Multimedia Architecture and Processing Laboratory (MAPL)

Dr. Wen-Hsiao Peng
wpeng@cs.nctu.edu.tw
https://mapl.cs.nycu.edu.tw

Computer Science Department
National Yang Ming Chiao Tung Univ. (NYCU), Taiwan

March 25, 2025

About Me

- **Director**, Institute of Data Science, NYCU, Taiwan
- **Director**, Computer Vision Research Center, NYCU ’21-’24
- **Associate Director**, Joint AI Research Labs, Univ. of Washington, USA and NYCU, Taiwan since ’21
- **Visiting Professor**, IBM T. J. Watson Research, New York, USA, ’15-’16
- **Intern**, Intel Microprocessor Research Lab, California, USA, ’00-’01
- **Delegate**, ISO/IEC Moving Picture Experts Group (MPEG), ’04 – Pres.
- **Ph.D.,** Institutes of Electronics Engineering, NCTU, Taiwan, ’05

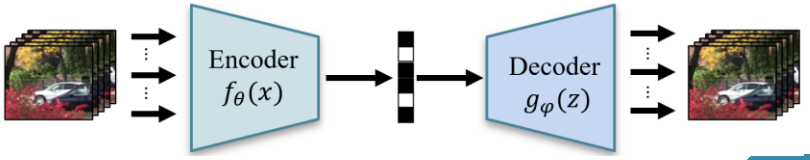
(Visit https://sites.google.com/g2.nctu.edu.tw/wpeng for more details)

Professional Activities in IEEE

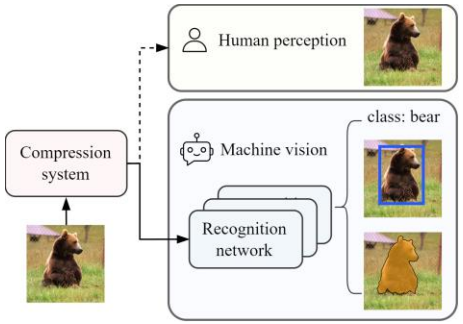
- **Fellow** of the IEEE,'25
- **Editor-in-Chief**, the IEEE Journal on Emerging and Selected Topics in Circuits and Systems (JETCAS), '24 – '25
- **Chair**, the IEEE CASS Visual Signal Processing and Communications (VSPC) Technical Committee, '21 – '22
- **Distinguished Lecturer**, IEEE CASS, '22 – '23; APSIPA, '17-'18
- **Associate Editor-in-Chief/SEB Member/Guest Editor,** IEEE Journal on Emerging and Selected Topics in Circuits and Systems (JETCAS), '16 – Pres.
- **(Senior) Associate Editor**, IEEE Transactions on Circuits and Systems for Video Technology (TCSVT), '19 – '20
- **Guest Editor**, IEEE Transactions on Circuits and Systems II (TCAS-II): Express Briefs, '19
- **Area Chair/Technical Program Chair/Publication Chair/Track Chair/Session Chair** for IEEE and APSIPA conferences

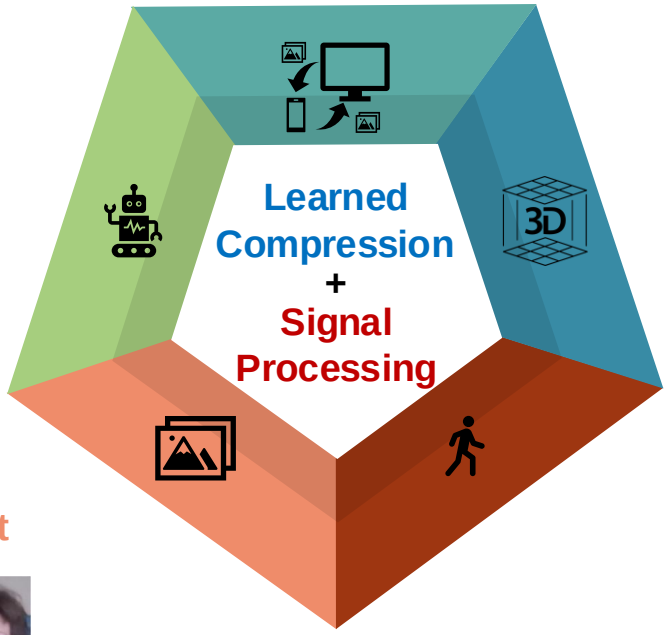
Visual Signal Representation, Processing and Understanding

Learned Image/Video Compression

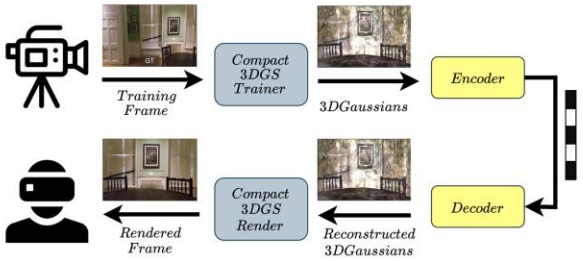


Coding for Machine

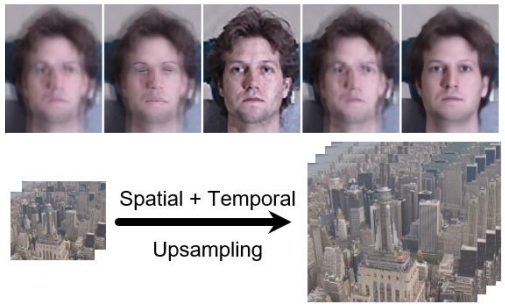




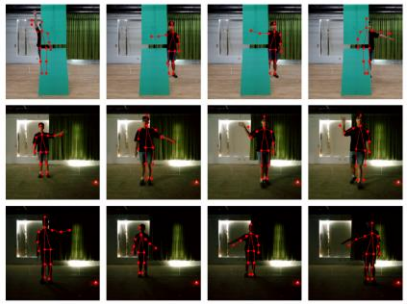
3D View Rendering & Coding



Image/Video Quality Enhancement



Radar-based Human Pose Estimation



Selected Publications

- **Learning-based Image and Video Compression**
  - Learning-based Image/Video Compression (OJCAS'21; **ECCV'22**; TCSVT'23; PCS'24, **CVPR'23, TCSVT'24, ICME'25**)
  - Deep Learning-assisted Video Compression (**DCC'21**)
  - Image/Video Coding for Machines/MLLM (**ICCV'23, ICLR'25**)
  - Rate-distortion-optimized 3DGS Coding (**ICLR'25**)
- **Visual Signal Processing and Computer Vision**
  - Diffusion models (**ICLR'25**)
  - Continuous-scale Video Super-resolution (**ICCV'23**)
  - Radar-based Human Pose Estimation (**WACV'23, BMVC'24**)
  - Video Rescaling (**CVPR'21**)
  - Reinforcement Learning-based Video Prediction (**ICCV'19**)
  - Incremental Learning (ACCV'21)
  - Weakly Supervised Semantic Segmentation (ICME'21)
  - Domain Adaptation for Semantic Segmentation (**CVPR'19**)

Active Contributor to ISO & ITU-T Video Standards

The Moving Picture Experts Group

Follow @MPEGgroup

HOME | STANDARDS | TECHNOLOGIES | MEETINGS | ABOUT MPEG

Our Collaborators

- Academia
  - Leibniz Universität Hannover, **Germany**
  - University of Brescia, **Italy**
  - Yokohama National University, **Japan**
  - Federal University of Pelotas - UFPel, **Brazil**
  - University of Washington, **USA**
  - Poznan University of Technology, **Poland**
- Industry
  - MediaTek, Taiwan
  - Qualcomm, USA

Student Internship

- 2024 Internship with University of Brescia, Italy
- 2024 Internship with Leibniz Universität Hannover, Germany
- 2020 Internship with University of Washington, USA
- 2019 Internship with Poznan University of Technology, Poland
- 2019 Internship with Max Planck Institute, Germany
- 2017 Student Exchange Program with RWTH Aachen University, Germany
- 2015 Internship with InterDigital, San Diego, USA
- 2014 Internship with INRIA, France

Awards

- First Academia-Industry Collaboration Excellent Research Award from the MediaTek Advanced Research Center (**聯發科技** 2024 前瞻研究中心產學合作傑出研究獎)
- The Sixteenth TSC Thesis Award (兩位學生分別榮獲第16屆崇越論文大賞碩士組 AI 資訊類**特優**及**佳作**)
- IEEE ISCAS 2023 Grand Challenge on Neural Network-based Video Coding **Top Creativity Award**
- 15th IPPR outstanding Ph.D. thesis award (中華民國影像處理與圖形識別學會第十五屆博碩士論文獎 - 博士**優等**論文獎)
- IEEE ISCAS 2022 Grand Challenge on Neural Network-based Video Coding **Top performance Award** in the End-to-end Track

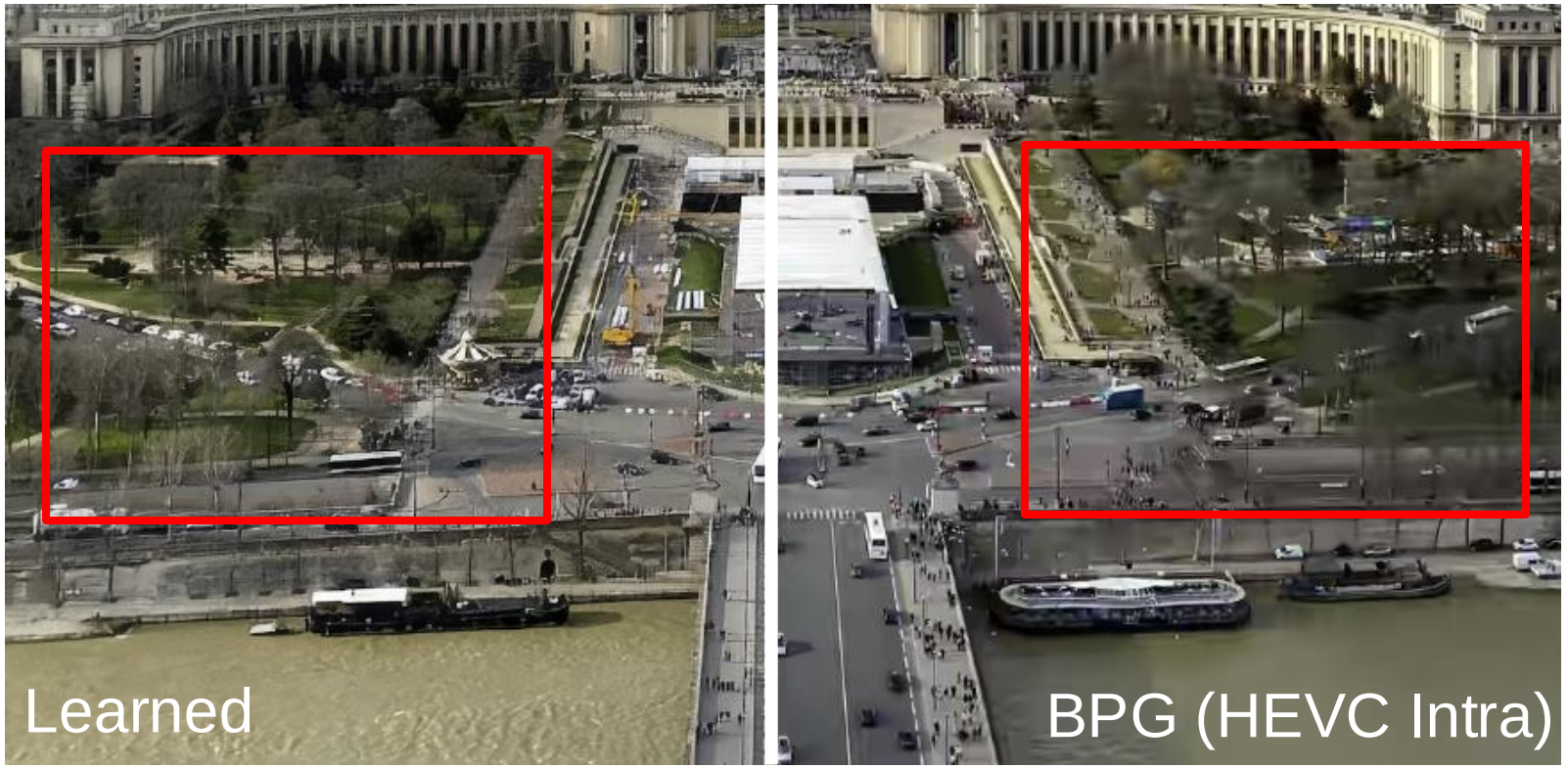
End-to-End Learned Image and Video Coding: Recent Advances and Beyond

Deep Compression Papers

- Deep image/video compression is attracting attention
- **150+** papers on deep image compression since 2017
  - Most adopt the **autoencoder-based** framework with **hyperprior**
- **40+** papers on deep video compression since 2019
  - Potential techniques are still being researched
  - Pixel/feature-domain **residual and conditional coding** are popular approaches

Neural Networks for Image Compression?

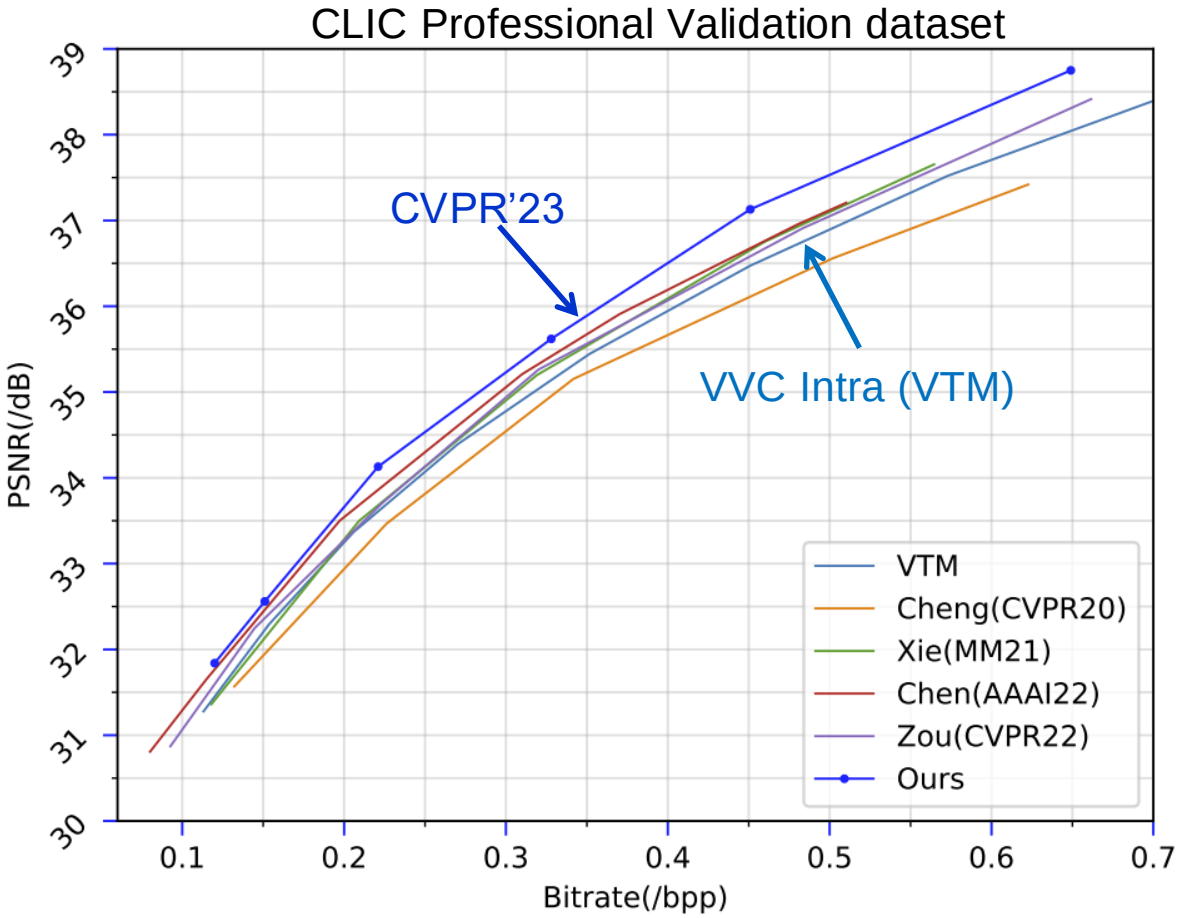
- Neural networks are good at synthesizing image detail
- Easily trained with any differentiable quality metric



https://hific.github.io/

Source: Mentzer et al., "High-Fidelity Generative Image Compression (HIFIC)," NeurIPS 2020

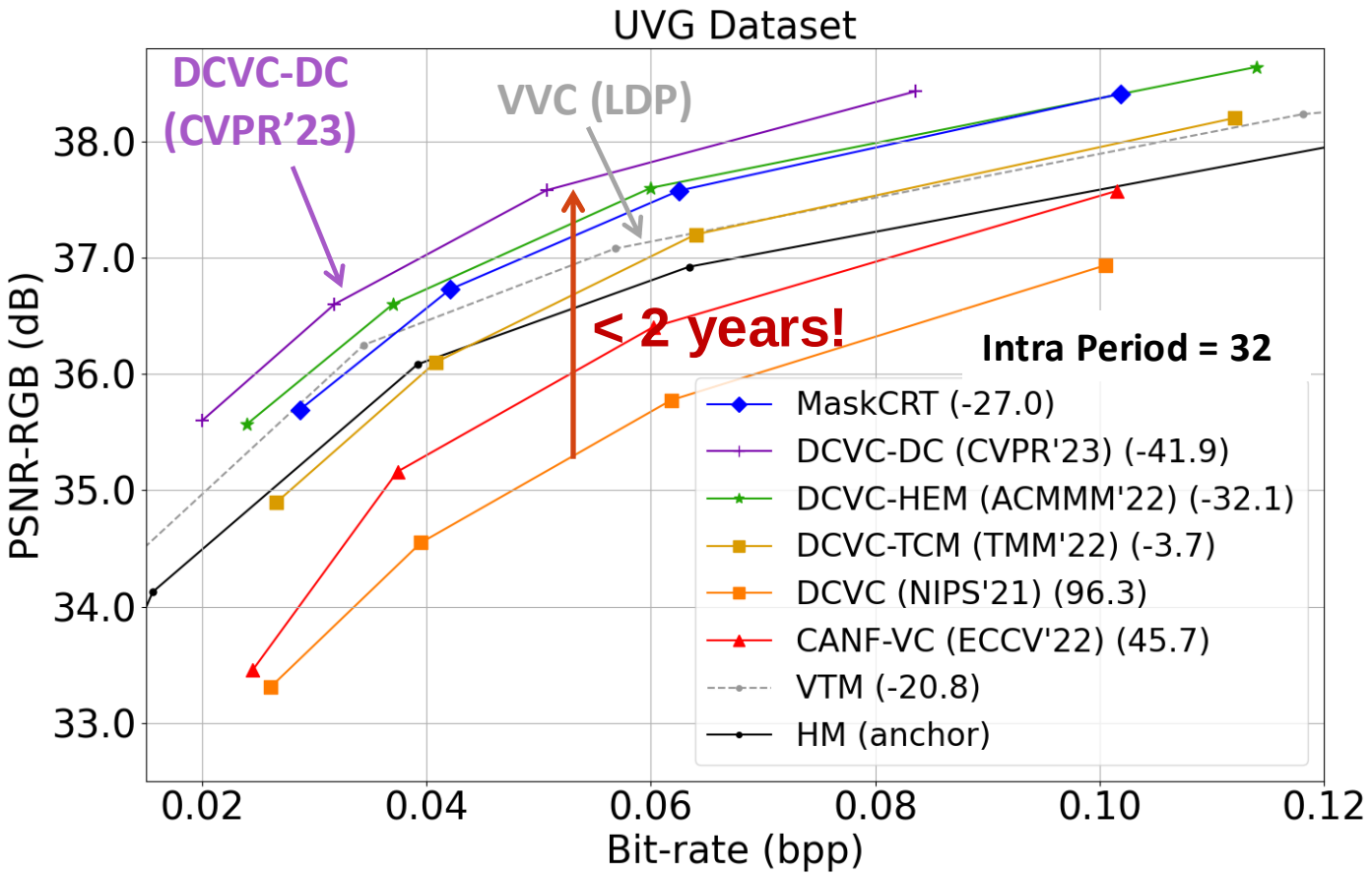
How Good is Learned Image Compression?



Top performer (CVPR'23): **11%** bit rate saving over **VVC Intra**

Source: Liu et al., "Learned Image Compression with Mixed Transformer-CNN Architectures," CVPR 2023

How Good is Learned Video Compression?



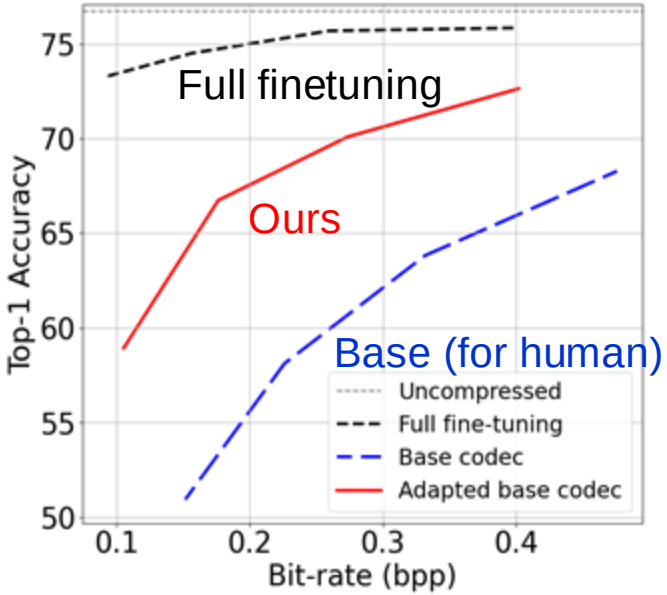
Top performer (CVPR'23): **41.9%** bit rate saving over **HM LDP**

Tutorial on Learned Image/Video Coding at ICCV'23

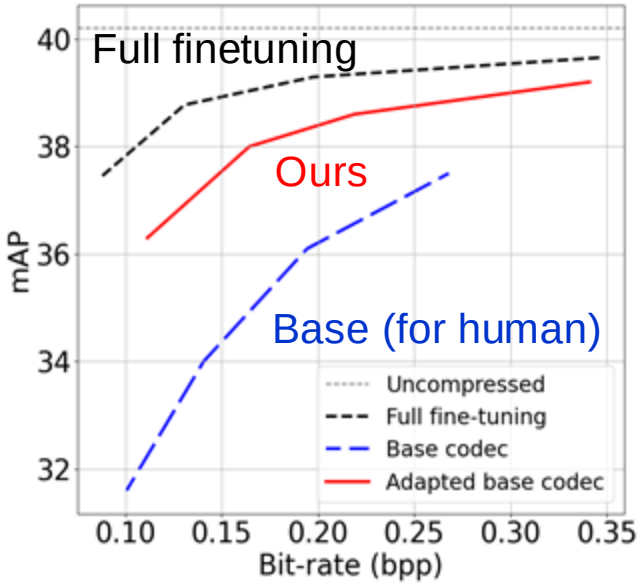
https://nycu-mapl.github.io/ICCV2023_Learned_Image_and_Video_Compression_Tutorial/

Image Coding for Machine Perception

Classification



Object Detection

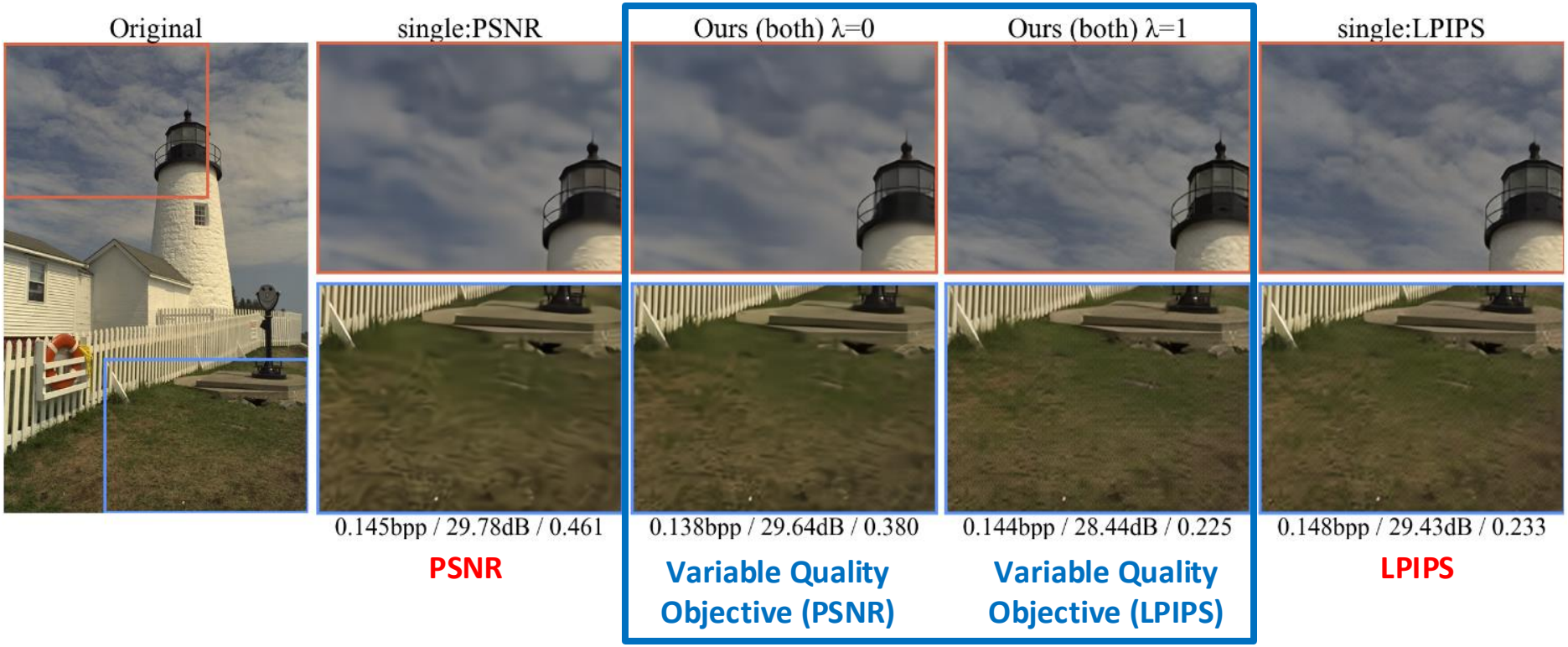


Base codec (for human)

Ours (for machine)

Full finetuning (for machine)

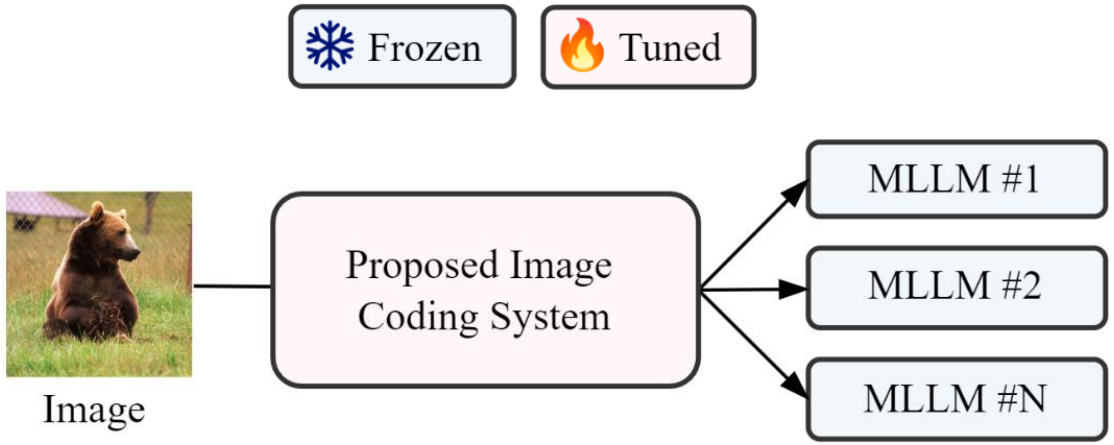
Image Coding with a Variable Quality Objective

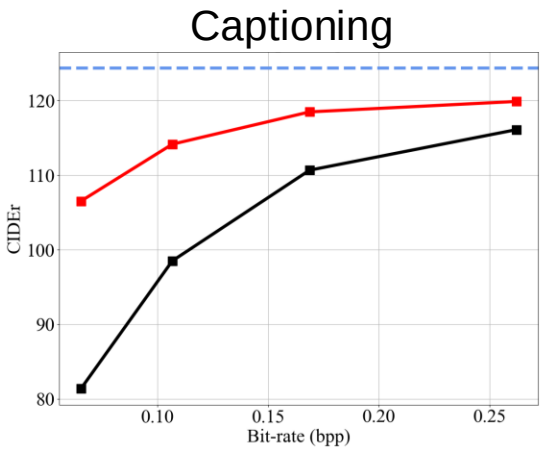


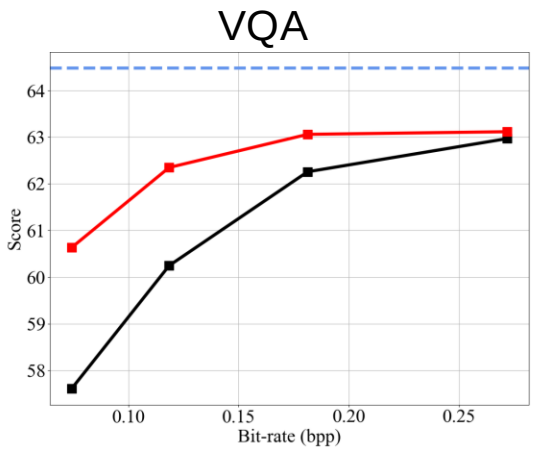
Codecs trained for ONE quality metric

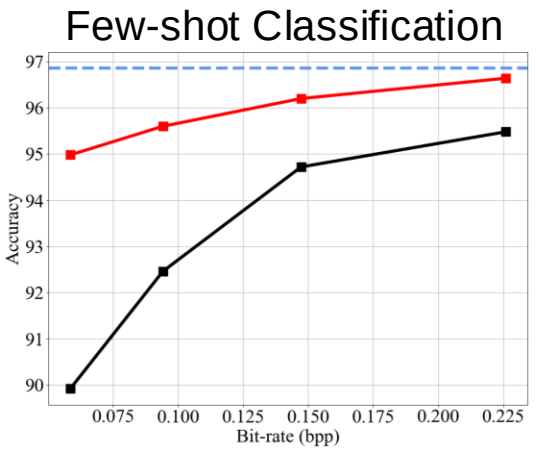
Codecs trained for MULTIPLE quality metrics

Learned Image Coding for Multimodal Large Language Models (MLLMs)









IEEE ISCAS'22 Grand Challenge

The IEEE International Symposium on Circuits and Systems 2022

Grand Challenge on Neural Network-based Video Coding

Top Performance Award in the End-to-end Track

is Presented to

Team NYCU_MAPL

With Team Members

Yung-Han Ho, Chih-Hsuan Lin, Peng-Yu Chen, Mu-Jung Chen, Chih-Peng Chang, Wen-Hsiao Peng and Hsueh-Ming Hang

Dr. Li Zhang, Head of Multimedia Lab of Bytedance Inc.
On Behalf of the Grand Challenge Organizers

ByteDance

IEEE ISCAS'23 Grand Challenge

IEEE

2023 IEEE International Symposium on Circuits and Systems

May 21 through May 25, Conference in Monterey, California, the United States

Top Creativity Award

of The Grand Challenge on Neural Network-based Video Coding

is presented to

Mu-Jung Chen, Hong-Sheng Xie, Cheng Chien , Wen-Hsiao Peng and Hsueh-Ming Hang

For the paper entitled

Learned Hierarchical B-Frame Coding with Adaptive Feature Modulation for YUV 4:2:0 Content

56TH EDITION IEEE ISCAS 2023

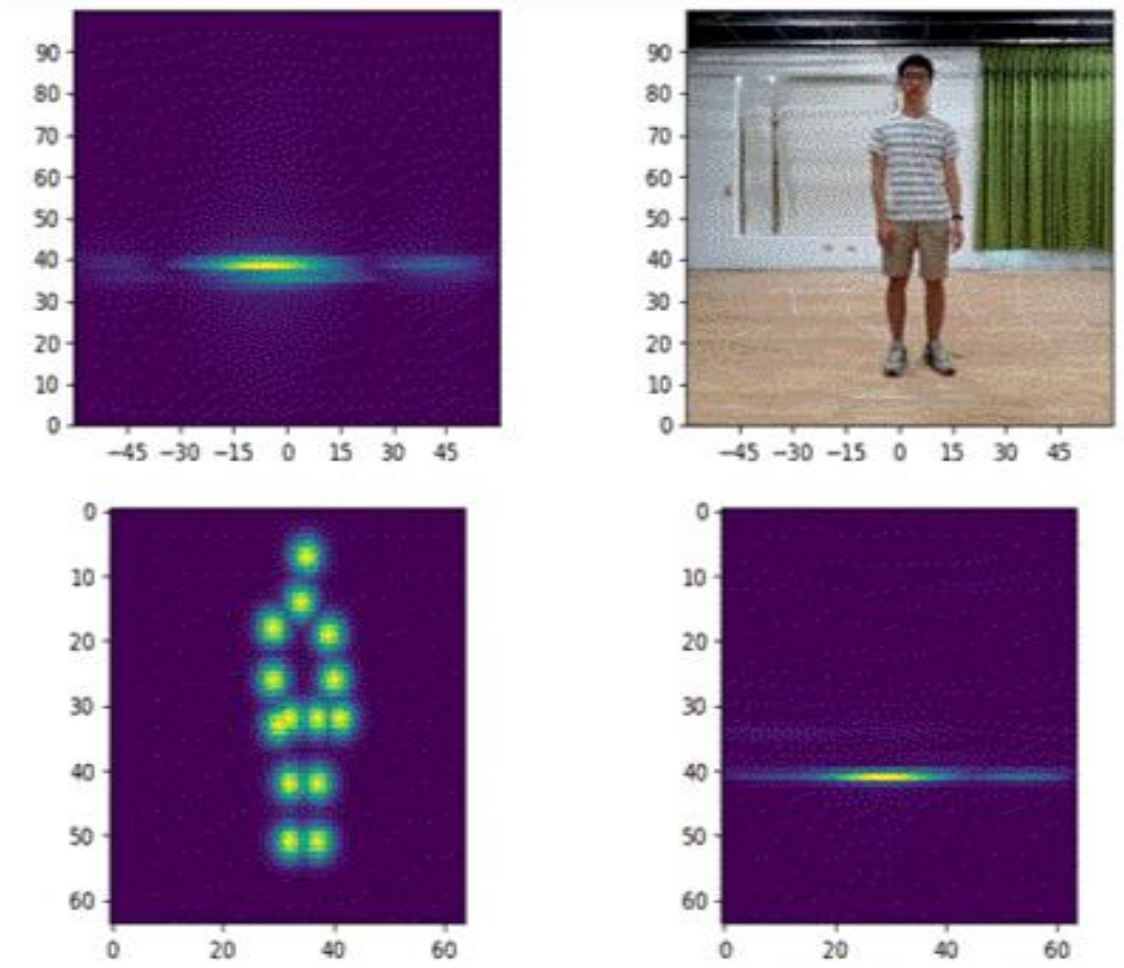
Li Zhang
on behalf of the grand challenge organizer

ByteDance

Visual Signal Processing and Computer Vision

Radar-based Human Pose Estimation

- **Task:** To predict human poses with **mmWave radar**



Radar-based Human Pose Estimation

- Preserving privacy and robustness in low-light conditions
- Suitable for home care or privacy-sensitive medical environments

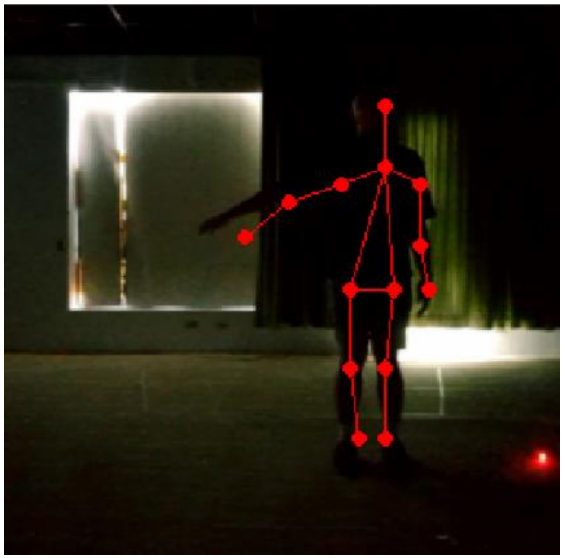
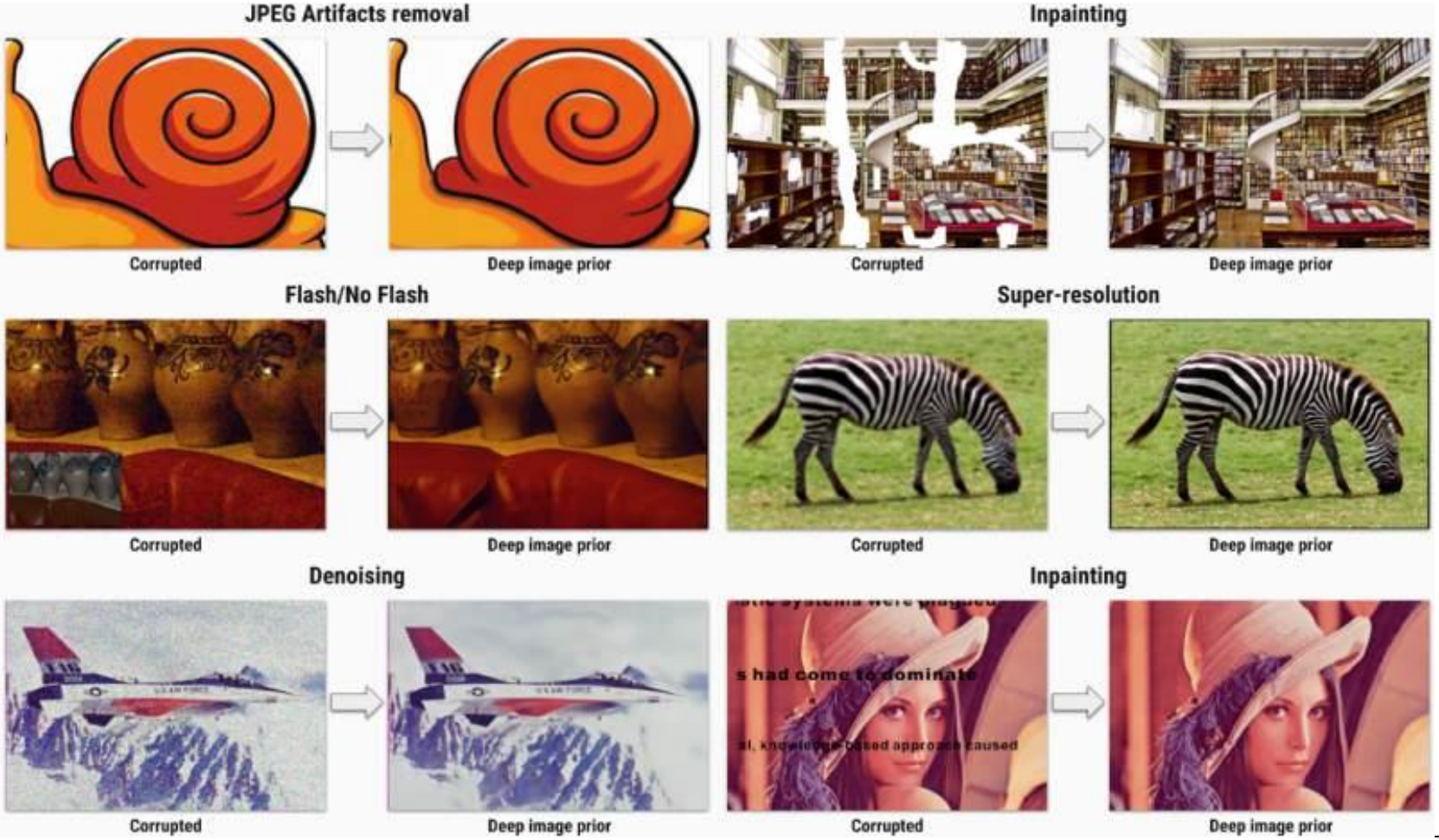


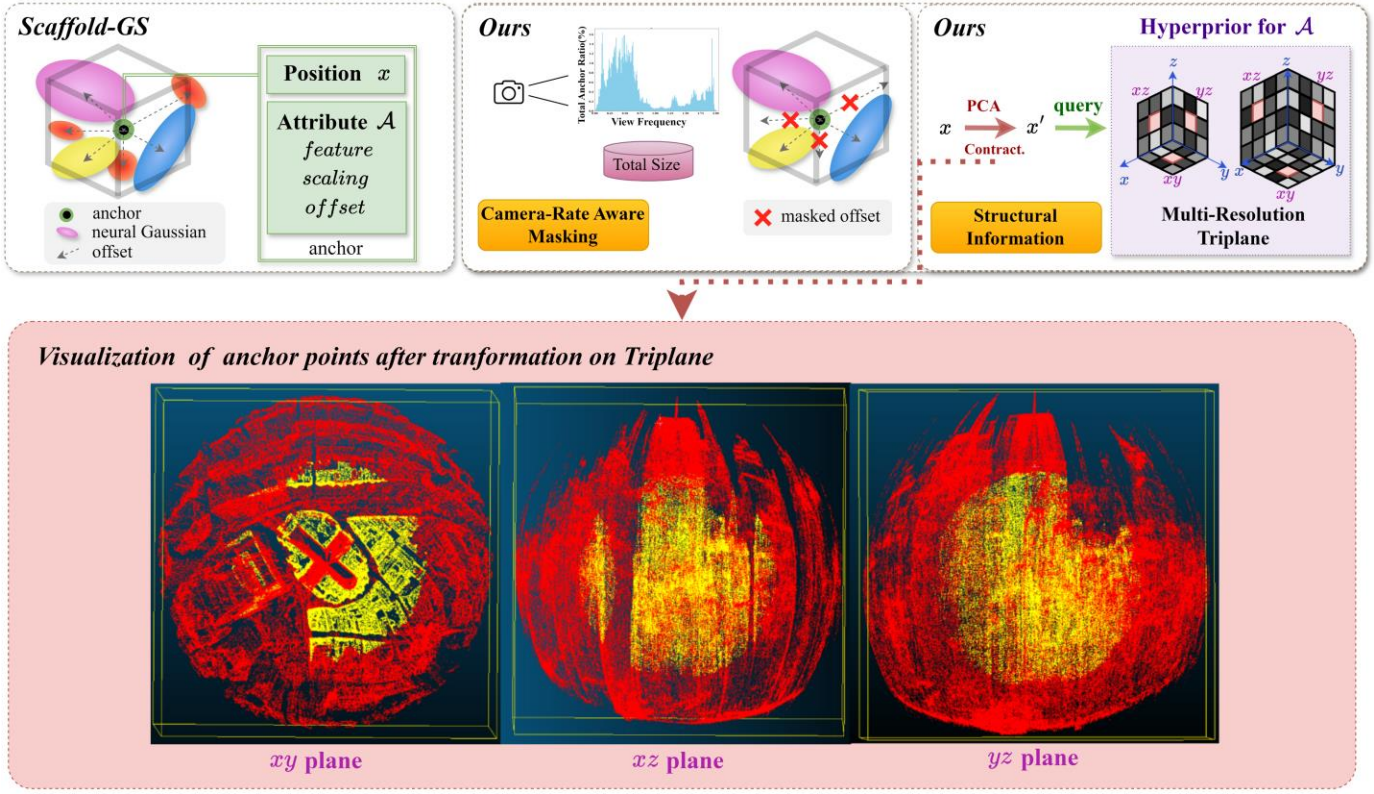
Image restoration

- Application: removal of compression artifacts, etc.



3D Gaussian Splatting Compression

- 3D Scene Compression
  - Grid-Based Implicit Neural Representation
  - 3D Gaussian Splatting for Novel View Synthesis



To Join MAPL

- We are recruiting **Ph.D./Master students**
  - Image processing
  - Machine & deep learning
  - Computer vision

- Monthly stipend: 30,000NTD (1000USD)

- Write to me (wpeng@cs.nctu.edu.tw) to schedule an interview