2023 18th International Conference on Machine Vision and Applications (MVA 2023)

Hamamatsu, Japan
23-25 July 2023

IEEE Catalog Number: CFP23A66-POD
ISBN: 979-8-3503-0297-4

Copyright © 2023, Institute of Electronics, Information and Communication Engineers (IEICE)
All Rights Reserved

****** This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number: CFP23A66-POD
ISBN (Print-On-Demand): 979-8-3503-0297-4
ISBN (Online): 978-4-88552-343-4

Additional Copies of This Publication Are Available From:

Curran Associates, Inc
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: (845) 758-0400
Fax: (845) 758-2633
E-mail: curran@proceedings.com
Web: www.proceedings.com

TABLE OF CONTENTS

ORAL 1-1 IMAGE ANALYSIS AND PERCEPTION

Weakly-Supervised Deep Image Hashing Based on Cross-Modal Transformer 1
Ching-Ching Yang, Wei-Ta Chu, Shiv Ram Dubey

Contrastive Knowledge Distillation for Anomaly Detection in Multi-Illumination/Focus Display Images 6
Jihyun Lee, Hangil Park, Yongmin Seo, Taewon Min, Joodong Yun, Jaewon Kim, Tae-Kyun Kim

Shape Preservation in Image Style Transfer for Gaze Estimation 11
Daiki Mushiake, Kentaro Otomo, Chihiro Nakatani, Norimichi Ukita

ORAL 1-2 NEURAL ARCHITECTURES AND CONTINUAL LEARNING

Hardware-Aware Zero-Shot Neural Architecture Search 16
Yutaka Yoshihama, Kenichi Yadani, Shota Isobe

Transformer with Task Selection for Continual Learning 21
Sheng-Kai Huang, Chun-Rong Huang

Joint Learning with Group Relation and Individual Action 26
Chihiro Nakatani, Hiroaki Kawashima, Norimichi Ukita

ORAL 1-3 COMPUTER VISION TECHNIQUES AND APPLICATIONS

Multi-Prior Based Multi-Scale Condition Network for Single-Image HDR Reconstruction 32
Haorong Jiang, Fengshan Zhao, Junda Liao, Qin Liu, Takeshi Ikenaga

Black-Box Adversarial Attack Against Visual Interpreters for Deep Neural Networks 37
Yudai Hirose, Satoshi Ono

Padding Investigations for CNNs in Scene Parsing Tasks 43
Yu-Hui Huang, Marc Proesmans, Luc Van Gool

ORAL 2-1 HUMAN-OBJECT INTERACTION DETECTION AND POSE ESTIMATION

QAHOI: Query-Based Anchors for Human-Object Interaction Detection 48
Junwen Chen, Keiji Yanai

Hierarchical Spatio-Temporal Neural Network with Displacement Based Refinement for Monocular Head Pose Prediction 53
Zhe Xu, Yuan Li, Yuhong Li, Songlin Du, Takeshi Ikenaga

QaQ: Robust 6D Pose Estimation Via Quality-Assessed RGB-D Fusion 58
Théo Petitjean, Zongwei Wu, Olivier Laligant, Cédric Demonceaux

ORAL 2-2 APPLICATIONS OF COMPUTER VISION

TomatoDIFF: On–plant Tomato Segmentation with Denoising Diffusion Models * 65
Marija Ivanovska, Vitomir Štruc, Janez Perš

Outline Generation Transformer for Bilingual Scene Text Recognition..... 71
Jui-Teng Ho, Gee-Sern Hsu, Svetlana Yanushkevich, Marina L. Gavrilova

Monocular Blind Spot Estimation with Occupancy Grid Mapping..... 76
Kazuya Odagiri, Kazunori Onoguchi

ORAL 2-3 MEDICAL IMAGING AND VIDEO ANALYSIS

Investigating Self-Supervised Learning for Skin Lesion Classification..... 82
Takumi Morita, Xian-Hua Han

Uncertainty Criteria in Active Transfer Learning for Efficient Video-Specific Human Pose Estimation..... 87
Hiromu Taketsugu, Norimichi Ukita

Leveraging Embedding Information to Create Video Capsule Endoscopy Datasets..... 92
Pere Gilabert, Carolina Malagelada, Hagen Wenzek, Jordi Vitrià, Santi Seguí

ORAL 3-1 VISUAL INTERPRETATION AND RECOGNITION

Interpreting Art by Leveraging Pre-Trained Models 97
Niklas Penzel, Joachim Denzler

Panoptic Segmentation of Galactic Structures in LSB Images..... 103
Felix Richards, Adeline Paiement, Xianghua Xie, Elisabeth Sola, Pierre-Alain Duc

Using Unconditional Diffusion Models in Level Generation for Super Mario Bros..... 109
Hyeon Joon Lee, Edgar Simo-Serra

POSTER 1

Combining Static Specular Flow and Highlight with Deep Features for Specular Surface Detection 114
Hirotaka Hachiya, Yuto Yoshimura

LOTS: Litter on the Sand Dataset for Litter Segmentation 119
Paola Barra, Alessia Auriemma Citarella, Giosué Orefice, Modesto Castrillón-Santana, Angelo Ciaramella

Age Prediction from Face Images Via Contrastive Learning 123
Yeongnam Chae, Poulami Raha, Mijung Kim, Björn Stenger

Bottleneck Transformer Model with Channel Self-Attention for Skin Lesion Classification 129
Masato Tada, Xian-Hua Han

Safe Height Estimation of Deformable Objects for Picking Robots by Detecting Multiple Potential Contact Points..... 134
Jaesung Yang, Daisuke Hagihara, Kiyoto Ito, Nobuhiro Chihara

Mixed Distillation for Unsupervised Anomaly Detection 139
Fuzhen Cai, Siyu Xia

Lifelong Change Detection: Continuous Domain Adaptation for Small Object Change Detection in Everyday Robot Navigation 144
Koji Takeda, Kanji Tanaka, Yoshimasa Nakamura

Tackling Face Verification Edge Cases: In-Depth Analysis and Human-Machine Fusion Approach..... 149
Martin Knoche, Gerhard Rigoll

A Hybrid Wheat Head Detection Model with Incorporated CNN and Transformer..... 154
Sho Harada, Xian-Hua Han

Combining Knowledge Distillation and Transfer Learning for Sensor Fusion in Visible and Thermal Camera-Based Person Classification 159
Vijay John, Yasutomo Kawanishi

MFFPN: An Anchor-Free Method for Patent Drawing Object Detection 164
Yu-Hsien Chen, Chih-Yi Chiu

CG-Based Dataset Generation and Adversarial Image Conversion for Deep Cucumber Recognition..... 169
Hiroaki Masuzawa, Chuo Nakano, Jun Miura

Image Impression Estimation by Clustering People with Similar Tastes 174
Banri Kojima, Takahiro Komamizu, Yasutomo Kawanishi, Keisuke Doman, Ichiro Ide

Generalizable Solar Irradiation Prediction Using Large Transformer Models with Sky Imagery..... 179
Kuber Reddy Gorantla, Aditi Roy

PALF: Pre-Annotation and Camera-LiDAR Late Fusion for the Easy Annotation of Point Clouds 184
Yucheng Zhang, Masaki Fukuda, Yasunori Ishii, Kyoko Ohshima, Takayoshi Yamashita

Safe Landing Zone Detection for UAVs Using Image Segmentation and Super Resolution 189
Anagh Benjwal, Prajwal Uday, Aditya Vadduri, Abhishek Pai

ASD-EVNet: An Ensemble Vision Network Based on Facial Expression for Autism Spectrum Disorder Recognition..... 195
Assil Jaby, Md Baharul Islam, Md Atiqur Rahman Ahad

ViTVO: Vision Transformer Based Visual Odometry with Attention Supervision 200
Chu-Chi Chiu, Hsuan-Kung Yang, Hao-Wei Chen, Yu-Wen Chen, Chun-Yi Lee

An X3D Neural Network Analysis for Runner's Performance Assessment in a Wild Sporting Environment 205
David Freire-Obregón, Javier Lorenzo-Navarro, Oliverio J. Santana, Daniel Hernández-Sosa, Modesto Castrillón-Santana

Grid Sample Based Temporal Iteration and Compactness-Coefficient Distance for High Frame and Ultra-low Delay SLIC Segmentation System..... 210
Yuan Li, Tingting Hu, Ryuji Fuchikami, Takeshi Ikenaga

Towards Achieving Lightweight Deep Neural Network for Precision Agriculture with Maize Disease Detection 215
Carlos Victorino Padeiro, Takahiro Komamizu, Ichiro Ide

Human Pose Prediction by Progressive Generation in Multi-Scale Frequency Domain..... 221
Tomohiro Fujita, Yasutomo Kawanishi

MS-VACSNet: A Network for Multi-scale Volcanic Ash Cloud Segmentation in Remote Sensing Images 226
G Swetha, Rajeshreddy Datla, Chalavadi Vishnu, Krishna Mohan C

POSTER 2

Multi-Class Semantic Segmentation of Tooth Pathologies and Anatomical Structures on Bitewing and Periapical Radiographs 232
James-Andrew Sarmiento, Liushifeng Chen, Prospero Naval

Dynamic Transfer for Domain Adaptation in Crowd Counting..... 237
Shekhor Chanda, Yang Wang

Joint Learning of Images and Videos with a Single Vision Transformer..... 242
Shuki Shimizu, Toru Tamaki

Intra-Frame Skeleton Constraints Modeling and Grouping Strategy Based Multi-Scale Graph Convolution Network for 3D Human Motion Prediction 248
Zhihan Zhuang, Yuan Li, Songlin Du, Takeshi Ikenaga

YOLOv5 with Mixed Backbone for Efficient Spatio-Temporal Hand Gesture Localization and Recognition 253
Luis Acevedo-Bringas, Gibran Benitez-Garcia, Jesus Olivares-Mercado, Hiroki Takahashi

Multi-Plane Projection for Extending Perspective Image Object Detection Models to 360° Images 258
Yasuto Nagase, Yasunori Babazaki, Katsuhiko Takahashi

Diabetic Retinopathy Grading Based on a Sparse Network Fusion of Heterogeneous ConvNeXt Models with Category Attention..... 263
Agustin Castillo-Munguia, Gibran Benitez-Garcia, Jesus Olivares-Mercado, Hiroki Takahashi

Video Anomaly Detection Using Encoder-Decoder Networks with Video Vision Transformer and Channel Attention Blocks..... 268
Shimpei Kobayashi, Akiyoshi Hizukuri, Ryohei Nakayama

Enhancing Retail Product Recognition: Fine-Grained Bottle Size Classification 272
Katarina Tolja, Marko Subašic, Zoran Kalafatic, Sven Loncaric

Low-Level Feature Aggregation Networks for Disease Severity Estimation of Coffee Leaves..... 277
Takuhiro Okada, Yuantian Huang, Guoqing Hao, Satoshi IIzuka, Kazuhiro Fukui

Can You Read Lips with a Masked Face? 282
Taiki Arakane, Chihiro Kai, Takeshi Saitoh

Self-Supervised Pre-Training Boosts Semantic Scene Segmentation on LiDAR Data 287
Mariona Carós, Ariadna Just, Santi Seguí, Jordi Vitrià

Word to Sentence Visual Semantic Similarity for Caption Generation: Lessons Learned 293
Ahmed Sabir

Object Detection for Embedded Systems Using Tiny Spiking Neural Networks: Filtering Noise Through Visual Attention 298
Hugo Bulzomi, Amélie Gruel, Jean Martinet, Takeshi Fujita, Yuta Nakano, Rémy Bendahan

Automatic Reconstruction of Semantic 3D Models from 2D Floor Plans..... 303
Aleixo Cambeiro Barreiro, Mariusz Trzeciakiewicz, Anna Hilsmann, Peter Eisert

Cross-Modal Manifold Cutmix for Self-supervised Video Representation Learning 308
Srijan Das, Michael Ryoo

Domain Adaptation from Visible-Light to FIR with Reliable Pseudo Labels 314
Juki Tanimoto, Haruya Kyutoku, Keisuke Doman, Yoshito Mekada

Quadruped Robot Platform for Selective Pesticide Spraying 319
Hansen Hendra, Yubin Liu, Ryoichi Ishikawa, Takeshi Oishi, Yoshihiro Sato

Generalization of Pixel-Wise Phase Estimation by CNN and Improvement of Phase-unwrapping by MRF Optimization for One-shot 3D Scan 325
Hiroto Harada, Michihiro Mikamo, Ryo Furukawa, Ryusuke Sagawa, Hiroshi Kawasaki

Unsupervised Fall Detection on Edge Devices 330
Takuya Nakabayashi, Hideo Saito

Deep Randomized Time Warping for Action Recognition 335
Yutaro Hiraoka, Kazuhiro Fukui

Malware Detection Using Kernel Constrained Subspace Method 340
Djafer Yahia Messaoud Benchadi, Bojan Batalo, Kazuhiro Fukui

TinyPedSeg: A Tiny Pedestrian Segmentation Benchmark for Top-Down Drone Images 345
Yusuf H. Sahin, Elvin Abdinli, M. Arda Aydin, Gozde Unal

Automated Identification of Surgical Instruments Without Tagging: Implementation in Real Hospital Work Environment 350
Rui Ishiyama, Per Helge Litzheim Frøiland, Stein-Asle Øvrebotn

TECHNICAL EVENT

MVA2023 Small Object Detection Challenge for Spotting Birds: Dataset, Methods, and Results 354
Yuki Kondo, Norimichi Ukita, Takayuki Yamaguchi, Hao-Yu Hou, Mu-Yi Shen, Chia-Chi Hsu, En-Ming Huang, Yu-Chen Huang, Yu-Cheng Xia, Chien-Yao Wang, Chun-Yi Lee, Da Huo, Marc A. Kastner, Tingwei Liu, Yasutomo Kawanishi, Takatsugu Hirayama, Takahiro Komamizu, Ichiro Ide, Yosuke Shinya, Xinyao Liu, Guang Liang, Syusuke Yasui

Ensemble Fusion for Small Object Detection 365
Hao-Yu Hou, Mu-Yi Shen, Chia-Chi Hsu, En-Ming Huang, Yu-Chen Huang, Yu-Cheng Xia, Chien-Yao Wang, Chun-Yi Lee

Small Object Detection for Birds with Swin Transformer 371
Da Huo, Marc A. Kastner, Tingwei Liu, Yasutomo Kawanishi, Takatsugu Hirayama, Takahiro Komamizu, Ichiro Ide

BandRe: Rethinking Band-Pass Filters for Scale-Wise Object Detection Evaluation 376
Yosuke Shinya

Author Index