

1. Record Nr.	UNINA9910629282503321
Titolo	Computer vision - ECCV 2022 . Part X : 17th European Conference, Tel Aviv, Israel, October 23-27, 2022 : proceedings // Shai Avidan [and four others]
Pubbl/distr/stampa	Cham, Switzerland : , : Springer, , [2022] ©2022
ISBN	3-031-20080-2
Descrizione fisica	1 online resource (815 pages)
Collana	Lecture Notes in Computer Science
Disciplina	006.37
Soggetti	Computer vision Pattern recognition systems
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di contenuto	Intro -- Foreword -- Preface -- Organization -- Contents - Part X -- DFNet: Enhance Absolute Pose Regression with Direct Feature Matching -- 1 Introduction -- 2 Related Work -- 3 Method -- 3.1 Direct Feature Matching for Pose Estimation -- 3.2 Histogram-Assisted NeRF -- 3.3 Random View Synthesis -- 4 Experiments -- 4.1 Implementation -- 4.2 Evaluation on the 7-Scenes Dataset -- 4.3 Evaluation on Cambridge Dataset -- 4.4 Comparison to Sequential APR and 3D Approaches -- 4.5 Ablation Study -- 5 Conclusion -- References -- Cornerformer: Purifying Instances for Corner-Based Detectors -- 1 Introduction -- 2 Related Work -- 2.1 Anchor-Based Detector -- 2.2 Anchor-Free Detector -- 3 Preliminary: The CornerNet Baseline -- 4 Cornerformer -- 4.1 Towards Better Instance Construction -- 4.2 Corner Transformer Encoder -- 4.3 Attenuation-Auto-Adjusted NMS -- 5 Experiments -- 5.1 Datasets, Metrics, and Implementation Details -- 5.2 Object Detection Results -- 5.3 Pose Estimation Results -- 5.4 Ablation Study -- 6 Conclusion -- References -- PillarNet: Real-Time and High-Performance Pillar-Based 3D Object Detection -- 1 Introduction -- 2 Related Works -- 2.1 Point Cloud 3D Object Detection -- 2.2 Multi-Sensor Based 3D Object Detection -- 3 PillarNet for 3D Object Detection -- 3.1 Preliminaries -- 3.2 PillarNet Design for 3D Object Detection -- 3.3 Orientation-Decoupled IoU Regression Loss -- 3.4

Overall Loss Function -- 4 Experiments -- 4.1 Overall Results -- 4.2 Ablation Studies -- 5 Conclusions -- References -- Robust Object Detection with Inaccurate Bounding Boxes -- 1 Introduction -- 2 Related Work -- 3 Object-Aware Multiple Instance Learning -- 3.1 Preliminaries -- 3.2 Object-Aware MIL Formulation -- 3.3 Deployment to Off-the-Shelf Object Detectors -- 4 Results and Discussions -- 4.1 Datasets and Evaluation Metrics -- 4.2 Comparison with State of the Art. 4.3 Ablation Study -- 5 Conclusion -- References -- Efficient Decoder-Free Object Detection with Transformers -- 1 Introduction -- 2 Related Work -- 2.1 One-Stage and Two-Stage Detectors -- 2.2 End-to-End Detectors -- 3 Method -- 3.1 Detection-Oriented Transformer Backbone -- 3.2 Encoder-Only Single-Level Dense Prediction -- 3.3 Miscellaneous -- 4 Experiments -- 4.1 Settings -- 4.2 Main Results -- 4.3 Ablation Study -- 5 Conclusion -- References -- Cross-Modality Knowledge Distillation Network for Monocular 3D Object Detection -- 1 Introduction -- 2 Related Works -- 3 Method -- 3.1 Framework Overview -- 3.2 BEV Feature Learning -- 3.3 Domain Adaptation via Self-calibration -- 3.4 Feature-Based Knowledge Distillation -- 3.5 Response-Based Knowledge Distillation -- 3.6 Loss Function -- 3.7 Extension: Distilling Unlabeled Data -- 4 Experiments -- 4.1 Datasets -- 4.2 Experiment Settings -- 4.3 Results on KITTI Test Set -- 4.4 Results on Waymo Open Dataset -- 4.5 Ablation Studies -- 5 Conclusion -- References -- ReAct: Temporal Action Detection with Relational Queries -- 1 Introduction -- 2 Related Work -- 3 Method -- 3.1 Relational Attention with IoU Decay -- 3.2 Action Classification Enhancement -- 3.3 Segment Quality Prediction -- 3.4 Training Losses -- 4 Experiment -- 4.1 Implementation Details -- 4.2 Main Results -- 4.3 Ablation Study -- 5 Conclusion -- References -- Towards Accurate Active Camera Localization -- 1 Introduction -- 2 Related Work -- 3 Approach -- 3.1 Task Setup -- 3.2 Passive Localization Module -- 3.3 Active Localization Module -- 4 Experiments -- 4.1 Experimental Setup -- 4.2 Compared Approaches -- 4.3 Results -- 5 Conclusion -- References -- Camera Pose Auto-encoders for Improving Pose Regression -- 1 Introduction -- 2 Related Work -- 2.1 Structure-Based Pose Estimation -- 2.2 Regression-Based Pose Estimation. 3 Absolute Pose Regression Using Pose Auto-encoders -- 3.1 Training Camera Pose Auto-encoders -- 3.2 Network Architecture -- 3.3 Applications of Camera Pose Auto-encoders -- 3.4 Implementation Details -- 4 Experimental Results -- 4.1 Experimental Setup -- 4.2 Evaluation of Camera Pose Auto-encoders (PAEs) -- 4.3 Ablation Study -- 4.4 Refining Position Estimation with Encoded Poses -- 4.5 Limitations and Future Research -- 5 Conclusions -- References -- Improving the Intra-class Long-Tail in 3D Detection via Rare Example Mining -- 1 Introduction -- 2 Related Work -- 3 Methods -- 3.1 Rare Example Mining -- 3.2 Track-Level REM for Active Learning -- 4 Experiments -- 4.1 Rare Example Mining Analysis -- 4.2 Rare Example Mining for Active Learning -- 5 Ablation Studies -- 6 Discussions and Future Work -- References -- Bagging Regional Classification Activation Maps for Weakly Supervised Object Localization -- 1 Introduction -- 2 Related Work -- 3 Methodology -- 3.1 Problem Definition -- 3.2 Regional Localizers Generation Strategy -- 3.3 Bagging Regional Classification Activation Maps -- 4 Experiments -- 4.1 Settings -- 4.2 Comparison with State-of-the-Arts -- 4.3 Discussion -- 4.4 Conclusion -- References -- UC-OWOD: Unknown-Classified Open World Object Detection -- 1 Introduction -- 2 Related Work -- 3 Unknown-Classified Open World Object Detection -- 3.1 Problem Formulation -- 3.2 Overall Architecture -- 3.3 Detection of

Unknown Objects -- 3.4 Similarity-Based Unknown Classification -- 3.5 Unknown Clustering Refinement -- 3.6 Training and Refining -- 4 Experiments -- 4.1 Preparation -- 4.2 Results and Analysis -- 4.3 Ablation Study -- 5 Conclusions and Future Work -- References -- RayTran: 3D Pose Estimation and Shape Reconstruction of Multiple Objects from Videos with Ray-Traced Transformers -- 1 Introduction -- 2 Related Work -- 3 Proposed Approach. 3.1 The RayTran Backbone -- 3.2 Task-Specific Heads on Top of the Backbone -- 4 Experiments -- 5 Conclusions -- References -- GTCaR: Graph Transformer for Camera Re-localization -- 1 Introduction -- 2 Related Work -- 3 Problem Formulation -- 4 GTCaR Architecture -- 4.1 Architecture Overview -- 4.2 Graph Embedding -- 4.3 Graph Transformer Layer -- 4.4 Graph Loss and Update -- 5 Experimental Results -- 5.1 Experiment Setting -- 5.2 Performance Evaluation -- 5.3 Ablation Study -- 5.4 Discussions and Limitations -- 6 Conclusion -- References -- 3D Object Detection with a Self-supervised Lidar Scene Flow Backbone -- 1 Introduction -- 2 Related Work -- 3 Method -- 3.1 Problem Definition -- 3.2 Self-supervised Scene Flow -- 3.3 Downstream Task: 3D Object Detection -- 4 Implementation Details -- 4.1 Pre-training with Self-supervised Scene Flow -- 4.2 3D Detection Fine-tuning -- 4.3 Datasets -- 4.4 Loss -- 4.5 Experiments -- 5 Results and Discussion -- 5.1 Point-GNN -- 5.2 CenterPoint -- 5.3 PointPillars -- 5.4 Ablation Study -- 5.5 Comparison with Other Self-supervised Learning Methods -- 5.6 Sparse Scene Flow Estimations -- 6 Conclusion -- References -- Open Vocabulary Object Detection with Pseudo Bounding-Box Labels -- 1 Introduction -- 2 Related Work -- 3 Our Approach -- 3.1 Generating Pseudo Box Labels -- 3.2 Open Vocabulary Object Detection with Pseudo Labels -- 4 Experiments -- 4.1 Datasets and Object Vocabulary for Training -- 4.2 Evaluation Benchmarks -- 4.3 Implementation Details -- 4.4 Experimental Results -- 4.5 Ablation Study -- 5 Closing Remarks -- References -- Few-Shot Object Detection by Knowledge Distillation Using Bag-of-Visual-Words Representations -- 1 Introduction -- 2 Related Work -- 3 Method -- 3.1 Position-Aware Bag-of-Visual-Words Model -- 3.2 Knowledge Distillation for Object Detection -- 4 Experiments -- 4.1 Experimental Setup. 4.2 Comparison with State-of-the-Art Methods -- 4.3 Ablation Studies -- 5 Conclusion -- References -- SALISA: Saliency-Based Input Sampling for Efficient Video Object Detection -- 1 Introduction -- 2 Related Work -- 3 SALISA -- 3.1 Saliency Map Generator -- 3.2 Resampling Module -- 3.3 Inverse Transformation Module -- 4 Experiments -- 4.1 Experimental Setup -- 4.2 Results -- 4.3 Ablation Study -- 5 Discussion and Conclusion -- References -- ECO-TR: Efficient Correspondences Finding via Coarse-to-Fine Refinement -- 1 Introduction -- 2 Related Work -- 3 Coarse-to-Fine Refinement Network -- 3.1 Overall Pipeline -- 3.2 Efficient Feature Extraction -- 3.3 Coarse-to-Fine Prediction Refinement -- 3.4 Adaptive Query-Clustering -- 3.5 Implementation Details -- 4 Experiments -- 4.1 Results on HPatches Dataset -- 4.2 Results on KITTI Dataset -- 4.3 Results on ETH3D Dataset -- 4.4 Results on Megadepth Dataset -- 4.5 Ablation Studies -- 5 Conclusions -- References -- Vote from the Center: 6 DoF Pose Estimation in RGB-D Images by Radial Keypoint Voting -- 1 Introduction -- 2 Related Work -- 3 Methodology -- 3.1 Keypoint Voting Scheme Alternatives -- 3.2 Keypoint Estimation Pipeline -- 3.3 RCVPose -- 4 Experiments -- 4.1 Datasets -- 4.2 Implementation Details -- 4.3 Evaluation Metrics -- 4.4 Comparison of Keypoint Voting Schemes -- 4.5 Keypoint Dispersion -- 4.6 Comparison with SOTA -- 5 Conclusion -- References -- Long-Tailed

Instance Segmentation Using Gumbel Optimized Loss -- 1 Introduction  
-- 2 Related Works -- 3 Problem Formulation -- 3.1 Object  
Distribution -- 3.2 Activations as Priors -- 4 Gumbel Activation for  
Long-Tailed Detection -- 4.1 Sigmoid Activation for Object  
Classification -- 4.2 Gumbel Activation for Rare-Class Segmentation --  
4.3 Gumbel Optimised Loss -- 5 Experimental Setup -- 6 Results --  
6.1 Results with Different Sampling Strategies.  
6.2 Integrating Gumbel Activation.

---