



Prof. Kiyoharu Aizawa
aizawa@hal.t.u-tokyo.ac.jp

Computer Vision,
Image Processing,
Multimedia

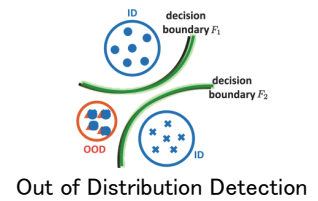
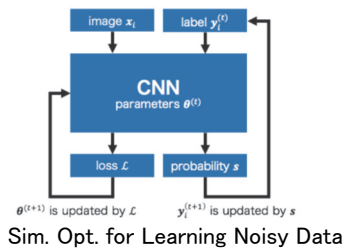
Hongo Campus,
Engineering Building #2
https://www.hal.t.u-tokyo.ac.jp

Image and Multimedia Processing

Image and multimedia technology is expanding to human centric computing. Specific issues and technical problems are diverse. The below are brief description of our current research topics. We operate jointly with Yamasaki, Yamakata and Matsui laboratories. Questions are welcome. Send them to aizawa@hal.t.u-tokyo.ac.jp. Our Web page https://www.hal.t.u-tokyo.ac.jp

1. Fundamentals of Recognition and Learning: Open World

Deep Learning accurately works for closed dataset containing large number of data per class. However, in reality, unknown classes and new classes with small amount of data frequently appear. We are investigating identification and recognition techniques for such open world situation. The topics are methodology for noisy training data, out-of-distribution detection, positive-unlabeled learning, open-set data learning, new category discovery, uncertainty estimation etc.



2. 360° Image Processing·3D: Movie Map

We are investigating 360° image processing. Specifically, we build “movie map” for walkers to explore in a city. Using 360° street videos, we work on many different research issues such as hyperlaspe 360°video, 360°image object detection, accurate vSLAM, intersection detection, depth from 360° image, RoI detection, route view generation, 360° super-resolution, detection of possible direction of travel, avatar in 360° video, real-world Metaverse etc. We are prototyping a platform of virtual exploration.



Route Synthesis of 360° video



MovieMap



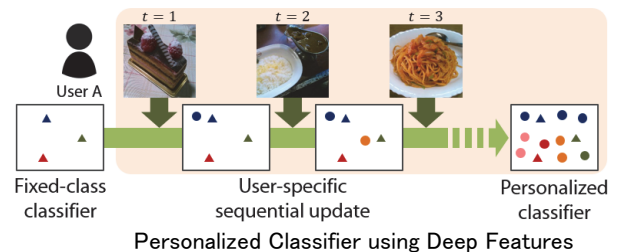
Real World Metaverse



New FoodLog Tool

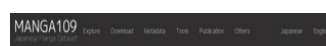
3. Multimedia Life Log: Food Computing

We have been pioneering life logging technology. To pursue specific purpose lifelogging, we focus on research on capture and analysis of our daily food logs (FoodLog). Using the app we developed, food records we captured exceeds 10 million. We are investigating various processing of FoodLog data, such as personalized food recognition, recipe and food record multimodal analysis, building a new tool for athletes and dietitians, prediction of healthy index, etc.

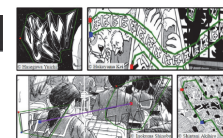


4. Manga & Comic Computing

Manga, our unique culture, is our research target. We have built a world largest scale Manga dataset, and investigate fundamentals of image processing techniques such as retrieval, segmentation, recognition, colorization, creator style transfer, etc. We are also investigating persons’ reading behaviors.



Manga109 Dataset



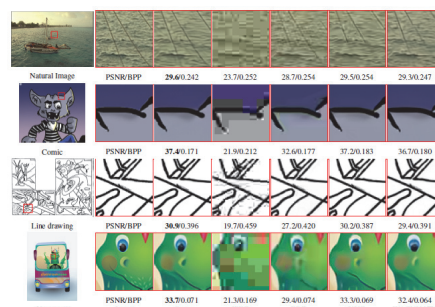
Onomatopoeia Dataset Scene Text Recognition

GT 誹弛湖掬漕
生成 誹弛湖掬漕

Font Generation from a Few Samples

5. Deep Image Compression, Image Generation, Scene Text Recognition

We investigate new image compression techniques using deep learning. Although image compression is a long traditional area of signal processing, a lots of research issues exist. We investigate a universal deep compression that adapts to variety of content outside the training domain. As for image generation, we investigate a Diffusion Model based framework that allows easy modification of the results. We have built a dataset of manga onomatopoeia for the most challenging problem for scene text recognition.



Universal Deep Image Compression