cvpaper.challenge in 2015
A review of CVPR2015 and DeepSurvey

Hirokatsu Kataoka · Yudai Miyashita · Tomoaki Yamabe · Soma Shirakabe · Shin’ichi Sato · Hironori Hoshino · Ryo Kato · Kaori Abe · Takaaki Imanari · Naomichi Kobayashi · Shinichiro Morita · Akio Nakamura

Received: date / Accepted: date

Abstract The “cvpaper.challenge” is a group composed of members from AIST, Tokyo Denki Univ. (TDU), and Univ. of Tsukuba that aims to systematically summarize papers on computer vision, pattern recognition, and related fields. For this particular review, we focused on reading the ALL 602 conference papers presented at the CVPR2015, the premier annual computer vision event held in June 2015, in order to grasp the trends in the field. Further, we are proposing “DeepSurvey” as a mechanism embodying the entire process from the reading through all the papers, the generation of ideas, and to the writing of paper.

1 Introduction

cvpaper.challenge is a joint project aimed at reading papers mainly in the field of computer vision and pattern recognition. Currently the project is run by around ten members representing different organizations; namely, AIST, TDU, and University of Tsukuba. Reading international conference papers clearly provides various advantages other than gaining an understanding of the current standing of your own research, such as acquiring ideas and methods used by researchers around the world. In reality, however, although this input of knowledge is important, researchers and engineers are too busy to have time to do it, and the process takes a great amount of time and effort for undergraduate and graduate students (particularly masters course students) who lack research experience and entails sacrificing their time for classes and research. Assigning this work, however, to non-experts who are not familiar with the field of computer vision, results in a great amount of time needed for interpreting the papers. As a way to address this problem, we believe that we can make it relatively easier to grasp advanced technologies if we share and systematize knowledge using the Japanese language. We therefore undertook to extensively read papers, summarize them, and share them with others working in the same field. The IEEE-sponsored Conference on Computer Vision and Pattern Recognition (CVPR) is known as the premier conference in the field of computer vision, pattern recognition, and related fields. CVPR, which is held annually in the U.S., has on average around 20% acceptance rate for submitted papers, making it a very difficult conference to hurdle, and pointing to the high quality of the accepted papers. Also, CVPR is also known to comprehensively cover papers in the different fields in computer vision and pattern recognition. A number of prominent international researchers and research groups choose their research themes after a comprehensive grasp of almost all papers presented in premier conferences and an understanding of research trends. We believe that the accuracy by which research themes are chosen can be improved by constantly being updated on cutting-edge technologies and discussing these new technology trends within the research groups as part of their regular activities. Further, a survey of papers presented in premier conferences is also an essential way to gather tools needed for research. We therefore believe that gaining an understanding of papers presented in premier conferences is the best method for authors to comprehend the lat-
est trends in computer vision, pattern recognition, and related fields. As the first step of this endeavor, we undertook to read all the 602 papers accepted during the CVPR2015 [1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99,100,101,102,103,104,105,106,107,108,109,110,111,112,113,114,115,116,117,118,119,120,121,122,123,124,125,126,127,128,129,130,131,132,133,134,135,136,137,138,139,140,141,142,143,144,145,146,147,148,149,150,151,152,153,154,155,156,157,158,159,160,161,162,163,164,165,166,167,168,169,170,171,172,173,174,175,176,177,178,179,180,181,182,183,184,185,186,187,188,189,190,191,192,193,194,195,196,197,198,199,200,201,202,203,204,205,206,207,208,209,210,211,212,213,214,215,216,217,218,219,220,221,222,223,224,225,226,227,228,229,230,231,232,233,234,235,236,237,238,239,240,241,242,243,244,245,246,247,248,249,250,251,252,253,254,255,256,257,258,259,260,261,262,263,264,265,266,267,268,269,270,271,272,273,274,275,276,277,278,279,280,281,282,283,284,285,286,287,288,289,290,291,292,293,294,295,296,297,298,299,300,301,302,303,304,305,306,307,308,309,310,311,312,313,314,315,316,317,318,319,320,321,322,323,324,325,326,327,328,329,330,331,332,333,334,335,336,337,338,339,340,341,342,343,344,345,346,347,348,349,350,351,352,353,354,355,356,357,358,359,360,361,362,363,364,365,366,367,368,369,370,371,372,373,374,375,376,377,378,379,380,381,382,383,384,385,386,387,388,389,390,391,392,393,394,395,396,397,398,399,400,401,402,403,404,405,406,407,408,409,410,411,412,413,414,415,416,417,418,419,420,421,422,423,424,425,426,427,428,429,430,431,432,433,434,435,436,437,438,439,440,441,442,443,444,445,446,447,448,449,450,451,452,453,454,455,456,457,458,459,460,461,462,463,464,465,466,467,468,469,470,471,472,473,474,475,476,477,478,479,480,481,482,483,484,485,486,487,488,489,490,491,492,493,494,495,496,497,498,499,500,501,502,503,504,505,506,507,508,509,510,511,512,513,514,515,516,517,518,519,520,521,522,523,524,525,526,527,528,529,530,531,532,533,534,535,536,537,538,539,540,541,542,543,544,545,546,547,548,549,550,551,552,553,554,555,556,557,558,559,560,561,562,563,564,565,566,567,568,569,570,571,572,573,574,575,576,577,578,579,580,581,582,583,584,585,586,587,588,589,590,591,592,593,594,595,596,597,598,599,600,601,602]. This review summarizes all the papers of CVPR2015 we read as the first project of cvpaper.challenge. In this paper, we will describe the characteristics of CVPR2015 and discuss the trends and leading methods used in three areas; namely, recognition, 3D, and imaging/image processing. Further, we will enumerate the proposed datasets and new research problems presented at the conference and propose the concept of “DeepSurvey”. Finally, we will give a summary and discuss future steps. We would like to stress, however, that this paper mainly focuses on a survey of the research trends, and does not cover the details of all the 602 papers, which are beyond the scope of this paper.

2 Related initiatives and positioning of this project

An example of a related initiative is the Japanese CHI Study Group that undertook to read all the papers presented at the ACM CHI Conference on Human Factors in Computing Systems, the top conference for user interfaces [603]. In 2015, the Study Group was held jointly in Tokyo and Hokkaido using a remote conference system to read within one day all the 485 papers presented at CHI2015. Speakers were assigned one session each and introduced one paper in approximately 30 seconds. The CHI Study Group started in 2006, and is on its 10th year in 2015. This initiative is a very effective way to grasp trends in the user interface domain, which is a very progressive field. It is noteworthy that several Japanese researchers have received the Best Paper and Honorable Mention awards at CHI2015.

Considering the rapid progress in the field, the authors focused on “reading all the papers included in the international conference and summarizing trends through the project,” as well as on listening to the introduction of the papers by project members. The CHI Study Group, therefore, serves its role in terms of covering all the papers. However, understanding the trends in a research field entails comprehensively reading all papers outside particular domains and holding discussions after reading the papers. We believe that trends can only be properly identified through discussions within the small group that undertook to read all the papers.

3 Trends at CVPR2015

This paper is mainly divided into three main areas; namely, (i) Recognition, (ii) 3D, and (iii) Imaging/Image processing. Before we proceed with discussing the details of each area, we will first explain the features and perspectives gleaned from the titles of the papers and of the papers selected for oral presentation.
3.1 CVPR2015, as seen from the titles of the papers

Figure 1 shows a visualization of the titles of papers at CVPR2015 using Wordle [608], a word-visualization service. We see that other than “image” as the most frequent word, which is expected, we also see that the words “deep,” “learning,” “object,” and “recognition” are very prominent. As can be seen from this word visualization and goes without saying, deep learning (DL) is a widely used tool in image recognition in researches presented at CVPR2015. Results of a search among CVPR2015 papers showed that the term was found in 250 out of the 602 papers. Although deep learning was mainly used for object recognition, since R-CNN [609] was proposed, it also came to be used more for object detection. Also, with detection and recognition becoming more accurate, there are now more research initiatives to address semantic segmentation, which is an even more difficult problem. Meanwhile, the paper that received the Best Paper Award dealt with the method called DynamicFusion [38], which pertains to real-time construction of high-resolution 3D models. The research was recognized for being able to successfully reconstruct non-rigid models in real time, in addition to being able to reconstruct in higher resolution than KinectFusion [610]. As shown by the prominence of the words “reconstruction,” “depth,” and “shape,” 3D research is also progressing at a steady pace. Also, deep learning is also being applied in stereo matching and 3D features, indicating the widening applications of deep learning. Even in imaging technologies, deblurring and sensing technologies are being updated, wherein examples of applications of feature extraction through deep learning were reported.

3.2 CVPR2015 as seen from the oral presentations

The 12 oral sessions are presented in CVPR2015 as below:

- CNN Architectures
- Depth and 3D Surfaces
- Discovery and Dense Correspondences
- 3D Shape: Matching, Recognition, Reconstruction
3.2.1 Recognition

With CNN as the most widely used approach in deep learning, the theme of the first oral presentation was on CNN Architecture. First we would like to mention about GoogleNet [1], the winner of the ILSVRC2014 with a record 6.67% top-5 error rate. GoogleNet is a 22-layers CNN architecture, where convolutional structures are recursively connected to make a deep structure. Some presentations dealt with addressing the conventional problems in image recognition through deep learning, such as methods to implement multi-layer or multi-instance learning in order to achieve flexibility through shape change [43], implementing optimization and repetition through Bayesian Optimization in the periphery of highly likely candidate regions in order to correct errors in object detection using R-CNN [28], proposal of a robust expression to withstand positional invariability and deformations [108], etc. Meanwhile, Nguyen et al. automatically generated features that are mistakenly recognized by deep learning and showed that CNN features are not universal [47].

3.2.2 3D

With DynamicFusion [38] at the head of the list, new methods on 3D were proposed. DynamicFusion is a method for conducting more precise 3D shape restoration in real time by chronologically integrating depth images obtained through Kinect and other methods. SUN RGB-D [62] was proposed as a large-scale dataset that captures indoor space in three-dimension. Their dataset contains a total of 10,335 RGB-D images, and they presented several important issues on the topic. Research on occlusion using 3D models has also progressed. Xiang et al. [207] used 3D Voxel Patterns (3DVP) to carry out 3D detection and enabled detection of missing parts of objects where occlusion or defects have occurred, using a model-based method. Reconstructing the World in Six Days is an example of research on large-scale space [356]. They carried out 3D reconstruction through world-scale SfM of 100 million images of worldwide landmarks found in flickr. Hengel et al. used silhouette obtained from multiple images to carry out meaningful resolution of each part of the 3D model [96]. The 3D structure was realized through Block World [611]. Richter et al. proposed a method for discriminatively resolving Shape-from-Shading [123]. Albl et al. came up with a mechanism for properly operating, in global shutters as well as in local shutters, the perspective-n-point problem (PnP), which is considered important in SfM, inferring camera position, and calculating odometry [249]. Due to the problem of having an arbitrary rotation matrix, in their proposed method, they proposed an R6P algorithm to make more dense calculation of rotation matrix. Song et al. proposed a method to infer the 3D location of vehicles using monocular cameras [404]. Kulkarni et al. proposed Picture (Probabilistic Programming Language), which is a stochastic expression of 3D modeling, to enable expressing a more complicated generation model [475]. Wang et
al. conducted 3D scene recognition in outdoor environment using GPS positional information as preliminary data [429]. CRF was applied to assign segments and their significance to 3D positional information. Barron et al. devised an optimized method to enable generation of effective stereo images [483]. Use of Defocus and Fast Bilateral Filter eliminates the need for calculating all corresponding points. Wang et al. devised a method for searching the 3D model from the 2D sketch [204].

A sketch image as seen from multiple perspectives is generated from the 3D model of one sample, and a 3D model is searched through comparison with inputted sketch image and presented to the user. Brubaker et al. carried out 3D molecular model reconstruction of high-resolution image from low-resolution image using electron cryomicroscopy [336]. Chin et al. realized improvement of robust matching such as RANSAC through optimization by A*search [262].

3.3.3 Image processing/imaging

In regard to image processing and imaging, advances in research through new themes were seen. For example, Tanaka et al. presented their research resolution of paintings that are physically separated into multiple layers, such as pencil sketches or colored paintings [592], enabling the extraction of even deeper components. [554] presented the problem of finding an efficient border ownership (where the borderline is, whether an area is part of the background or foreground) in 2D images. The authors addressed the problem by using structural random forests (SRF) to differentiate borders. The problem regarding realizing photometric stereo under natural light rather than controlled light sources was also presented [489]. In order to apply photometric stereo in outdoor environment, the authors assumed a hemispherical experimental space and used GPS timestamp as preliminary information, and separately carried out light source estimation of sunlight. There were several proposals regarding the problem of inferring depth images from input images and videos, as well as a paper on simultaneous solution for image correction from fogged images and for inference of depth images [540].

Research on super-resolution was also included in the oral presentations [563]. The authors used self-similarity based super-resolution, and at the same time carried out inference of affine transformation parameters and localized shape variations.

3.3 CVPR2015 as seen by area of study

In the previous section we looked at CVPR2015 based on the titles and papers selected for oral presentation. In this section we will enumerate papers in more detail by area of study. Here we will comprehend the current trends in the field of computer vision by looking at all papers, regardless of whether they were presented orally or as posters.

3.3.1 Recognition

Deep learning architecture. We will cite papers that discuss the overall structure, as well as those that deal with problem-based structures, parameter adjustments, and architecture evaluation. Two examples of papers that discuss overall structure are on GoogLeNet [1] and DeepID-Net [261]. DeepID-Net uses Deformation Constrained (Def) pooling as alternative to max pooling and average pooling in order to improve expressiveness against changes in shape and position, as in DPM [616], contributing to improvement of accuracy in object detection. There were also many examples of attempts to carry out improvements under the framework of existing CNN methods [48, 405, 93]. Wan et al. combined the advantages of DPM and CNN and, further, implemented Non-maximum Suppression (NMS) in order to correct effects of positional discrepancies [93]. DPM is a method for preserving parts and position in latent variables, while CNN has the advantage of being able to automatically learn features that are useful for object recognition. Other papers dealt with the characteristics of CNN [47, 43, 108], increasing speed of learning [88], initiatives to search for parameters [580], and visualization of features [562]. Leon et al. carried out robust CNN feature expression to address image rotation by implementing a transformation layer for geometric transformation of convoluted features [108]. Liu et al. succeeded in reducing computational complexity and CNN calculation time by implementing sparse representation to address convolution [88]. They succeeded in significantly reducing calculation time by sparsifying kernels computed at every convolution, and improved calculation to enable operation even on a CPU. He et al. studied depth of structure, filter size, stride, and other trade-offs pertaining to CCN architectural parameters [580], and showed that depth is important. Other papers dealt with improvement of convolution layers [365], method to calculate similarity of patches [355, 471], and research on morphing under the CNN framework [168]. Liang et al. claimed that better features can be obtained if CNN convolution frameworks are recursively convoluted [365]. This structure is called Recurrent Convolutional Layet (RCL). In MatchNet, architecture is configured for the purpose of measuring similarity between patches, and is partitioned to a network for generating features through pooling and convolution of four layers.
and a network for evaluating similarity through total combination of three layers [355]. Zagoruyko et al. also discussed a framework for calculating patch similarities in CNN [471]. They extracted the features based on convolutions of paired patches and calculated similarity in the later layers.

**Human recognition.** We will introduce papers in Human Recognition by dividing them into Face Recognition, Pedestrian Detection, Human Tracking, Pose Estimation, Action Recognition, Event Recognition, Crowd Analysis, Egocentric Vision, and Person Re-identification.

First, in **face recognition**, FaceNet was presented as a system for handling high-precision recognition [89]. DeepFace, which has been recently proposed in 2014 [619], brought about significant improvements in accuracy, but FaceNet has achieved an even higher accuracy than DeepFace. Sun et al. improved their conventional face recognition, DeepNet [623], and applied features extracted from early convolution layers to improve face recognition accuracy particularly of face profiles and occlusions [314].

In **pedestrian detection**, Tian et al. were able to improve accuracy by combining CNN features and attributes for detection of pedestrians [550]. They accomplished this by including other attributes, such as positional relationships between pedestrians and environment, as well as learning of pedestrians and backgrounds. Honsang et al. implemented evaluation of features using CNN to carry out pedestrian detection [441].

In **pose estimation**, a research on marker-less motion capture using CNN features was presented [412]. For practical use, it is possible to significantly reduce installation costs if estimation can be implemented through marker-less MoCap using 23 cameras.

In **human tracking**, there were reports featuring more advanced methods. Milan et al. were able to simultaneously carry out tasks of chronological area estimation and positioning by using Superpixel and CRF [585]. They established a method for combining low- and high-level information and finely dividing background and foreground. A method for carrying out accurate tracking of multiple objects using Target Identity-aware Network Flow (TINF), which probabilistically resolves network nodes, was also presented [125]. The method constructs the optimum network using graph theory and carries out optimization through Lagrangian relaxation optimization. In action recognition, Gkioxari et al. used R-CNN [609] as basis for proposing a mechanism for recognizing actions, including position of the human subjects [83]. In order to extract the action area, candidate areas where extracted from an assembly of optical flows to extract CNN-based features. And in order to extract features from chronological actions, convolution was implemented for chronological images that stored optical flows and RGB visible images. To improve accuracy, researchers proposed a method based on Dense Trajectories (DT) [620, 621] and on TDD, an action descriptor that combines CNN features [583]. In regard to the DT-based method, researchers adopted HOG, HOF, and MBH to accurately recognize actions, as well as applied CNN features to action recognition through normalization of the feature map. Lan et al. proposed Multi-skip Feature Stacking (MIFS), a method for extract features by configuring multiple gradations to a chronological offset [23].

In **event recognition**, architecture specialized for event recognition called Deep Event Network (DevNet) was proposed [279]. The system enabled extracting not only pre-defined events, but also clues for important chronological events. Xiong et al. carried out recognition of complex events by combining multiple identification results and factors for still images and combined CNN features and results of object/human/face detection results to recognize events [175]. Shu et al. carried out event recognition from aerial images taken using unmanned aerial vehicles (UAV) [495]. They proposed a Space-time AND-OR Graph to analyze various clues from images of drones, such as positional adjustment of images containing egomotion, group action recognition, and human interaction.

In **crowd analysis**, a mechanism that allows cross-scene crowd counting was proposed [91]. They used a CNN model that allows switching the crowd density map and human count model. Although these two models are different, they are correlated and complement each others accuracy. Yi et al. analyzed crowd models from videos taken from surveillance cameras and measured routine pedestrian path directions [378]. They predicted crowd attributes and pedestrian destinations and enabled detection of abnormal actions as well as prediction of paths taken to reach destinations.

A method for editing ones own videos taken using **egocentric vision** was also proposed [590]. Research to solve face recognition problems, such as recognition of severely occluded faces and small and far faces in images, has progressed. Huang et al. proposed a hand region segmentation method for egocentric vision to determine what tasks the person taking the video is performing [73].

**Person re-identification** deals with the problem of personal authentication between different cameras for surveillance and other cameras. Shi et al. inferred semantic attributes regarding humans and clothing at the patch level, and applied them in person re-identification [453]. They obtained clothing and other external appearance features and were able to improve expressivity by us-
ing attributes. Chen et al. carried Multiple Similarity Function Learning using PCA compression color and texture features from images with segregated regions [171]. Zheng et al. evaluated effectiveness of features and enabled feature integration needed for Re-ID using Late Fusion [190]. Person re-identification using low-resolution images was also addressed [76]. Generally, images from surveillance cameras are of poor quality, and to address this, Jing et al. carried out super-resolution to propose a mechanism for improving performance even for low-resolution images. Neural network architecture to improve robustness against feature variations between cameras was also proposed [423]. Given a pair of images as input, the authors used the difference of activation functions extracted from each patch after convolution and pooling as features for recognition.

Object recognition and detection. The problem of recognizing objects appearing in images is currently an intensively studied area. This section also deals with object detection that includes recognition of position, scene recognition, search of hashed images, as well as fine-grained image recognition. Papers on object recognition have dramatically increased after AlexNet was proposed [614] at ILSVRC2012, and object recognition has also been applied to scene recognition and other problems. Research on object detection expanded after the proposal of R-CNN [609]. These trends are clearly evident in CVPR2015.

A study was conducted to improve accuracy and streamline recognition by carrying out selection of CNN factors [106]. Association Rules [626] widely used in the data mining field were applied, and only features that are useful for identification were selected as a subset from among the CNN feature space. In object detection, there were many researches addressing the problem of inaccurate localization, which is one of the vulnerabilities of R-CNN. As previously mentioned, Zhang et al. proposed a method for optimization to correct inaccurate localization in R-CNN to address this vulnerability [28]. Tsai et al. considered the diversity of internal changes and variations of objects for detection, and compensated for inaccurate localization by improving feature pooling [80]. Oquab et al. used weakly supervised learning to investigate solutions for discrimination and localization of objects based only on labeling of image levels [75]. Fine-grained image discrimination is a problem that entails more detailed classification of objects, such as dog breeds or vehicle types. Due to high visual similarity of objects, such detailed classification is very difficult to carry out. It was found that adaptively extracting features useful for discrimination by dividing images into parts and extracting features only from particular regions is an effective method [630]. Using CNN architecture, Xiao et al. extracted candidate patches from major categories (e.g. dog, bird) and detailed categories (e.g. fine classification of dogs and birds) in a layered structure, and simultaneously implemented feature selection and discrimination [92]. Xie et al. carried out learning by applying multitask learning in multiple structured classes as well as in limited task data extensions [287]. They succeeded in simultaneously learning relationships through multitask learning of major and minor classifications. Lin et al. [182] proposed Deep Localization, Alignment and Classification (DeepLAC) as a mechanism to correct changes in regional position and angles, which is needed for fine-grained image recognition, within the back-propagation algorithm framework.

Segmentation. Segmentation requires implementing object recognition at the pixel level, making it a difficult procedure in terms of distinguishing borders between foreground and background. The number of papers dealing with semantic segmentation, which deals with assignment of meaning to segmentation areas, has increased.

Hariharan et al. demonstrated the increase in accuracy of semantic segregation by using features extracted in the middle layers, not only from the fully connected layer, in regard to CNN architecture [49]. In particular, they used the 2nd pooling layer, the 4th convolution layer, and the 7th fully connected layer, and by combining these they were able to simultaneously implement low-, mid-, and high-level feature expression. In saliency-based segmentation, a method was proposed for extracting multi-scale CNN features [591]. Itti et al’s saliency model is well known [627], and, although they conducted multi-scale calculations, Li et al. extracted saliency and applied it segmentation by replacing CNN features. Although it overlaps with 3D reconstruction, we would like to mention here that Martinovic et al. proposed research for implementing semantic segmentation of 3D urban models [482].

Data generation. Data generation is an important issue in addressing recognition problems. In this section we will cite papers on data collection and selection. Hattori et al. generated learning images for pedestrian detection [413]. They conducted learning of 36 types of pedestrians, various kinds of walking, and occlusion patterns using CG. Russakovsky et al. cited an annotation method leveraging crowdsourcing, in order to efficiently and accurately detect objects [231]. The method deals with the usability and accuracy of labeling and is aimed at minimizing human annotation costs, wherein machines and humans interactively carry out annotation based on results from baseline recognition equip-
ment. Xiao et al. discussed a framework for efficient labeling and learning, in an effort to reduce annotation operations for massive data [292].

3.3.2 3-Dimension

There were also many examples of applications of CNN even for 3D object recognition. Fang et al. proposed Deep Shape Descriptor (DeepSD) as a method for expressing 3D shapes [252]. They proposed a robust 3D feature that can handle structural variations in shape, noise, and shapes that include three-dimensional incompleteness. Xie et al. proposed DeepShape, a CNN feature to address problems in 3D object matching and retrieval [139]. They used a shape descriptor based on an auto-encoder to search 3D shapes. Abdelrahman et al. proposed a 3D non-rigid texture descriptor based on Weighted Heat Kernel Signature (W-HKS) [21]. There was also a proposal for a mechanism to extract information useful for recognition even from a limited learning sample using Deep Boltzmann Machine (DBM) and design of object recognition features through RGB-D [327]. They proposed an effective descriptor even for complex 3D objects by combining geometric shape information as well as color information.

In RGB-D input, a problem was reported in giving tasks, such as 3D recognition and inferring positions that can be grasped by robots, in complex indoor environment [498]. Superpixel was applied as a preliminary processing step, and recognition of cuboid models and spatial smoothing through Conditional Random Fields (CRF) was carried out. Matsuo et al. also proposed a method for enhancing depth images (particularly planes) by combining low-resolution depth images and high-resolution RGB images [387]. They adjusted position and connection of tangent planes in 3D space and used JBU filter to reconstruct rough surfaces. Gupta et al. conducted research on extracting object position and 3D segmentation results from RGB-D image input [512]. They expressed object features through learning by CNN of surface normal line images. They then roughly estimated object pose based on 3-layered CNN and inferred detailed object pose and segment by comparison with the 3D model.

3.3.3 Image processing/imaging

CNN was also used for blur removal [84]. Non-uniform motion blurs arising from shaking of camera, etc. were corrected through learning of blurred/non-blurred patch pairs. There was also a research on fusion of multiple kernels [41]. The authors adopted a method using kernels for fusing multiple deblurring methods in order to develop a more advanced blur removal method. By using Gaussian Conditional Random Fields (GCRF), they were able to carry out kernel fusion based on learning. Eriksson et al. proposed a method for noise removal that takes sparsity into consideration [363]. To solve the k-support norm optimization and normalization problem, Eriksson et al. carried optimization by considering this problem as the minimum convex set that includes the set given as Convex Envelopes. Research on blur removal for videos was also reported [437]. There are two methods for blur removal for videos. One is by independently removing blur within the frames and splicing the frames together. The other is by inferring camera motion between frames. Zhang et al. combined these two methods.

In regard to the problem of super-resolution, a method using Self-Similarity based Super-Resolution was reported [563]. The method simultaneously infers affine transformations and localized shape variations. Comparison with external/internal dictionaries enabled mapping to clear images. A method using a reference dictionary that accommodates shape variations was also reported to address the super-resolution problem for single images [587]. Gradient Ridge Image processing was performed as a preliminary processing step, and resolution was enhanced through matching with the dictionary. Schuler et al. solved the single-image super-resolution problem as a linear regression problem using Random Forests [410].

A method for inferring shadow regions using CNN was reported for basic algorithms in image processing [225]. Shen et al. also proposed DeepContour, which is a CNN architecture for contour detection [431]. DeepContour involves learning contour/non-contour regions and is composed of a 6-layered architecture (four convolution layers and two fully connected layers). DeepEdge was also proposed as an application of CNN architecture for edge detection [474]. DeepEdge carries out more accurate edge detection by using higher-level features. Experimentally, they were able to show that unlike CannyEdge, where there was noise contamination, DeepEdge was able to better remove backgrounds as well as extract edges from objects. Teo et al. also proposed a method for effectively extracting borderlines in 2D images [554]. By using Structural Random Forests (SRF), they were able to rapidly determine where the borders are, and whether the area belongs to the background or the foreground. A research on the application of Linear Spectral Clustering (LSC) to Superpixel methods was also presented [148].

In device research, a hyperspectral camera that can acquire chronological images was proposed [535]. Sequences of multiple hyperspectral cameras were alter-
nately complemented, and image reconstruction based on dictionary learning was conducted, in order to obtain clear images even at high-speed (100 fps) observation. Ti et al. developed a ToF sensor using a monocular camera and LED [469]. They developed the ToF sensor by attaching a total of four LEDs to the upper, lower, right, and left sides of the camera and capturing the reflection of LED light using the camera. To improve accuracy of ToF cameras, Naik et al. resolved the problem of Multipath Interference (MPI), where multiple optical reflections appear and are mixed up in the pixel [9]. MPI also occurs in natural scenes, such as in an environment where multiple reflected lights occur or reflected light is diffused. These reflections were divided into Phase and Amplitude, both directly and globally, in order to reduce depth image errors due to MPI. Ye et al. proposed an enhanced Kinect sensor by attaching Ultrasonic Sensor to Kinect [529]. They inferred the plane by applying Bayesian Network to the inference point obtained through the Ultrasonic Sensor.

3.3.4 Datasets

In this section, we will discuss new research problems as well as research on datasets.

Datasets. An example of a dataset is the SUN RGB-D, an expansion of SUNdatabase (which is a problem in scene recognition mentioned in the previous section) to RGB-D [62]. It is an attempt to expand the data set into more advanced scene recognition, such as segmentation and detection of objects within scenes, other than merely for recognition. A similar research problem is on the dataset for estimating indoor layout proposed by Liu et al [370]. The dataset for indoor environment included information on the entire room, walls, doors, windows, and their positional information. A research for outputting detailed explanations of medical images was also reported [119]. This research problem pertained to outputting sentence descriptions from an input of medical images. Detailed explanations of symptoms are generated by learning in pairs the actual medical images and the corresponding medical examination results. There was also an attempt to increase recognition capability by creating a much larger-scale dataset in the field of fine-grained recognition [65]. NABirds is a dataset for fine-grained recognition of birds, the scope of which was expanded by increasing the number of classes. There was also a report on a dataset for categorizing cars [430]. The study provided data for fine-grained classification of cars, which previously were only categorized into the class called cars.

There was also a study on creating data for detection of pedestrians through the use of images containing a higher amount of information. Hwang et al. used a hyperspectral camera to acquire richer image information in order to improve detection of pedestrians at nighttime as well as daytime [113]. A dataset was also proposed for analyzing each person in a crowd by focusing on the spectators rather than on the sport itself [222]. They analyzed individual reactions of persons in a crowd, categorized crowds, and determined the type of spectators. In regard to pedestrian detection, a dataset was proposed for estimating gender, age, weight, clothing, etc., of pedestrians as well their location [594]. This dataset is intended for fine-grained recognition of persons. Thus, there was more focus on addressing fine-grained detection of pedestrians. There is more research being conducted on generation of image descriptions, with one oral session devoted to the topic. In particular, Rohrbach et al. proposed a dataset for movie description [347]. For action recognition datasets, Heilbron et al. published a dataset called ActivityNet [105], which is a large-scale dataset similar to ImageNet and includes a significantly greater amount of data and action variations. The dataset includes 203 trimmed data classes and 137 untrimmed classes, for a total of 849 video hours. Also in action recognition, Xu et al. proposed a dataset that maps attributes in advance to actors and actions [246].

New research problems. Here we introduce new research problems proposed at CVPR2015. Lin et al. proposed the research problem of identifying locations of aerial images using images taken on the streets as query [542]. Although ground images and aerial images are completely different in nature, the authors presented a possible approach to the problem by proposing Where-CNN. Akhter et al. conducted estimation of 3D human pose from 2D joint angles, and by adding a joint angle limit they were able to add a process for reducing poses with inscrutable motion [158]. Peng et al. proposed two new aspects on human emotions predicted from images [94]. There was a paper on detecting persons or animals in a best relationship, i.e., with a high co-occurrence relationship with another person or animal based on Best-Buddies Similarity [220]. The authors proposed a method based on template matching to visualize the co-occurrence relationship. There was also a paper that addressed the problem of identifying very important people (VIP) within a group [526]. The authors used im2text to solve the problem by classifying level of importance of images and texts. Traditional machine learning methods map input and output vectors as pairs, but Wang et al. assigned hidden information to images to further improve flexibility [538]. On the basis of this concept, they proposed that hidden information be handled as features or second objective functions.
Zhang et al. proposed a method to address the problem of counting items in an image as well as finding saliency from images [438]. They claimed that it can be used for egocentric lifelogs and image thumbnails. Not only in sensing, but there will also be a need to carry out person recognition in next-generation camera images whose resolution has been lowered for security and privacy protection. This problem is addressed by Pittaluga et al. by carrying out face and pose recognition that can handle low-resolution images and resist changes in light source, proposing the method to be used for privacy protection [35]. There was also a proposal on object recognition that takes into consideration what kind of tasks are completed using particular tools [310]. The authors constructed 3D models of objects using 3D sensors and inferred the position by which the person carries the object based on joint angle, and measured how the task is being carried out. Measurement was made not only on 2D and 3D images, but they also calculated the impulse strength using voice data. Handling of the tool was inferred based on joint angle trajectory. There was also a proposal for inferring what a store is selling based on the storefront image [185]. Streetview images were used to extract characters through OCR, and ontology from those characters was used to classify stores according to business category.

4 DeepSurvey

We are proposing DeepSurvey (see Figure 2) as a mechanism for the systematization of knowledge, the generation of ideas, and as well as the writing of papers (especially for new research problems) based on an extensive reading of papers. DeepSurvey architecture is devised based on DeepLearning, which has flourished in recent years, and is composed of the following elements:

- Input: Input the papers read (knowledge)
- 1st ideas: Individually generate ideas (from knowledge to ideas)
- 1st discussion: Group discussion (consolidation of ideas)
- 2nd ideas: Generate more ideas based on consolidated ideas
5 Summary and future trends

In this survey we comprehensively read papers presented at CVPR2015 to gain an understanding of the trends in computer vision. Further, we devised DeepSurvey as a mechanism to generate ideas from knowledge and eventually write a paper. We divided the papers into three areas; namely, recognition, 3D, and imaging/image processing, and sought to identify new research areas, as a means to expand the limits of the field. Here we are proposing DeepSurvey, and, going forward, we have started addressing some of its problems.

The authors are sorting out the current issues and believe that conducting surveys that include a study of technologies is essential also for identifying the next research problems. Further, there is a need to gain the ability to view the field from a wider perspective aside from actually testing the survey results to better understand the issues. We hope that this initiative would serve as a useful step towards that end.

References

1. Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich, “Going Deeper With Convolutions”, in CVPR2015.
2. Jen-Hao Rick Chang, Yu-Chiang Frank Wang, “Propagated Image Filtering”, in CVPR2015.
3. Yunchao Gong, Marcin Pawlowski, Fei Yang, Louis Brandy, Lubomir Bourdev, Rob Fergus, “Web Scale Photo Hash Clustering on A Single Machine”, in CVPR2015.
4. Alina Kuznetsova, Sung Ju Hwang, Bodo Rosenhahn, Leonid Sigal, “Expanding Object Detector’s Horizon: Incremental Learning Framework for Object Detection in Videos”, in CVPR2015.
5. Fumin Shen, Chunhai Shen, Wei Liu, Heng Tao Shen, “Supervised Discrete Hashing”, in CVPR2015.
6. Mihir Jain, Jan C. van Gemert, Cees G. M. Snoek, “What do 15,000 Object Categories Tell Us About Classifying and Localizing Actions?”, in CVPR2015.
7. Rahaf Aljundi, Remi Emonet, Damien Muselet, Marc Sebban, “Landmarks-Based Kernelized Subspace Alignment for Unsupervised Domain Adaptation”, in CVPR2015.
8. Wei-Sheng Lai, Juan-Jiun Ding, Yen-Yu Lin, Yung-Yu Chuang, “Blur Kernel Estimation Using Normalized Color-Line Prior”, in CVPR2015.
9. Nikhil Naik, Achuta Kadambi, Christoph Rhemann, Shahram Izadi, Ramesh Raskar, Sing Bing Kang, “A Light Transport Model for Mitigating Multipath Interference in Time-of-Flight Sensors”, in CVPR2015.
10. Simone Frintrop, Thomas Werner, German Martin Garcia, “Traditional Saliency Reloaded: A Good Old Model in New Shape”, in CVPR2015.
11. Patrick Snape, Yannis Panagakis, Stefanos Zafeiriou, “Automatic Construction Of Robust Spherical Harmonic Subspaces”, in CVPR2015.
12. Min-Gyu Park, Kuk-Jin Yoon, “Leveraging Stereo Matching With Learning-Based Confidence Measures”, in CVPR2015.
13. Yao Qin, Huchuan Lu, Yiqun Xu, He Wang, “Saliency Detection via Cellular Automata”, in CVPR2015.
14. Jonas Wülff, Michael J. Black, “Efficient Sparse-to-Dense Optical Flow Estimation Using a Learned Basis and Layers”, in CVPR2015.
15. Carlo Ciliberto, Lorenzo Rosasco, Silvia Villa, “Learning Multiple Visual Tasks While Discovering Their Structure”, in CVPR2015.
16. Zhiwu Huang, Ruiping Wang, Shiguang Shan, Xilin Chen, “Projection Metric Learning on Grassmann Manifold With Application to Video Based Face Recognition”, in CVPR2015.
17. Tianzhu Zhang, Si Liu, Changsheng Xu, Shuicheng Yan, Bernard Ghanem, Narenda Aluja, Ming-Hsuan Yang, “Structural Sparse Tracking”, in CVPR2015.
18. HyeokHyen Kwon, Yu-Wing Tai, Stephen Lin, “Data-Driven Depth Map Refined via Multi-Scale Sparse Representation”, in CVPR2015.
19. Feng Lu, Imari Sato, Yoichi Sato, “Uncalibrated Photometric Stereo Based on Elevation Angle Recovery From BRDF Symmetry of Isotropic Materials”, in CVPR2015.
20. Ran Tao, Arnold W.M. Smeulders, Shih-Fu Chang, “Attributes and Categories for Generic Instance Search From One Example”, in CVPR2015.
21. Mostafa Abdelrahman, Aly Farag, David Swanson, Moumen T. El-Melegy, “Heat Diffusion Over Weighted Manifolds: A New Descriptor for Textured 3D Non-Rigid Shapes”, in CVPR2015.
22. Christopher Zach, Adrian Penate-Sanchez, Minh-Tri Pham, “A Dynamic Programming Approach for Fast and Robust Object Pose Recognition From Range Images”, in CVPR2015.
23. Zhengzhong Lan, Ming Lin, Xuanchong Li, Alex G. Hauptmann, Bhiksha Raj, “Beyond Gaussian Pyramid: Multi-Skip Feature Stacking for Action Recognition”, in CVPR2015.
24. Dongping Li, Kaiming He, Jian Sun, Kun Zhou, “A Geodesic-Preserving Method for Image Warping”, in CVPR2015.
25. Shaoxin Li, Junliang Xing, Shiguang Shan, Honglak Lee, “Improving Object Detection With Deep Convolutional Networks via Bayesian Optimization and Structured Prediction”, in CVPR2015.
26. Marko Ristin, Juergen Gall, Matthieu Guillaumin, Luc Van Gool, “From Categories to Subcategories: Large-Scale Image Classification With Partial Class Label Refinement”, in CVPR2015.
27. Yunsheng Jiang, Jinwen Ma, “Combination Features and Models for Human Detection”, in CVPR2015.
28. Yuting Zhang, Kihyuk Sohn, Ruben Villegas, Gang Pan, Honglak Lee, “Improving Object Detection With Deep Convolutional Networks via Bayesian Optimization and Structured Prediction”, in CVPR2015.
29. Spyridon Leonardos, Roberto Tron, Kostas Daniilidis, Shuicheng Yan, “Shape Driven Kernel Adaptation in Convolutional Neural Network for Robust Facial Traits Recognition”, in CVPR2015.
30. Marko Ristin, Juergen Gall, Matthieu Guillaumin, Luc Van Gool, “From Categories to Subcategories: Large-Scale Image Classification With Partial Class Label Refinement”, in CVPR2015.
31. Chun-Guang Li, Rene Vidal, “Structured Sparse Subspace Clustering: A Unified Optimization Framework”, in CVPR2015.
32. Dongping Li, Kaiming He, Jian Sun, Kun Zhou, “A Geodesic-Preserving Method for Image Warping”, in CVPR2015.
33. Sebastian Kaltwang, Sinisa Todorovic, Maja Pantic, “Lamda-Fusion: Reconstruction and Tracking of Non-Rigid Scenes in Real-Time”, in CVPR2015.
34. Hui Wu, Richard Souvenir, “Robust Regression on Image Manifolds for Ordered Label Denoising”, in CVPR2015.
35. Francesco Pittaluga, Sanjeev J. Koppal, “Privacy Preserving Optics for Miniature Vision Sensors”, in CVPR2015.
36. Junlin Hu, Jiwen Lu, Yap-Peng Tan, “Deep Transfer Metric Learning”, in CVPR2015.
37. Julian Straub, Trevor Campbell, Jonathan P. How, John W. Fisher III, “Small-Variance Nonparametric Clustering on the Hypersphere”, in CVPR2015.
38. Richard A. Newcombe, Dieter Fox, Steven M. Seitz, “DynamicFusion: Reconstruction and Tracking of Non-Rigid Scenes in Real-Time”, in CVPR2015.
39. Yang Li, Jianke Zhu, Steven C.H. Hoi, “Reliable Patch Trackers: Robust Visual Tracking by Exploiting Reliable Patches”, in CVPR2015.
40. Nian Liu, Junwei Han, Dingwen Zhang, Shifeng Wen, Tianming Liu, “Predicting Eye Fixations Using Convolutional Neural Networks”, in CVPR2015.
41. Long Mai, Feng Liu, “Kernel Fusion for Better Image Deblurring”, in CVPR2015.
42. Christian Hane, ?ubor Ladicky, Marc Pollefeys, “Direction Matters: Depth Estimation With a Surface Normal Classifier”, in CVPR2015.
43. George Papandreou, Iasonas Kokkinos, Pierre-Andre Savalle, “Untangling Local and Global Deformations in Deep Learning: Epistemic Convolution, Multiple Instance Learning, and Sliding Window Detection”, in CVPR2015.
44. Yezhou Yang, Cornelia Fermüller, Yi Li, Yiannis Aloimonos, “Grasp Type Revisited: A Modern Perspective on a Classical Feature for Vision”, in CVPR2015.
45. Sheng Huang, Mohamed Elhoseiny, Ahmed Elgammal, Dan Yang, “Learning Hypergraph-Regularized Attribute Predictors”, in CVPR2015.
46. Roozbeh Mottaghi, Yu Xiang, Silvio Savaresi, “A Coarse-to-Fine Model for 3D Pose Estimation and Sub-Category Recognition”, in CVPR2015.
47. Anh Nguyen, Jason Yosinski, Jeff Clune, “Deep Neural Networks Are Easily Fooled: High Confidence Predictions For Unrecognizable Images”, in CVPR2015.
48. Ross Girshick, Forrest Iandola, Trevor Darrell, Jitendra Malik, “Deformable Part Models are Convolutional Neural Networks”, in CVPR2015.
49. Bharath Hariharan, Pablo Arbelaez, Ross Girshick, Jitendra Malik, “Hypercolumns for Object Segmentation and Fine-Grained Localization”, in CVPR2015.
50. Johannes Hofmanninger, Georg Langs, “Mapping Visual Features to Semantic Profiles for Retrieval in Medical Imaging”, in CVPR2015.
51. Stephan Schraml, Ahmed Nabil Belbachir, Horst Bischof, “Event-Driven Stereo Matching for Real-Time 3D Panoramic Vision”, in CVPR2015.
52. Daniel Prusa, “Graph-Based Simplex Method for Pairwise Energy Minimization With Binary Variables”, in CVPR2015.
53. Hangfan Liu, Ruejinn Xiong, Jian Zhang, Wen Gao, “Image Denoising via Adaptive Soft-Thresholding Based on Non-Local Samples”, in CVPR2015.
54. Mingsong Dou, Jonathan Taylor, Henry Fuchs, Andrew Fitzgibbon, Shahram Izadi, “3D Scanning Deformable Objects With a Single RGBD Sensor”, in CVPR2015.
55. Jeffrey Byrne, “Nested Motion Descriptors”, in CVPR2015.
56. Gottfried Graber, Jonathan Balzer, Stefano Soatto, Thomas Pock, “Efficient Minimal-Surface Regularization of Perspective Depth Maps in Variational Stereo”, in CVPR2015.
57. Alexander Shekhovtsov, Paul Swoboda, Bogdan Savchynskyy, “Maximum Persistency via Iterative Relaxed Inference With Graphical Models”, in CVPR2015.
58. Abhishek Sharma, Oncel Tuzel, David W. Jacobs, “Deep Hierarchical Parsing for Semantic Segmentation”, in CVPR2015.
59. Xiaolong Wang, David Fouhey, Abhinav Gupta, “Designing Deep Networks for Surface Normal Estimation”, in CVPR2015.
60. Deqing Sun, Erik B. Sudderth, Hanspeter Pfister, “Layered RGBD Scene Flow Estimation”, in CVPR2015.
61. Miguel A. Carreira-Perpinan, Ramin Raziperchikolaei, “Hashing With Binary Autoencoders”, in CVPR2015.
62. Shuran Song, Samuel P. Lichtenberg, Jianxiong Xiao, “SUN RGB-D: A RGB-D Scene Understanding Benchmark Suite”, in CVPR2015.
63. Chen Fang, Haillin Jin, Jianchao Yang, Zhe Lin, “Collaborative Feature Learning From Social Media”, in CVPR2015.
64. Xiaochun Cao, Chaoqun Zhang, Huazhu Fu, Si Liu, Hua Zhang, “Diversity-Induced Multi-View Subspace Clustering”, in CVPR2015.
65. Grant Van Horn, Steve Branson, Ryan Farrell, Scott Haber, Jessie Barry, Panos Ipeirotis, Pietro Perona, Serge Belongie, “Building a Bird Recognition App and Large Scale Dataset With Citizen Scientists: The Fine Print in Fine-Grained Dataset Collection”, in CVPR2015.
66. Mingqin Shi, Yannis Avrithis, Herve Jegou, “Early Burst Detection for Memory-Efficient Image Retrieval”, in CVPR2015.
67. Wei Zhuo, Mathieu Salzmann, Xuming He, Miaomiao Liu, “Indoor Scene Structure Analysis for Single Image Depth Estimation”, in CVPR2015.
68. Juliet Fiss, Brian Curless, Rick Szeliski, “Light Field Layer Matting”, in CVPR2015.
69. Qian-Yi Zhou, Vladlen Koltun, “Depth Camera Tracking With Contour Cues”, in CVPR2015.
70. Zuzana Kukelova, Jan Heller, Martin Bujnak, Tomas Pajdla, “Radial Distortion Homography”, in CVPR2015.
71. Jonathan Tompson, Ross Goroshin, Arjun Jain, Yann LeCun, Christoph Bregler, “Efficient Object Localization Using Convolutional Networks”, in CVPR2015.
72. Jianping Shi, Li Xu, Jiaya Jia, “Just Noticeable Defocus Blur Detection and Estimation”, in CVPR2015.
73. De-An Huang, Minghuang Ma, Wei-Chiu Ma, Kris M. Kitani, “How Do We Use Our Hands? Discovering a Diverse Set of Common Grasps”, in CVPR2015.
74. Junho Yim, Heechul Jung, ByungIn Yoo, Changkyu Choi, Dusik Park, Junmo Kim, “Rotating Your Face Using Multi-Task Deep Neural Network”, in CVPR2015.
75. Maxime Oquab, Leon Bottou, Ivan Laptev, Josef Sivic, “Is Object Localization for Free? - Weakly-Supervised Learning With Convolutional Neural Networks”, in CVPR2015.
76. Xiao-Yuan Jing, Xiaoke Zhu, Fei Wu, Xinge You, Qiong Long Liu, Dong Yue, Ruimin Hu, Baowen Xu, “Super-Resolution Person Re-Identification With Semi-Coupled Low-Rank Discriminant Dictionary Learning”, in CVPR2015.
77. Hang Yang, Ming Zhu, Yan Niu, Yujing Guan, Zhongbo Zhang, “Dual Domain Filters Based Texture and Structure Preserving Image Non-Blind Deconvolution”, in CVPR2015.
78. Xuan Dong, Boyan Bove, Yu Zhu, Alan L. Yuille, “Region-Based Temporally Consistent Video Post-Processing”, in CVPR2015.
79. Shaqing Ren, Xudong Cao, Yi Chen Wei, Jian Sun, “Global Refinement of Random Forest”, in CVPR2015.
80. Yi-Hsuan Tsai, Onur C. Hamsici, Ming-Hsuan Yang, “Adaptive Region Pooling for Object Detection”, in CVPR2015.
81. Mohammad Rastegari, Hannaneh Hajishirzi, Ali Farhadi, “Discriminative and Consistent Similarities in Instance-Level Multiple Instance Learning”, in CVPR2015.
82. Zhibin Hong, Zhe Chen, Chaohui Wang, Xue Mei, Danil Prokhorov, Dachen Tao, “MULTi-Store Tracker (MUSTer): A Cognitive Psychology Inspired Approach to Object Tracking”, in CVPR2015.
83. Georgia Gkiouzari, Jitendra Malik, “Finding Action Tubes”, in CVPR2015.
84. Jian Sun, Wenfei Cao, Zongben Xu, Jean Ponce, “Learning a Convolutional Neural Network for Non-Uniform Motion Blur Removal”, in CVPR2015.
85. Yao Xiao, Cewu Lu, Efstratios Tsougenis, Yongyi Lu, Chi-Keung Tang, “Complexity-Adaptive Distance Metric for Object Proposals Generation”, in CVPR2015.
86. Xiangyu Zhu, Zhen Lei, Junjie Yan, Dong Yi, Stan Z. Li, “High-Fidelity Pose and Expression Normalization for Face Recognition in the Wild”, in CVPR2015.
87. Masaki Saito, Takayuki Okatani, “Transformation of Markov Random Fields for Marginal Distribution Estimation”, in CVPR2015.
88. Baoyuan Liu, Min Wang, Hassan Foroosh, Marshall Tappen, Marimana Pensky, “Sparse Convolutional Neural Networks”, in CVPR2015.
89. Florian Schroff, Dmitry Kalenichenko, James Philbin, “FaceNet: A Unified Embedding for Face Recognition and Clustering”, in CVPR2015.
90. Xiao Sun, Yichen Wei, Shuang Liang, Xiaou Tang, Jian Sun, “Cascaded Hand Pose Regression”, in CVPR2015.
91. Cong Zhang, Hongsheng Li, Xiaogang Wang, Xiaokang Yang, “Cross-Scale Crowd Counting via Deep Convolutional Neural Networks”, in CVPR2015.
92. Tianjun Xiao, Yichong Xu, Kuiyuan Yang, Jiaxing Zhang, Xuyin Peng, Zheng Zhang, “The Application of Two-Level Attention Models in Deep Convolutional Neural Network for Fine-Grained Image Classification”, in CVPR2015.
93. Li Wan, David Eigen, Rob Fergus, “End-to-End Integration of a Convolution Network, Deformable Parts Model and Non-Maximum Suppression”, in CVPR2015.
94. Kuan-Chuan Peng, Tsuhan Chen, Amir Sadovnik, Andrew C. Gallagher, “A Mixed Bag of Emotions: Model, Predict, and Transfer Emotion Distributions”, in CVPR2015.
95. Edgar Simo-Serra, Sanja Fidler, Francesc Moreno-Noguer, Raquel Urtasun, “Neuroaesthetics in Fashion: Modeling the Perception of Fashionability”, in CVPR2015.
96. Anton von den Hengel, Chris Russell, Anthony Dick, John Bastian, Daniel Pooley, Lachlan Fleming, Lourdes Agapito, “Part-Based Modelling of Compound Scenes From Images”, in CVPR2015.
97. Olga Vekeler, “Efficient Parallel Optimization for Potts Energy With Hierarchical Fusion”, in CVPR2015.
98. Michael S. Ryoo, Brandon Rothrock, Larry Matthies, “Regularization of Singular Values for Robust Subspace Matrix Completion”, in CVPR2015.
99. Artiom Kovnatsky, Michael M. Bronstein, Xavier Bresson, Pierre Vandergheynst, “Functional Correspondence by Compliant Approximate Image Codes”, in CVPR2015.
100. Eunwoo Kim, Minsk Lee, Soughwai Oh, “Elastic-Net Regularization of Singular Values for Robust Subspace Learning”, in CVPR2015.
101. Da Kuang, Alex Gittens, Raffay Hamid, “Hardware Compliant Approximate Image Codes”, in CVPR2015.
102. Avishek Chatterjee, Venu Madhav Govindu, “Photometric Refinement of Depth Maps for Multi-Albedo Objects”, in CVPR2015.
103. Christoph H. Lampert, “Predicting the Future Behavior of a Time-Varying Probability Distribution”, in CVPR2015.
104. Anna Khoreva, Fabio Galasso, Matthias Hein, Bernt Schiele, “Classifier Based Graph Construction for Video Segmentation”, in CVPR2015.
105. Fabian Caba Heilbron, Victor Escorcia, Bernard Ghanem, Juan Carlos Niebles, “ActivityNet: A Large-Scale Video Benchmark for Human Activity Understanding”, in CVPR2015.
150. Guangcong Zhang, Patricio A. Vela, “Good Features to Track for Visual SLAM”, in CVPR2015.
151. Phillip Isola, Joseph J. Lim, Edward H. Adelson, “Discovering States and Transformations in Image Collections”, in CVPR2015.
152. Junhwa Hur, Hwasup Lim, Changsoo Park, Sang Chul Ahn, “Generalized Deformable Spatial Pyramid: Geometry-Preserving Dense Correspondence Estimation”, in CVPR2015.
153. Amelie Royer, Christoph H. Lampert, “Classifier Adaptation at Prediction Time”, in CVPR2015.
154. Simone Meyer, Oliver Wang, Henning Zimmer, Max Gross, Alexander Sorkine-Hornung, “Phase-Based Frame Interpolation for Video”, in CVPR2015.
155. Si Liu, Xiaodan Liang, Luoqi Liu, Xiaohui Shen, Jianchao Yang, Changsheng Xu, Liang Lin, Xiaochun Cao, Shuicheng Yang, “Matching-CNN Meets KNN: Quasi-Parametric Human Parsing”, in CVPR2015.
156. Sebastian Haner, Kalle Astrom, “Absolute Pose for Cameras Under Flat Refractive Interfaces”, in CVPR2015.
157. Alex Yong-Sang Chia, Udana Bandara, Xiangyu Wang, Parametric Human Parsing”, in CVPR2015.
158. Fereshteh Sadeghi, Santosh K. Kumar Divvala, Ali Farhadi, “ViSKE: Visual Knowledge Extraction and Question Answering by Visual Verification of Relation Phrases”, in CVPR2015.
159. Xinzhi Du, David Doermann, Wael Abd-Almageed, “A Graphical Model Approach for Matching Partial Signatures”, in CVPR2015.
160. Hao Fang, Saurabh Gupta, Forrest Islanda, Rupesh K. Srivastava, Li Deng, Piotr Dollar, Jianfeng Gao, Xiaoqiang He, Margaret Mitchell, John C. Platt, C. Lawrence Zitnick, Geoffrey Zweig, “From Captions to Visual Concepts and Back”, in CVPR2015.
161. Liping Jing, Liu Yang, Jian Yu, Michael K. Ng, “Semi-Supervised Low-Rank Mapping Learning for Multi-Label Classification”, in CVPR2015.
162. Bolei Zhou, Vignesh Jagadeesh, Robinson Piramuthu, “ConceptListener: Discovering Visual Concepts From Weakly Labeled Image Collections”, in CVPR2015.
163. Mohammad Rastegari, Cem Keskin, Pushmeet Kohli, Ashish Shahram Izadi, “Computationally Bounded Retrieval”, in CVPR2015.
164. Shuang Yang, Chunfeng Yuan, Baoxin Wu, Weiming Hu, Fangshi Wang, “Multi-Feature Max-Margin Hierarchical Bayesian Model for Action Recognition”, in CVPR2015.
165. Yu-Xiong Wang, Martial Hebert, “Model Recommendation: Generating Object Detectors From Few Samples”, in CVPR2015.
166. Alexey Dosovitskiy, Jost Tobias Springenberg, Thomas Brox, “Learning to Generate Chairs With Convolutional Neural Networks”, in CVPR2015.
167. Philipp Krahenbuhl, Vladlen Koltun, “Learning to Propose Objects”, in CVPR2015.
168. Guangcong Zhang, Patricio A. Vela, “Good Features to Track for Visual SLAM”, in CVPR2015.
169. Hae-Gon Jeon, Jaesik Park, Gyeongmin Choe, Jinsun Park, Yunsu Bok, Yu-Wing Tai, In So Kweon, “Accurate Depth Map Estimation From a Lenslet Light Field Camera”, in CVPR2015.
170. Fang Zhao, Yongzheng Huang, Liang Wang, Tieniu Tan, “Deep Semantic Ranking Based Hashing for Multi-Label Image Retrieval”, in CVPR2015.
171. Dapeng Chen, Zejian Yuan, Gang Hua, Nanming Zheng, Jingdong Wang, “Similarity Learning on an Explicit Polynomial Kernel Feature Map for Person Re-Identification”, in CVPR2015.
195. Jianyu Wang, Alan L. Yuille, “Semantic Part Segmentation Using Compositional Model Combining Shape and Appearance”, in CVPR2015.
196. Zhongwen Xu, Yi Yang, Alex G. Hauptmann, “A Discriminative CNN Video Representation for Event Detection”, in CVPR2015.
197. Akihiko Torii, Relja Arandjelovic?, Josef Sivic, Masatoshi Okutama, Tomas Pajdla, “24/7 Place Recognition by View Synthesis”, in CVPR2015.
198. Arturo Deza, Devi Parikh, “Understanding Image Vitality”, in CVPR2015.
199. Makarand Tapaswi, Martin Bauml, Rainer Stiefelhagen, “Book2Movie: Aligning Video Scenes With Book Chapters”, in CVPR2015.
200. Hui Chen, Jiangdong Li, Fengjun Zhang, Yang Li, Hongan Wang, “3D Model-Based Continuous Emotion Recognition”, in CVPR2015.
201. Sakrapee Paisitkriangkrai, Chunhua Shen, Anton van den Hengel, “Learning to Rank in Person Re-Identification With Metric Ensembles”, in CVPR2015.
202. Yonggang Qi, Yi-Zhe Song, Tao Xiang, Honggang Zhang, Timothy Hospedales, Yi Li, Jun Guo, “Making Better Use of Edges via Perceptual Grouping”, in CVPR2015.
203. Jeong-Ryun Lee, Kuk-Jin Yoon, “Real-Time Joint Estimation of Camera Orientation and Vanishing Points”, in CVPR2015.
204. Fang Wang, Le Kang, Yi Li, “Sketch-Based 3D Shape Retrieval Using Convolutional Neural Networks”, in CVPR2015.
205. Na Tong, Huchuan Lu, Xiang Ruan, Ming-Hsuan Yang, “Salient Object Detection via Bootstrap Learning”, in CVPR2015.
206. Abhijit Bendale, Terrance Boult, “Towards Open World Recognition”, in CVPR2015.
207. Yu Xiang, Wengun Choi, Yunqiang Lin, Silvio Savarese, “Data-Driven 3D VoCoal Patterns for Object Category Recognition”, in CVPR2015.
208. Zhirong Wu, Shuran Song, Aditya Khosla, Fisher Yu, Lingxi Zhuang, Xiaoou Tang, Jianxiong Xiao, “3D ShapeNets: A Deep Representation for Volumetric Shapes”, in CVPR2015.
209. Kuang-Jui Hsu, Yen-Yu Lin, Yung-Yu Chuang, “Robust Image Alignment With Multiple Feature Descriptors and Matched-Reduced Neighborhoods”, in CVPR2015.
210. Brendan F. Klare, Ben Klein, Emma Taborsky, Austin Blanton, Jordan Cheney, Kristen Allen, Patrick Grother, Alan Mah, Mark Burge, Anil K. Jain, “Pushing the Frontiers of Unconstrained Face Detection and Recognition: IARPA Janus Benchmark A”, in CVPR2015.
211. Michael W. Tao, Pratul P. Srinivasan, Jitendra Malik, Szymon Rusinkiewicz, Ravi Ramamoorthi, “Depth From Shading, Defocus, and Correspondence Using Light-Field Angular Coherence”, in CVPR2015.
212. Xiao-Ming Wu, Zhenguo Li, Shih-Fu Chang, “New Insights Into Laplacian Similarity Search”, in CVPR2015.
213. Swara Tariq, Hassan Foroosh, “Feature-Independent Context Estimation for Automatic Image Annotation”, in CVPR2015.
214. Abhishek Kar, Shubham Tulsiani, Joao Carreira, Jitendra Malik, “Category-Specific Object Reconstruction From a Single Image”, in CVPR2015.
215. Hang Su, Zhaozheng Yin, Takeo Kanade, Seungil Huh, “Active Sample Selection and Correction Propagation on a Gradually-Augmented Graph”, in CVPR2015.
216. Xiangyu Zhang, Jianhua Zou, Xiang Ming, Kaiming He, Jian Sun, “Efficient and Accurate Approximations of Non-linear Convolutional Networks”, in CVPR2015.
217. Gunhee Kim, Seungwhan Moon, Leonid Sigal, “Ranking and Retrieval of Image Sequences From Multiple Paragraph Queries”, in CVPR2015.
218. Fan Zhang, Feng Liu, “Casual Stereoscopic Panorama Stitching”, in CVPR2015.
219. Andras Bodis-Szomoruz, Hayko Riemenschneider, Luc Van Gool, “Superpixel Meshes for Fast Edge-Preserving Surface Reconstruction”, in CVPR2015.
220. Tali Dekel, Shaol Oron, Michael Rubinstein, Shai Avidan, William T. Freeman, “Best-Buddies Similarity for Robust Template Matching”, in CVPR2015.
221. Tatsunori Tanai, Yasuyuki Matsushita, Takeshi Naeura, “Superdifferential Cuts for Binary Energies”, in CVPR2015.
222. Davide Coniglio, Paolo Rota, Francesco Setti, Chiara Bassetti, Nicola Conci, Nicu Sebe, Marco Cristani, “The S-Hock Dataset: Analyzing Crowds at the Stadium”, in CVPR2015.
223. Wen Wang, Ruizping Wang, Zhigui Huang, Shiguang Shan, Xinlin Chen, “Discriminant Analysis on Riemannian Manifold of Gaussian Distributions for Face Recognition With Image Sets”, in CVPR2015.
224. Georgios Georgiadis, Alessandro Chiuso, Stefano Soatto, “Texture Representations for Image and Video Synthesis”, in CVPR2015.
225. Li Shen, Teck Wee Chua, Karianto Leman, “Shadow Optimization From Structured Deep Edge Detection”, in CVPR2015.
226. Maximillian Baust, Laurent Demaret, Martin Storath, Nassir Navab, Andreas Weinmann, “Total Variation Regularization of Shape Signals”, in CVPR2015.
227. Damien Teney, Matthew Brown, Dmitry Kit, Peter Hall, “Learning Similarity Metrics for Dynamic Scene Segmentation”, in CVPR2015.
228. Baohua Li, Ying Zhang, Zhouchen Lin, Huchuan Lu, Yang Li, “Subspace Clustering by Mixture of Gaussian Regression”, in CVPR2015.
229. Seungryong Kim, Dongbo Min, Bumsub Ham, Seunghul Ryu, Minh N. Do, Kwanghoon Sohn, “DASC: Dense Adaptive Self-Correlation Descriptor for Multi-Modal and Multi-Spectral Correspondence”, in CVPR2015.
230. Horst Possegger, Thomas Mauthner, Horst Bischof, “In Defense of Color-Based Model-Free Tracking”, in CVPR2015.
231. Olga Russakovsky, Li-Jia Li, Fei-Fei Li, “Best of Both Worlds: Human-Machine Collaboration for Object Annotation”, in CVPR2015.
232. Zygmunt L. Szpak, Wojciech Chojnacki, Anton van den Hengel, “Robust Multiple Homography Estimation: An Ill-Solved Problem”, in CVPR2015.
233. Ting Yao, Yingwei Pan, Chong-Wah Ng, Houqiang Li, Tao Mei, “Semi-Supervised Domain Adaptation With Subspace Learning for Visual Recognition”, in CVPR2015.
234. Luca Del Pero, Susanna Ricco, Rahul Sukthankar, Victorio Ferrari, “Articulated Motion Discovery Using Pairs of Trajectories”, in CVPR2015.
235. Florian Bernard, Johan Thunberg, Peter Gennar, Frank Hertel, Andreas Hesch, Jorge Goncalves, “A Solution for Multi-Alignment by Transformation Synchronisation”, in CVPR2015.
236. Florian Bernard, Johan Thunberg, Peter Gennar, Frank Hertel, Andreas Hesch, Jorge Goncalves, “A Solution for Multi-Alignment by Transformation Synchronisation”, in CVPR2015.
237. Antonio Agudo, Francesc Moreno-Noguer, “Simultaneous Pose and Non-Rigid Shape With Particle Dynamics”, in CVPR2015.
238. Kwang In Kim, James Tompkin, Hanspeter Pfister, Christian Theobalt, “Semi-Supervised Learning With Explicit Relationship Regularization”, in CVPR2015.

239. Shengcai Liao, Yang Hu, Xiangyu Zhu, Stan Z. Li, “Person Re-Identification by Local Maximal Occurrence Representation and Metric Learning”, in CVPR2015.

240. Kaihil Zhao, Wen-Sheng Chu, Fernando De la Torre, Jefrey F. Cohn, Honggang Zhang, “Joint Patch and Multi-Label Learning for Facial Action Unit Detection”, in CVPR2015.

241. Chao Liu, Hernando Gomez, Srinivasa Narasimhan, Arthur Dubrawski, Michael R. Finksky, Brian Zisserbraun, “Real-Time Visual Analysis of Microvascular Blood Flow for Critical Care”, in CVPR2015.

242. Longqin Wen, Dawei Du, Zhen Lei, Stan Z. Li, Ming-Hsuan Yang, “JOTS: Joint Online Tracking and Segmentation”, in CVPR2015.

243. Jia Xu, Lopamudra Mukherjee, Yin Li, Jamieson Warner, James M. Rehg, Vikas Singh, “Gaze-Enabled Ego-centric Video Summarization via Constrained Submodular Maximization”, in CVPR2015.

244. Jianjun Lu, David Forsyth, “Sparse Depth Super Resolution”, in CVPR2015.

245. Kai-Fu Yang, Shao-Bing Gao, Yong-Jie Li, “Efficient Illuminant Estimation for Color Constancy Using Grey Pixels”, in CVPR2015.

246. Chenxiang Xu, Shao-Hang Hsieh, Caiming Xiong, Jason J. Corso, “Can Humans Fly? Action Understanding With Multiple Classes of Actors”, in CVPR2015.

247. Lei Zhang, Wei Wei, Yanning Zhang, Chunna Tian, Fei-Cheng Wang, “ Efficient Multi-Modal Fusion With Binary Features”, in CVPR2015.

248. Cenek Albl, Zuzana Kukelova, Tomas Pajdla, “R6P - Rolling Shutter Absolute Camera Pose”, in CVPR2015.

249. Daniel Moreno, Kilho Son, Fernando De la Torre, Jie Yang, “Online Sketching Hashing”, in CVPR2015.

250. Jason Rock, Tanmay Gupta, Justin Thorsen, JunYoung Gwak, Daeyun Shin, Derek Hoiem, “Completing 3D Object Shape From One Depth Image”, in CVPR2015.

251. Thomas Mauthner, Horst Possegger, Georg Waltner, Horst Bischof, “Encoding Based Saliency Detection for Videos and Images”, in CVPR2015.

252. Cong Leng, Jiaxiang Wu, Jian Cheng, Xiao Bai, Hanqing Lu, “Online Sketching Hashing”, in CVPR2015.

253. Christopher Bongsoo Choy, Michael Stark, Sam Corbett-Davies, Silvio Savarese, “Enriching Object Detection With 2D-3D Registration and Continuous Viewpoint Estimation”, in CVPR2015.

254. Naoufel Werghi, Claudio Tortorici, Stefano Berretti, Alberto Del Bimbo, “Representing 3D Texture on Mesh Manifolds for Retrieval and Recognition Applications”, in CVPR2015.

255. Wanli Ouyang, Xiaogang Wang, Xingyu Zeng, Shih Qin, Ping Luo, Yonglong Tian, Hongsheng Li, Shaohyang, Zhe Wang, Chen-Change Loy, Xiaou Tang, “DeepID-Net: Deformable Deep Convolutional Neural Networks for Object Detection”, in CVPR2015.

256. Tat-Jun Chin, Pulak Purskait, Anders Eriksson, David Suter, “Efficient Globally Optimal Consensus Maximisation With Tree Search”, in CVPR2015.

257. Xinlei Chen, C. Lawrence Zitnick, “Mind’s Eye: A Recurrent Visual Representation for Image Caption Generation”, in CVPR2015.

258. Raghuraman Gopalan, “Hierarchical Sparse Coding With Geometric Prior For Visual Geo-Location”, in CVPR2015.

259. Hongchang Wu, “P3.5P: Pose Estimation With Unknown Focal Length”, in CVPR2015.

260. Tae-Hyun Koo, Daejin Dai, Duc Vi, “Joint Vanishing Point Extraction and Tracking”, in CVPR2015.

261. Changchang Wu, “P3.5P: Pose Estimation With Unknown Focal Length”, in CVPR2015.

262. Christopher Bongsoo Choy, Michael Stark, Sam Corbett-Davies, Silvio Savarese, “Enriching Object Detection With 2D-3D Registration and Continuous Viewpoint Estimation”, in CVPR2015.

263. Naoufel Werghi, Claudio Tortorici, Stefano Berretti, Alberto Del Bimbo, “Representing 3D Texture on Mesh Manifolds for Retrieval and Recognition Applications”, in CVPR2015.

264. Cong Leng, Jiaxiang Wu, Jian Cheng, Xiao Bai, Hanqing Lu, “Online Sketching Hashing”, in CVPR2015.

265. Thomas Mauthner, Horst Possegger, Georg Waltner, Horst Bischof, “Encoding Based Saliency Detection for Videos and Images”, in CVPR2015.
280. Philippe Weinzaepfel, Jerome Revaud, Zaid Harchaoui, Cordelia Schmid, “Learning to Detect Motion Boundaries”, in CVPR2015.
281. Xiaozhi Chen, Huimin Ma, Xiang Wang, Zhichen Zhao, “Improving Object Proposals With Multi-Thresholding Straddling Expansion”, in CVPR2015.
282. Hossein Hajimirsadeghi, Wang Yan, Arash Vahdat, Greg Mori, “Visual Recognition by Counting Instances: A Multi-Instance Cardinality Potential Kernel”, in CVPR2015.
283. Joseph Roth, Yiyong Tong, Xiaoming Liu, “Unconstrained 3D Face Reconstruction”, in CVPR2015.
284. Edward Johns, Oisin Mac Aodha, Gabriel J. Brestow, “Becoming the Expert - Interactive Multi-Class Machine Teaching”, in CVPR2015.
285. Jeffrey Donahue, Lisa Anne Hendricks, Sergio Guadarrama, Marcus Rohrbach, Subhashini Venugopalan, Kate Saenko, Trevor Darrell, “Long-Term Recurrent Convolutional Networks for Visual Recognition and Description”, in CVPR2015.
286. Zhenyong Fu, Tao Xiang, Elyor Kodirov, Shaogang Gong, “Zero-Shot Object Recognition by Semantic Manifold Distance”, in CVPR2015.
287. Saining Xie, Tianbao Yang, Xiaoyu Wang, Yuanqing Lin, “Hyper-Class Augmented and Regularized Deep Learning for Fine-Grained Image Classification”, in CVPR2015.
288. Nianjuan Jiang, Daniel Lin, Minh N. Do, Jianguo Lu, “Direct Structure Estimation for 3D Reconstruction”, in CVPR2015.
289. Xuehan Xiong, Fernando De la Torre, “Global Supervised Descent Method”, in CVPR2015.
290. Onur Ozayesil, Amit Singer, “Robust Camera Location Estimation by Convex Programming”, in CVPR2015.
291. Johan Fredriksson, Viktor Larsson, Carl Olsson, “Practical Robust Two-View Translation Estimation”, in CVPR2015.
292. Tong Xiao, Tian Xia, Yi Yang, Chang Huang, Xiaogang Wang, “Learning From Massive Noisy Labeled Data for Image Classification”, in CVPR2015.
293. Mithun Das Gupta, Srinidhi Srinivasa, Madhukara J., Meryl Antony, “KL Divergence Based Agglomerative Clustering for Automated Vitiiligo Grading”, in CVPR2015.
294. Changyang Li, Yuchen Yuan, Weidong, Cai, Yong Xia, David Dagan Feng, “Robust Salience Detection via Regularized Random Walks Ranking”, in CVPR2015.
295. Wei Zhang, Sheng Zeng, Dequan Wang, Xiangyang Xue, “Weakly Supervised Semantic Segmentation for Social Image Annotation”, in CVPR2015.
296. Mainak Jas, Devi Parikh, “Image Specificity”, in CVPR2015.
297. Neel Shah, Vladimir Kolmogorov, Christoph H. Lampert, “A Multi-Plane Block-Coordinate Frank-Wolfe Algorithm for Training Structural SVMs With a Costly Max-Oracle”, in CVPR2015.
298. Yaniv Taigman, Ming Yang, Marc’Aurelio Ranzato, Lior Wolf, “Web-Scale Training for Face Identification”, in CVPR2015.
299. Christoph Feichtenhofer, Axel Pinz, Richard P. Wildes, “Dynamically Encoded Actions Based on Spacetime Saliency”, in CVPR2015.
300. Takami Kobayashi, “Three Viewpoints Toward Exemplar SVM”, in CVPR2015.
301. Li Niu, Wen Li, Dong Xu, “Visual Recognition by Learning From Web Data: A Weakly Supervised Domain Generalization Approach”, in CVPR2015.
302. Georg Nebehay, Roman Pflugfelder, “Clustering of Static-Adaptive Correspondences for Deformable Object Tracking”, in CVPR2015.
303. Sherwin Ardehshir, Kofi Malcolm Collins-Sibley, Mubarak Shah, “Geo-Semantic Segmentation”, in CVPR2015.
304. Peng Wang, Xiaohui Shen, Zhe Lin, Scott Cohen, Brian Price, Alan L. Yuille, “Towards Unified Depth and Semantic Prediction From a Single Image”, in CVPR2015.
305. Tu-Hoa Pham, Abderrahmane Kheddar, Ammar Qammar, Antonis A. Argyros, “Towards Force Sensing From Vision: Observing Hand-Object Interactions to Infer Manipulation Forces”, in CVPR2015.
306. Mateusz Koziński, Raghuveer Gadde, Sergey Zagoruyko, Guillaume Obozinski, Renaud Marlet, “A MRF Shape Prior for Facade Parsing With Occlusions”, in CVPR2015.
307. Timur Bagautdinov, Francois Fleuret, Pascal Fua, “Probability Occupancy Maps for Occluded Depth Images”, in CVPR2015.
308. Rabeeh Karimi Mahabadi, Christian Hane, Marc Pollefeys, “Segment Based 3D Object Shape Priors”, in CVPR2015.
309. Mathias Gallardo, Daniel Pizarro, Adrién Bartoli, Toby Collins, “Shape-From-Template in Flatland”, in CVPR2015.
310. Yixin Zhu, Yibiao Zhao, Song Chun Zhu, “Understanding Tools: Task-Oriented Object Modeling, Learning and Recognition”, in CVPR2015.
311. Edouard Oyallon, Stephane Mallat, “Deep Roto-Translation Scattering for Object Classification”, in CVPR2015.
312. Hong-Ren Su, Shang-Hong Lai, “Non-Rigid Registration of Images With Geometric and Photometric Deformation by Using Local Affine Fourier-Moment Matching”, in CVPR2015.
313. Judy Hoffman, Deepak Pathak, Trevor Darrell, Kate Saenko, “Detector Discovery in the Wild: Joint Multiple Instance and Representation Learning”, in CVPR2015.
314. Yi Sun, Xiaogang Wang, Xiaoou Tang, “Deeply Learned Face Representations Are Sparse, Selective, and Robust”, in CVPR2015.
315. Fatemeh Shokrollahi Yancheshmeh, Ke Chen, Joni-Kristian Kamarainen, “Unsupervised Visual Alignment With Similarity Graphs”, in CVPR2015.
316. Kai-Wen Cheng, Yie-Tarng Chen, “Video Anomaly Detection and Localization Using Hierarchical Feature Representation and Gaussian Process Regression”, in CVPR2015.
317. Yiyan Pan, Martial Hebert, Takeo Kanade, “Inferring 3D Layout of Building Facades From a Single Image”, in CVPR2015.
318. Zeynep Akata, Scott Reed, Daniel Walter, Honglak Lee, Bernt Schiele, “Evaluation of Output Embeddings for Fine-Grained Image Classification”, in CVPR2015.
319. Joao Carreira, Abhishek Mall, Shubham Tulsiani, Jitendra Malik, “Virtual View Networks for Object Reconstruction”, in CVPR2015.
320. Jian Yao, Marko Boben, Sanja Fidler, Raquel Urtasun, “Real-Time Coarse-to-Fine Topologically Preserving Segmentation”, in CVPR2015.
321. Albert Gordo, “Supervised Mid-Level Features for Word Image Representation”, in CVPR2015.
322. Takuya Naribari, Michael Maire, Stella X. Yu, “Learning Lightness From Human Judgement on Relative Reflectance”, in CVPR2015.
323. Mandar Dixit, Si Chen, Dashan Gao, Nikhil Riasisaria, Nuno Vasconcelos, “Scene Classification With Semantic Fisher Vectors”, in CVPR2015.
324. Xiao Lin, Devi Parikh, “Don’t Just Listen, Use Your Imagination: Leveraging Visual Common Sense for Non-Visual Tasks”, in CVPR2015.
325. Dingwen Zhang, Junwei Han, Chao Li, Jingdong Wang, “Co-Saliency Detection via Looking Deep and Wide”, in CVPR2015.
326. Filippo Bergamasco, Andrea Albareli, Luca Cosmo, Andrea Torsello, Emanuele Rodola, Daniel Cremers, “Adopting an Unconstrained Ray Model in Light-Field Cameras for 3D Shape Reconstruction”, in CVPR2015.
327. Wei Liu, Rongrong Ji, Shaozi Li, “Towards 3D Object Detection With Bimodal Deep Boltzmann Machines Over RGBD Imagery”, in CVPR2015.
328. Abel Gonzalez-Garcia, Alexander Vezhnevets, Vittorio Ferrari, “An Active Search Strategy for Efficient Object Class Detection”, in CVPR2015.
329. Aasa Feragen, Francois Lauze, Soren Hauberg, “Geodesic Exponential Kernels: When Curvature and Linearity Conflict”, in CVPR2015.
330. Dmitry Laptev, Joachim M. Buhmann, “Transformation-Invariant Convolutional Jungles”, in CVPR2015.
331. Joaquin Zepeda, Patrick Perez, “Exemplar SVMs as Visual Feature Encoders”, in CVPR2015.
332. Moritz Menze, Andreas Geiger, “Object Scene Flow for Autonomous Vehicles”, in CVPR2015.
333. Hang Zhang, Kristin Dana, Ko Nishino, “Reflectance Hasing for Material Recognition”, in CVPR2015.
334. Gunhee Kim, Seungwhan Moon, Leonid Sigal, “Joint Photo Stream and Blog Post Summarization and Exploration”, in CVPR2015.
335. Michael Gygli, Helmut Grabner, Luc Van Gool, “Video Summarization by Learning Submodular Mixtures of Objectives”, in CVPR2015.
336. Marcus A. Brubaker, Ali Punjani, David J. Fleet, “Building Proteins in a Day: Efficient 3D Molecular Reconstruction”, in CVPR2015.
337. Paul Wohlhart, Vincent Lepetit, “Co-Saliency Detection via Looking Deep and Wide”, in CVPR2015.
338. Anna Rohrbach, Marcus Rohrbach, Niket Tandon, Bernt Schiele, “A Dataset for Movie Description”, in CVPR2015.
339. Shinath Sridhar, Franziska Mueller, Antti Oulasvirta, Christian Theobalt, “Fast and Robust Hand Tracking Using Detection-Guided Optimization”, in CVPR2015.
340. Peng Wang, Chunhua Shen, Anton van den Hengel, “Efficient SDP Inference for Fully-Connected CRFs Based on Low-Rank Decomposition”, in CVPR2015.
341. Rushil Anirudh, Pavan Turaga, Jingyong Su, Anuj Srivastava, “Computing Similarity Transformations From Only Image Correspondences”, in CVPR2015.
342. Xufeng Han, Thomas Leung, Yangqing Jia, Rahul Sukthankar, Alexander C. Berg, “MatchNet: Unifying Feature and Metric Learning for Patch-Based Matching”, in CVPR2015.
343. Karihaley Shanmuga Vadivel, Thuyen Ngo, Miguel Eckstein, B.S. Manjunath, “Eye Tracking Assisted Extraction of Attentionally Important Objects From Videos”, in CVPR2015.
344. YiChang Shih, Dilip Krishnan, Fredo Durand, William T. Freeman, “Reflection Removal Using Ghosting Cues”, in CVPR2015.
345. YiChang Shih, Dilip Krishnan, Fredo Durand, William T. Freeman, “Reflection Removal Using Ghosting Cues”, in CVPR2015.
413. Hironori Hattori, Vishnu Naresh Boddeti, Kris M. Kitani, Takeo Kanade, “Learning Scene-Specific Pedestrian Detectors Without Real Data”, in CVPR2015.

414. Mireia Cimpoli, Subhransu Maji, Andrea Vedaldi, “Deep Filter Banks for Texture Recognition and Segmentation”, in CVPR2015.

415. Chulwoo Lee, Won-Dong Jang, Jae-Young Sim, Chang-Su Kim, “Multiple Random Walkers and Their Application to Image Cosegmentation”, in CVPR2015.

416. Rui Caseiro, Joao F. Henriques, Pedro Martins, Jorge Batista, “Beyond the Shortest Path : Unsupervised Domain Adaptation by Sampling Subspaces Along the Spline Flow”, in CVPR2015.

417. Etao Littwin, Hadar Averbuch-Elor, Daniel Cohen-Or, “Spherical Embedding of Inlier Silhouette Dissimilarities”, in CVPR2015.

418. Zijia Lin, Guiguang Ding, Mingqing Hu, Jianmin Wang, “Semantics-Preserving Hashing for Cross-View Retrieval”, in CVPR2015.

419. Chaoyang Wang, Long Zhao, Shuang Liang, Linqing Zhang, Jinyuan Jia, Yichen Wei, “Object Proposal by Multi-Branch Hierarchical Segmentation”, in CVPR2015.

420. Wei Yang, Yu Ji, Haiting Lin, Yang Yang, Sing Bing Kang, Jiayi Yu, “Ambient Occlusion via Compressive Visibility Estimation”, in CVPR2015.

421. Naemullah Khan, Marei Algarni, Anthony Yezzi, Ganesh Sundaramoorthy, “Shape-Tailored Local Descriptors and Their Application to Segmentation and Tracking”, in CVPR2015.

422. Ting-Hsuan Chao, Yen-Liang Lin, Yin-Hsi Kuo, Winston H. Hsu, “Scalable Object Detection by Filter Compressions With Regularized Sparse Coding”, in CVPR2015.

423. Ejaz Ahmed, Michael Jones, Tim K. Marks, “An Improved Deep Learning Architecture for Person Re-Identification”, in CVPR2015.

424. Mayank Kabra, Alice Robie, Kristin Branson, “Understanding Classifier Errors by Examining Influenal Neighbors”, in CVPR2015.

425. Mehrtash Harandi, Mathieu Salzmann, “Riemannian Coding and Dictionary Learning: Kernels to the Rescue”, in CVPR2015.

426. Benjamin Resch, Hendrik A. A. Lensch, Oliver Wang, Marc Pollefeys, Alexander Sorkine-Hornung, “Scalable Structure From Motion for Densely Sampled Videos”, in CVPR2015.

427. Xianjie Chen, Alan L. Yuille, “Parsing Occluded People by Flexible Compositions”, in CVPR2015.

428. Davide Modolo, Alexander Veeranees, Olga Ruscakovsky, Vittorio Ferrari, “Joint Calibration of Ensemble of Exemplar SVMs”, in CVPR2015.

429. Shenlong Wang, Sanja Fidler, P. A. Lensch, Oliver Wang, Marc Pollefeys, Alexander Sorkine-Hornung, “Scalable Structure From Motion for Densely Sampled Videos”, in CVPR2015.

430. Xianjie Chen, Alan L. Yuille, “Parsing Occluded People by Flexible Compositions”, in CVPR2015.

431. Wei Shen, Xinggang Wang, Yan Wang, Xiang Bai, Zhiqiang Zhang, “DeepContour: A Deep Convolutional Feature Learned by Positive-Sharing Loss for Contour Detection”, in CVPR2015.

432. Jifeng Dai, Kaiming He, Jian Sun, “Convolutional Feature Masking for Joint Object and Stuff Segmentation”, in CVPR2015.

433. Kai Han, Kwan-Yee K. Wong, Miaomiao Liu, “A Fixed Viewpoint Approach for Dense Reconstruction of Transparent Objects”, in CVPR2015.
476. Julien Valentin, Matthias Niesner, Jamie Shotton, An-
474. Gedas Bertasius, Jianbo Shi, Lorenzo Torresani,
472. Chenxia Wu, Jiemi Zhang, Silvio Savarese, Ashutosh
470. Christoph Kading, Alexander Freytag, Erik Rodner,
467. Mrigank Rochan, Yang Wang, “Weakly Supervised Lo-
464. Buyu Liu, Xuming He, “Multiclass Semantic Video
468. Gregory Rogez, James S. Supan?i? III, Deva Ra-
466. Limin Wang, Yu Qiao, Xiaoou Tang, “Action Recogni-
465. Tal Hassner, Shai Harel, Eran Paz, Roee Enbar, “Ef-
461. Yu-Wei Chao, Zhan Wang, Rada Mihalcea, Jia Deng,
459. Artem Babenko, Victor Lempitsky, “Tree Quantization
458. Christian Simon, In Kyu Park, “Reflection Removal for
457. Gunay Do?an, Javier Bernal, Charles R. Hagwood, “A
455. Donghoon Lee, Hyunsin Park, Chang D. Yoo, “Face
454. Jose C. Rubio, Bjorn Ommer, “Regularizing Max-
452. Brian Taylor, Vasiliy Karasev, Stefano Soatto, “Causa
450. Yu-Sung Xie, Ting-Yi Li, Yi-Chuan Chiang, “First-
448. Xiaoyang Wang, Qiang Ji, “Video Event Recognition
447. Yang Song, Weidong Cai, Qing Li, Fan Zhang, David
509. Yukun Zhu, Raquel Urtasun, Ruslan Salakhutdinov, Heng Yang, Ioannis Patras, “Mirror, Mirror on the Wall,” in CVPR2015.

505. Daniil Kononenko, Victor Lempitsky, “Learning To Tell Me, Is the Error Small?”, in CVPR2015.

500. Chen Li, Kun Zhou, Stephen Lin, “Simulating Makeup Ground Modeling”, in CVPR2015.

499. Shijie Xiao, Wen Li, Dong Xu, Dacheng Tao, “FaLRR: A Fast Low Rank Representation Solver”, in CVPR2015.

498. Philip Saponaro, Scott Sorensen, Abhishek Kolagunda, Chavdar Papazov, Tim K. Marks, Michael Jones, “Real-time 3D Head Pose and Facial Landmark Estimation From Depth Images Using Triangular Surface Patch Features”, in CVPR2015.

497. Heng Yang, Ioannis Patras, “Mirror, Mirror on the Wall, Tell Me, Is the Error Small?”, in CVPR2015.

496. Bo Xin, Yuan Tian, Yizhou Wang, Wen Gao, “Back-ground Subtraction via Generalized Fused Lasso Foreground Modeling”, in CVPR2015.

495. Ying Li, Ruiping Wang, Zhiwu Huang, Shiguang Shan, “A Stable Multi-Scale Kernel for Topological Machine learning”, in CVPR2015.

494. Lingqiao Liu, Chunhua Shen, Anton van den Hengel, “The Treasure Beneath Convolutional Layers: Cross-Convolutional-Layer Pooling for Image Classification”, in CVPR2015.

493. Jan Reininghaus, Stefan Huber, Ulrich Bauer, Roland Kwitt, “A Stable Multi-Scale Kernel for Topological Machine learning”, in CVPR2015.

492. Syed Zulqarnain Gilani, Faisal Shafait, Ajmal Mian, Ke Jiang, Qichao Que, Brian Kulis, “Revisiting Kernelized Locality-Sensitive Hashing for Improved Large-Scale Image Retrieval”, in CVPR2015.

491. Tsung-Yi Lin, Yin Cui, Serge Belongie, James Hays, “3D Reconstruction in the Presence of Glasses by Acoustic and Stereo Fusion”, in CVPR2015.

490. Ziheng Wang, Qiang Ji, “Classifier Learning With Hidden Information”, in CVPR2015.

489. Ali Mosleh, Paul Green, Emmanuel Onzon, Isabelle Béggin, J.M. Pierre Langlois, “Camera Intrinsic Blur Kernel Estimation: A Reliable Framework”, in CVPR2015.

488. Zheng Wang, Qiang Ji, “Classifier Learning With Hidden Information”, in CVPR2015.
543. Dongyoon Han, Junno Kim, “Unsupervised Simultaneous Orthogonal Basis Clustering Feature Selection”, in CVPR2015.
544. Shugao Ma, Leonid Sigal, Stan Sclaroff, “Space-Time Tree Ensemble for Action Recognition”, in CVPR2015.
545. Siyu Tang, Bjorn Andres, Miikhyayo Andriluka, Bernt Schiele, “Subgraph Decomposition for Multi-Target Tracking”, in CVPR2015.
546. Xian-Ming Liu, Rongrong Ji, Changhu Wang, Wei Liu, Siyu Tang, Bjoern Andres, Miykhaylo Andriluka, Bernt Schiele, “Subgraph Decomposition for Multi-Target Tracking”, in CVPR2015.
547. Yanchao Yang, Zhaojin Lu, Ganesh Sundaramoorthi, “Coarse-To-Fine Region Selection and Matching”, in CVPR2015.
548. Yan Luo, Yongkang Wang, Qi Zhao, “Label Consistent Quadratic Surrogate Model for Visual Saliency Prediction”, in CVPR2015.
549. Yumin Suh, Kamil Adamczewski, Kyoung Mu Lee, “Subgraph Matching Using Compactness Prior for Robust Feature Correspondence”, in CVPR2015.
550. Yongjoung Ttan, Ping Loo, Xiaogang Wang, Xiaoo Tang, “Pedestrian Detection Aided by Deep Learning Semantic Tasks”, in CVPR2015.
551. Dae-Youn Lee, Jae-Young Sim, Chang-Su Kim, “Multi-hypothesis Trajectory Analysis for Robust Visual Tracking”, in CVPR2015.
552. Jingming Dong, Stefano Soatto, “Domain-Size Pooling in Local Descriptors: DSP-SIFT”, in CVPR2015.
553. Junjie Yan, Yinan Yu, Xiangyu Zhu, Zhen Lei, Stan Z. Li, “Object Detection by Labeling Superpixels”, in CVPR2015.
554. Ching Teo, Cornelia Fermuller, Yiannis Aloimonos, “Fast 2D Border Ownership Assignment”, in CVPR2015.
555. Johannes L. Schonberger, Filip Radenovi?, Ondrej Chum, Jan-Michael Frahm, “From Single Image Query to Detailed 3D Reconstruction”, in CVPR2015.
556. Felix Heide, Wolfgang Heidrich, Gordon Wetzstein, “Fast and Flexible Convolutional Sparse Coding”, in CVPR2015.
557. Thalaiyasingam Ajanthan, Richard Hartley, Mathieu Salzmann, Hongdong Li, “Iteratively Reweighted Graph Cut for Multi-Label MRFs With Non-Convex Priors”, in CVPR2015.
558. Xishuo Li, Martha Larson, Alan Hanjalic, “Pairwise Geometric Matching for Large-Scale Object Retrieval”, in CVPR2015.
559. Fayao Liu, Chunhua Shen, Guosheng Lin, “Deep Convolutional Neural Fields for Depth Estimation From a Single Image”, in CVPR2015.
560. Xianning Liu, Xiaolin Wu, Jiantao Zhou, Debin Zhao, “Data-Driven Sparsity-Based Restoration of JPEG-Compressed Images in Dual Transform-Pixel Domain”, in CVPR2015.
561. Yale Song, Jordi Vallmitjana, Amanda Stent, Alejandro Jaimes, “TVSum: Summarizing Web Videos Using Titles”, in CVPR2015.
562. Aravindh Mahendran, Andrea Vedaldi, “Understanding Deep Image Representations by Inverting Them”, in CVPR2015.
563. Jia-Bin Huang, Abhishek Singh, Narendra Ahuja, “Single Image Super-Resolution From Transformed Self-Exemplars”, in CVPR2015.
564. Markus Schoeler, Jeremie Papon, Florentin Worgotter, “Constrained Planar Cuts - Object Partitioning for Point Clouds”, in CVPR2015.
565. Nianyi Li, Bilin Sun, Jingyi Yu, “A Weighted Sparse Coding Framework for Saliency Detection”, in CVPR2015.
566. Ziyang Ma, Renjie Liao, Xin Tao, Li Xu, Jiaya Jia, Enhua Wu, “Handling Motion Blur in Multi-Frame Super-Resolution”, in CVPR2015.
567. Nir Ben-Zrihem, Lihi Zelnik-Manor, “Approximate Nearest Neighbor Fields in Video”, in CVPR2015.
568. Roee Litman, Simon Korman, Alexander Bronstein, Shai Avidan, “Inverting RANSAC: Global Model Detection via Inlier Rate Estimation”, in CVPR2015.
569. Yonggang Jin, Jianzhuang Liu, XueLong Li, “A Maximum Entropy Feature Descriptor for Age Invariant Face Recognition”, in CVPR2015.
570. Xunli Chen, Alan Ritter, Abhinav Gupta, Tom Mitchell, “Sense Discovery via Co-Clustering on Images and Text”, in CVPR2015.
571. Zicheng Liao, Kevin Karsch, David Forsyth, “An Approximate Shading Model for Object Relighting”, in CVPR2015.
572. Xing Chen, Junshi Huang, Rogerio Feris, Lisa M. Brown, Jian Dong, Shuicheng Yan, “Deep Domain Adaptation for Describing People Based on Fine-Grained Clothing Attributes”, in CVPR2015.
573. Haoxiang Li, Zhe Lin, Xiaohui Shen, Jonathan Brandt, Gang Hua, “A Convolutional Neural Network Cascade for Face Detection”, in CVPR2015.
574. Abe Davis, Katherine L. Bouman, Justin G. Chen, Michael Rubinstein, Fredo Durand, William T. Freeman, “Visual Vibrometry: Estimating Material Properties From Small Motion in Video”, in CVPR2015.
575. Tian-Fang Hu, Wei-Shi Zheng, Jianhuang Lai, Jianguo Zhang, “Jointly Learning Heterogeneous Features for RGBD Activity Recognition”, in CVPR2015.
576. Quynh Nguyen, Antoine Gautier, Matthias Hein, “TILDE: A Temporally Invariant Learned DETector”, in CVPR2015.
577. Xiaohan Zhang, Hai Su, Lin Yang, Shuangting Zhang, “Fine-Grained Histopathological Image Analysis via Robust Segmentation and Large-Scale Retrieval”, in CVPR2015.
578. Xiaofan Zhang, Hai Su, Lin Yang, Shuangting Zhang, “Fine-Grained Histopathological Image Analysis via Robust Segmentation and Large-Scale Retrieval”, in CVPR2015.
588. Tae Hyun Kim, Kyoung Mu Lee, “Generalized Video Deblurring for Dynamic Scenes”, in CVPR2015.
589. Epameinondas Antonakos, Joan Alabort-i-Medina, Stefanos Zafeiriou, “Active Pictorial Structures”, in CVPR2015.
590. Ryo Yonetani, Kris M. Kitani, Yoichi Sato, “Ego-Surfing First-Person Videos”, in CVPR2015.
591. Guanbin Li, Yizhou Yu, “Visual Saliency Based on Multiscale Deep Features”, in CVPR2015.
592. Kenichiro Tanaka, Yasuhiro Mukaigawa, Hiroiuki Kubo, Yasuyuki Matsushita, Yasushi Yagi, “Recovering Inner Slices of Translucent Objects by Multi-Frequency Illumination”, in CVPR2015.
593. Kwang In Kim, James Tompkin, Hanspeter Pfister, Christian Theobalt, “Local High-Order Regularization on Data Manifolds”, in CVPR2015.
594. David Hall, Pietro Perona, “Fine-Grained Classification of Pedestrians in Video: Benchmark and State of the Art”, in CVPR2015.
595. Anastasia Pentina, Viktoriia Sharmanska, Christoph H. Lampert, “Curriculum Learning of Multiple Tasks”, in CVPR2015.
596. Sayed Hossein Khatoonabadi, Nuno Vasconcelos, Ivan V. Bajic, Yufeng Shan, “How Many Bits Does it Take for a Stimulus to Be Salient?”, in CVPR2015.
597. Nikolay Savinov, Lubar Ladicky, Christian Hane, Marc Pollefeys, “Discrete Optimization of Ray Potentials for Semantic 3D Reconstruction”, in CVPR2015.
598. Chenglong Li, Liang Lin, Wangmeng Zuo, Shuicheng Yan, Jin Tang, “SOLD: Sub-Optimal Low-rank Decomposition for Efficient Video Segmentation”, in CVPR2015.
599. Ioannis Gkioulekas, Bruce Walter, Edward H. Adelson, Kavita Bala, Todd Zickler, “On the Appearance of Translucent Edges”, in CVPR2015.
600. Vissesh Chari, Simon Lacoste-Julien, Ivan Laptev, Josef Sivic, “On Pairwise Costs for Network Flow Multi-Object Tracking”, in CVPR2015.
601. Jonathan Krause, Hailin Jin, Jianchao Yang, Li Fei-Fei, “Fine-Grained Recognition Without Part Annotations”, in CVPR2015.
602. Sungjoon Choi, Qian-Yi Zhou, Vladlen Koltun, “Robust Reconstruction of Indoor Scenes”, in CVPR2015.
603. http://hci.tokyo/seminar/chi2015
604. Matthew D. Zeiler, Rob Fergus, “Visualizing and Understanding Convolutional Networks”, in ECCV2014.
605. https://sites.google.com/site/cvsaisentan/
606. https://twitter.com/nagoyacv
607. https://twitter.com/kansaicvprml
608. Wordle, http://www.wordle.net/
609. R. Girshick, J. Donahue, T. Darrell, J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation”, in CVPR2014.
610. R. A. Newcombe, S. Izadi, O. Hilliges, D. Molyneaux, D. Kim, A. J. Davison, P. Kohli, J. Shotton, S. Hodges, A. Fitzgibbon, “KinectFusion: Real-Time Dense Surface Mapping and Tracking”, in IEEE ISMAR, 2011.
611. A. Gupta, A. A. Efros, M. Hebert, “Blocks World Revisited: Image Understanding Using Qualitative Geometry and Mechanics”, European Conference on Computer Vision (ECCV), 2010.
612. J. Shotton, M. Johnson, R. Cipolla, “Semantic Texture Forests for Image Categorization and Segmentation”, in CVPR2008.
613. J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei, “ImageNet: A Large-Scale Hierarchical Image Database”, in CVPR2009.
614. A. Krizhevsky, I. Sutskever, G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks”, in NIPS2012.
615. J. Xiao, K. A. Ehinger, J. Hays, A. Torralba, A. Oliva, “SUN Database: Exploring a Large Collection of Scene Categories”, International Journal of Computer Vision (IJCV), 2014.
616. P. Felzenszwalb, R. Girshick, D. McAllester, D. Ramanan, “Object Detection with Discriminatively Trained Part Based Models”, IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), Vol. 32, No. 9, 2010.
617. N. Dalal, B. Trigg, “Histograms of oriented gradients for human detection”, in CVPR2005.
618. C. Vondrick, A. Khosla, T. Malisiewicz, A. Torralba, “HOGGles: Visualizing Object Detection Features”, International Conference on Computer Vision (ICCV), 2013.
619. Y. Taigman, M. Yang, M.A. Ranzato and L. Wolf, DeepFace: Closing the Gap to Human-Level Performance in Face Verification, CVPR2014.
620. H. Wang, A. Kraer, C. Schmid, C.-L. Liu, “Action recognition by dense trajectories”, in CVPR2011.
621. H. Wang, C. Schmid, “Action recognition with improved trajectories”, in ICCV2013.
622. K. Simonyan, A. Zisserman, “Two-Stream Convolutional Networks for Action Recognition in Videos”, in NIPS2014.
623. Y. Sun, Y. Chen, X. Wang, and X. Tang, “Deep Learning Face Representation by Joint Identification-Verification”, in NIPS2014.
624. P. Dollar, C. Wojek, B. Schiele and P. Perona, “Pedestrian Detection: An Evaluation of the State of the Art”, IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2012.
625. A. Geiger, P. Lenz, R. Urtasun, “Are we ready for Autonomous Driving? The KITTI Vision Benchmark Suite”, in CVPR2012.
626. R. Agrawal, T. Imieliński, A. Swami, “Mining association rules between sets of items in large databases”, in ACM SIGMOD1993.
627. L. Itti, C. Koch, “A saliency-based search mechanism for overt and covert shifts of visual attention”, in Vision Research, 40(10-12), pp.1489–1506, 2000.
628. T.-Y. Liu, “Learning to Rank for Information Retrieval”, Springer, 2011.
629. D. Parikh, K. Grauman, “Relative Attribute”, in ICCV2011.
630. Fine-grained Competition 2013, https://sites.google.com/site/fgcomp2013/results
631. Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, “Gradient-Based Learning Applied to Document Recognition”, Proceedings of the IEEE, 86(11):2278-2324, 1998.
632. Simonany, A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition”, ICLR, 2015.
633. K. He, X. Zhang, S. Ren, J. Sun, “Deep Residual Learning for Image Recognition”, in CVPR, 2016.
634. Hirokatsu Kataoka, Soma Shirakabe, Yudai Miyashita, Akio Nakamura, Kenji Iwata, Yutaka Satoh, “Semantic Change Detection with Hypermaps”, in arXiv pre-print 1604.07513, 2016.