TO SHARE IN MULTI-TASK NETS

Vandenhende S., Georgoulis S., De Brabandere B., Van Gool L.

Abstract: The construction of multi-task networks requires to decide on the layers that can be shared between tasks. As the number of possible configurations is combinatorially large, prior work often relied on ad hoc methods to determine the layer sharing structure. We propose a principled method for the automated construction of branched multi-task networks, based on the usefulness of a set of features of one task for the other, and vice versa. We tested our method on the CelebA, Cityscapes and SUN RGB-D datasets.

Contact: simon.vandenhende@kuleuven.be

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

SCENE UNDERSTANDING FOR SELF-DRIVING CARS

Van Gansbeke W., De Brabandere B., Neven D., Van Gool L.

Abstract: This paper focuses on two aspects of scene understanding for self-driving cars. On the one hand we propose an end-to-end lane detection method to directly regress towards the final lane line coefficients. We model the lane lines by a second degree polynomial in a bird's eye view and obtain better results than the conventional two-step pipeline: segmentation followed by a line fitting module. On the other hand, we propose a novel depth completion method. The goal is to generate dense depth maps from sparse and noisy depth measurements (e.g. LiDAR). The architecture is specifically designed to handle noisy data. LiDAR and RGB frames are fused by exploiting global, local information and attention maps. This technique ranks first on the KITTI depth completion benchmark taking into account no extra data nor post-processing.

Contact: wouter.vangansbeke@esat.kuleuven.be

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

MOTS: MULTI-OBJECT TRACKING AND SEGMENTATION

Voigtlaender P., Krause M., Osep A., Luiten J., Gnana B., Geiger A., Leibe B.

Abstract: We extend the popular task of multi-object tracking to multi-object tracking and segmentation (MOTS). We provide pixel-level annotations for KITTI and MOTChallenge comprising 65,213 pixel masks for 977 objects (cars and pedestrians) in 10,870 video frames. We propose a new baseline method to address detection, tracking, and segmentation with a single convolutional network. We make our annotations, code, and models available at https://www.vision.rwth-aachen.de/page/mots.

Contact: voigtlaender@vision.rwth-aachen.de

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

MACHINE LEARNING FOR 3D RECONSTRUCTION

Weilharter R., Fraundorfer F.

Abstract: Accurate, dense 3D reconstruction from images has a large number of applications, ranging from augmented reality over 3D modeling to structure inspection. The latter can be of particular importance for structural vulnerability assessment, especially related to infrastructure (e.g. bridges). For this purpose and within the context of the RESIST project we are assembling an UAV mounted photogrammetric stereo vision system, capable to accurately compute 3D measurements.

Contact: rafael.weilharter@icg.tugraz.at

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

6DOF VEHICLE POSE ESTIMATION FOR MONOCULAR RGB IMAGES

WU DI

Abstract: A framework called 6D-VNet is proposed for 6DoF object pose estimation, especially for autonomous driving scenario. 6D- VNet efficiently detects traffic participants in a monocular RGB image while simultaneously regressing their 3D translation and rotation vectors. The method extends Mask R-CNN by adding customised heads for predicting vehicle's finer class, rotation and translation. The proposed 6D-VNet reaches the 1st place in ApolloScape challenge 3D Car Instance.

Contact: stevenwudi@gmail.com

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

PHOTO-GEOMETRIC AUTOENCODING TO LEARN 3D OBJECTS FROM UNLABELLED IMAGES

Wu S., Rupprecht C., Vedaldi A.

Abstract: We show that generative models can be used to capture visual geometry constraints statistically. We use this fact to infer the 3D shape of object categories from raw single-view images, without any external supervision, such as keypoints, multiple views or videos. We achieve this by a simple reconstruction task, exploiting the symmetry of the objects' shape and albedo. Our experiments show that this method can recover the 3D shape of human faces, cat faces, and cars from single-view images.

Contact: szwu@robots.ox.ac.uk

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

SCSLAM - ON-THE-FLY SCENE COMPLETION WITH RGBD SLAM

Wu S., Tateno K., Tombari F., Navab N.

Abstract: Path planning is one of the keys to achieving autonomous navigation. When ground truth map is not available, autonomous exploration relies on real-time map reconstruction and short distance path planning. In this stage, replanning often occurs due to the previously unseen obstacles and roads which causes the drop in efficiency. We proposed to use scene completion network to predict the unseen regions as prior knowledge to reduce the amount and magnitude required for replanning.

Contact: shuncheng.wu@tum.de

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

DEEP LEARNING FOR MATERIAL CLASSIFICATION BASED ON LOCAL AND GLOBAL FEATURES

Xu S., Tremeau A., Muselet D.

Abstract: Material recognition in real-world images is a challenging task. One of the difficulties is that unlike object recognition where one category of objects has some common features(e.g. shape), both the shape of materials in one category and their local features are quite different. In this poster,using an available material dataset,FMD(Flickr Material Database), we do several tests to find the contribution of 1.Object's global information 2.material's local information for material classification.

Contact: sixiang.xu@univ-st-etienne.fr

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

AUTOMATIC PAIN INTENSITY ESTIMATION FROM FACIAL VIDEOS

Yang R.J., Hong X.P., Peng J.Y., Zhao G.Y.

Abstract: Contriving machine learning models for estimating abnormal behaviors from faces has attracted lots of attentions in the field of automatic face analysis. Such models has paramount potential values for medical diagnosis and treatment. In this work, we focus on automatic pain estimation from faces, where pain is expressed by a set of deformations of facial muscles. We present a novel model that unites the low-level local descriptors and high-level deep features as facial representations for continuous pain intensity estimation. To obtain a powerful and compact low-level representation, we explore the way of using second-order pooling over the local descriptors. Our method is evaluated on the benchmark pain databases and achieves promising results.

Contact: 15619389313@163.com

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

STRUCTUREVO: VISUAL ODOMETRY IN STRUCTURED ENVIRONMENTS

Yanyan Li, Nikolas Brasch, Federico Tombari, Nassir Navab

Abstract: We propose a robust RGB-D visual odometry system for structured environments. By separately estimating rotational and translational motion, we take full advantage of the geometric information available in the scene, including the orthogonality of planes, parallel lines and feature points which are common in man-made environments, leading to reduced computation compared to traditional dense RGB-D SLAM systems. The proposed method is evaluated on the ICL-NUIM and TUM RGB-D benchmarks and obtains better results than other state-of-the-art methods.

Contact: yanyan.li.camp@gmail.com

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

GAN-BASED TEXTURE INTERPOLATION AND GENERATED IMAGE ATTRIBUTION

Yu N., Barnes C., Shechtman E., Amirghodsi S., Lukáč M., Davis L., Fritz M.

Abstract: Rearch project I: We present to address the problem of controllable texture interpolation. We show our method outperforms a number of baselines. We further show several applications based on our technique, which include texture brush, texture dissolve, and animal hybridization.

Research project II: We present the first study of learning GAN fingerprints towards image attribution: we systematically investigating the accuracy of classifying an image as real or GAN-generated. For GANgenerated images, we further identify their sources.

Contact: ningyu@cs.umd.edu

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2

FONT ADAPTIVE TEXT RECOGNITION NETWORK

ZHANG C, GUPTA A, ZISSERMAN A

Abstract: Our objective in this work is to be able to recognise lines of text written in previously unseen fonts through one-shot learning. We propose a new text recognition neural network architecture where the task of reading a new font simply involves enrolling the font by supplying glyphs of the alphabet and learn to match them to glyphs in text-lines. Our approach enables quick test-time adaptation of our visual encoders without disturbing the learnt language models, which is not feasible in conventional architectures.

Contact: czhang@robots.ox.ac.uk

Presentation Type: Poster

Date: Tuesday 9 July 2019

Time: 21:30

Poster Session: 2