Multimedia Architecture and Processing Laboratory (MAPL)

Dr. Wen-Hsiao Peng

wpeng@cs.nctu.edu.tw

http://mapl.nctu.edu.tw

Computer Science Department

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

Oct. 17, 2023

About Me (1/2)

- **Director**, Institute of Data Science, NYCU, Taiwan
- **Director**, Computer Vision Research Center, NYCU since '21
- **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 (2/2)

- **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.

- **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

- **Senior Member** of the IEEE since '13

- **Area Chair/Technical Program Chair/Publication Chair/Track Chair/Session Chair** for IEEE and APSIPA conferences

MAPL Research Areas

- **Learning-based Image and Video Compression**
  - Learning-based Image/Video Compression (OJCAS'21; **ECCV'22**; TCSVT'23; **CVPR'23**)
  - Deep Learning-assisted Video Compression (**DCC'21**)
  - Image/Video Coding for Machines (**ICCV'23**)
- **Visual Signal Processing and Computer Vision**
  - Continuous-scale Video Super-resolution (**ICCV'23**)
  - Radar-based Human Pose Estimation (**WACV'23**)
  - 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**)
- **Robotics**
  - Autonomous drone (VCIP'21)

Our Collaborators

- Academia

  - University of Washington, USA

  - Poznan University of Technology, Poland

  - Leibniz Universität Hannover, Germany

  - University of Brescia, Italy

  - Yokohama National University, Japan

- Industry

  - MediaTek, Taiwan

  - Qualcomm, USA

  - Facebook, USA

Tutorial on Learned Image/Video Coding at ICCV'23

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

Tutorial on Learned Image/Video Coding at ICCV'23

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

Visit to Leibniz Universität Hannover, Germany (Oct. 2023)

Visit to Univ. of Brescia, Italy (Oct. 2023)
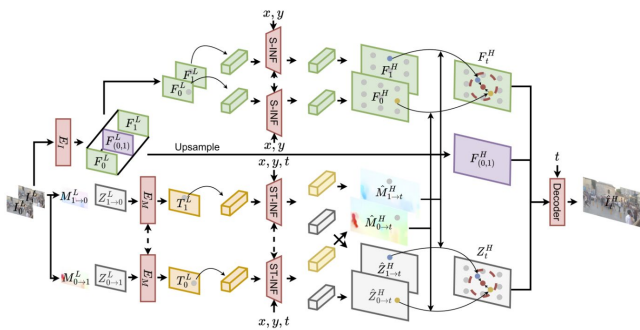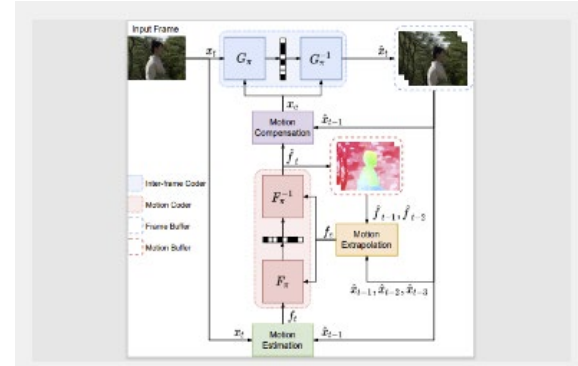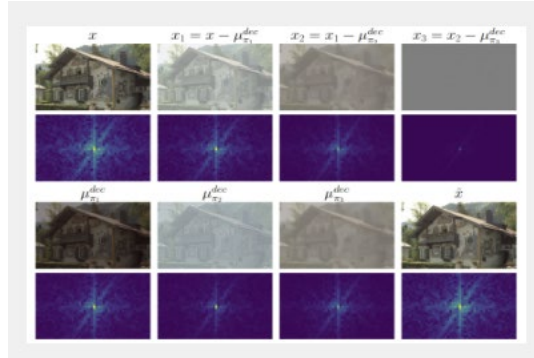
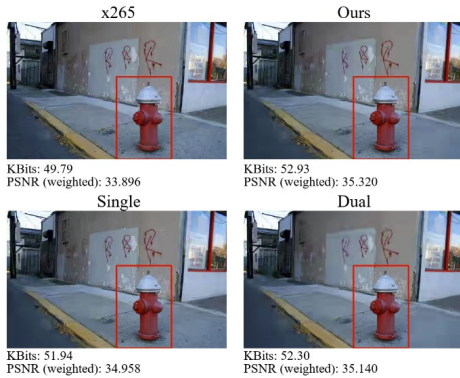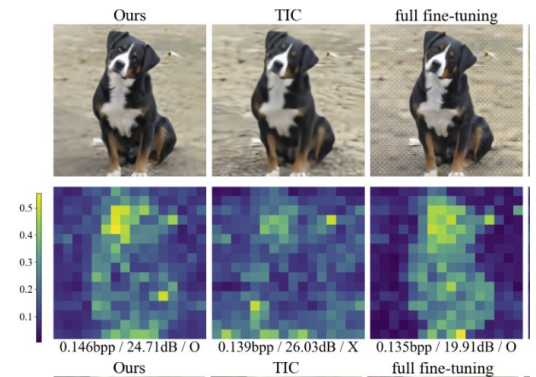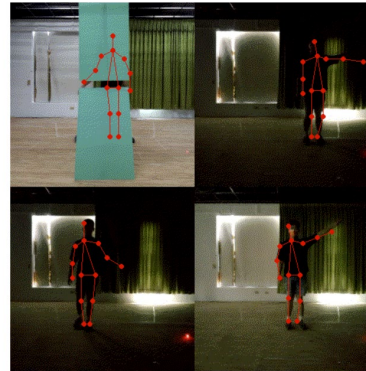# Visit our research project pages

-

Figure 4: The proposed MoTIF for C-STVSR, where the dash double arrows represent the shared-weight networks.

Active Contributor to ISO & ITU-T Video Standards (1/2)

MPEG Products
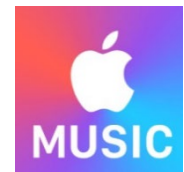
Active Contributor to ISO & ITU-T Video Standards (2/2)

- We were heavily involved in the development of
  - Scalable Video Coding (2004)
  - High-Efficiency Video Coding (2011)
  - HEVC Screen Content Coding Extensions (2014)
  - **JPEG AI learning-based image coding (2022)**

The 90th MPEG Meeting, Xi'an, China

JPEG AI

- To create a **learning-based image coding standard by 2024,** targeting both **human and machine visions**
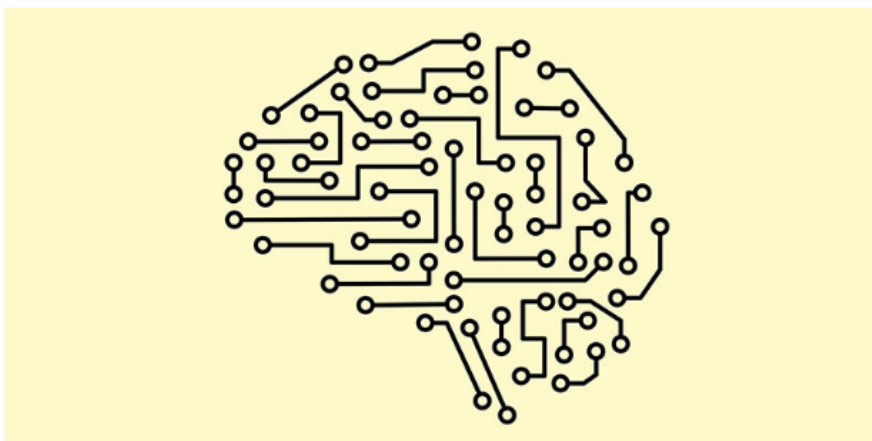
The scope of JPEG AI is the creation of a learning-based image coding standard offering a single-stream, compact compressed domain representation, targeting both human visualization, with significant compression efficiency improvement over image coding standards in common use at equivalent subjective quality, and effective performance for image processing and computer vision tasks, with the goal of supporting a royalty-free baseline.

JPEG AI targets a wide range of applications such as cloud storage, visual surveillance, autonomous vehicles and devices, image collection storage and management, live monitoring of visual data and media distribution. The objective is to design a coding solution that requires significant compression efficiency improvement over coding standards in common use at equivalent subjective quality as well as an effective compressed domain processing for machine learning-based image processing and computer vision tasks. Other key requirements include hardware/software implementation-friendly encoding and decoding, support for 8- and 10-bit depth, efficient coding of images with text and graphics, and progressive decoding.

The Final Call for Proposals for JPEG AI was issued as an outcome of the 94th JPEG Meeting. All documents and practical information can be found in the documentation section.

https://jpeg.org/jpegai/

# 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

Learned   BPG (HEVC Intra)

https://hific.github.io/
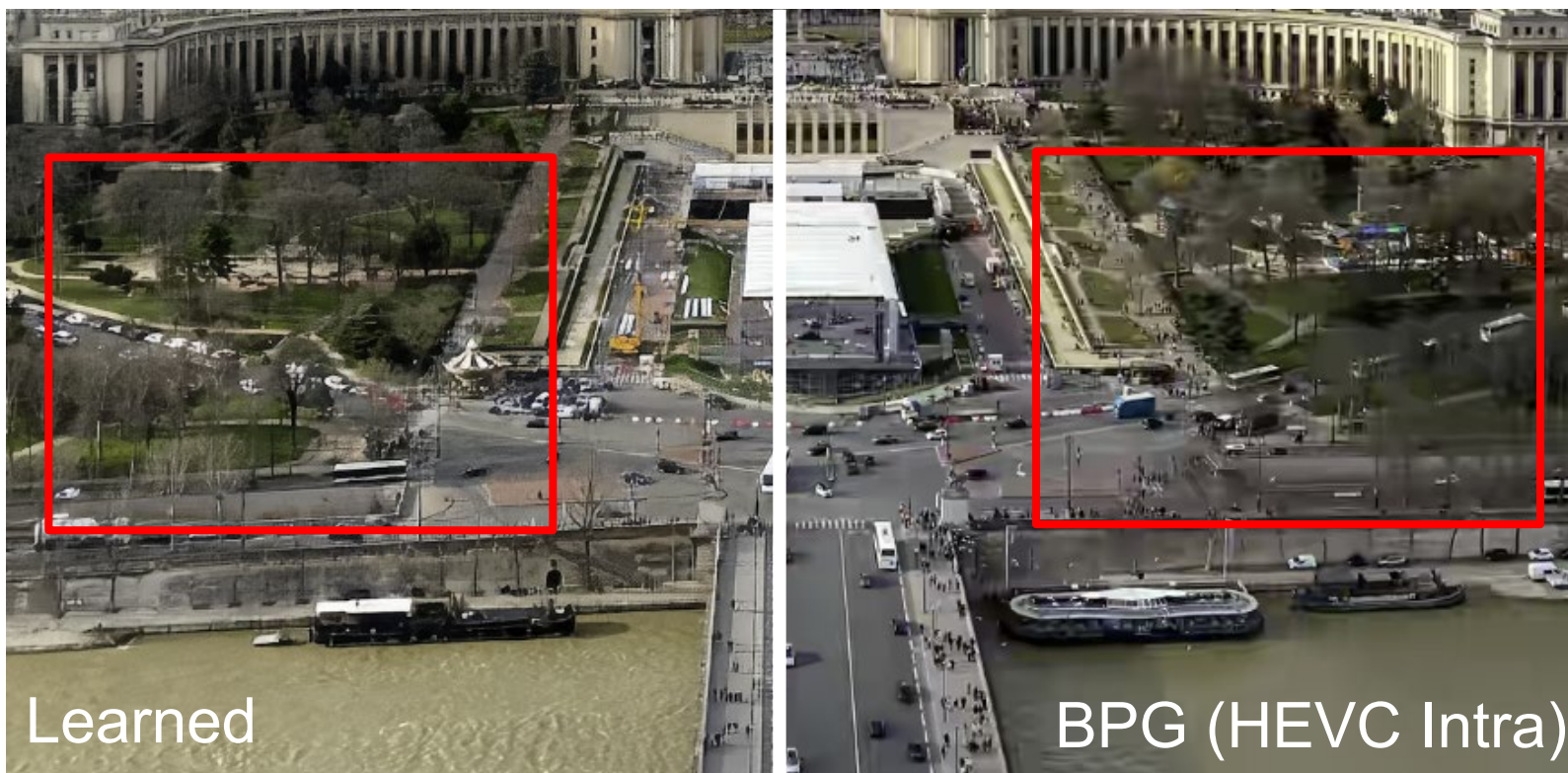
Source: Mentzer et al., "High-Fidelity Generative Image Compression (HIFIC)," NIPS 2020

16

Machine Task: Classification

Base codec (Human)   TransTIC (Machine)   Full finetuning (Machine)

17
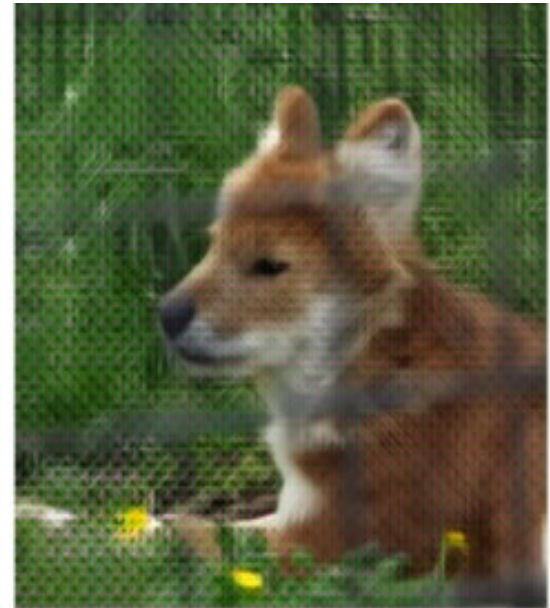
Image Coding with Variable Quality Objective

Original	single:PSNR	Ours (both) λ=0	Ours (both) λ=1	single:LPIPS
	0.145bpp / 29.78dB / 0.461	0.138bpp / 29.64dB / 0.380	0.144bpp / 28.44dB / 0.225	0.148bpp / 29.43dB / 0.233
	PSNR	**Variable Quality Objective (PSNR)**	**Variable Quality Objective (LPIPS)**	**LPIPS**
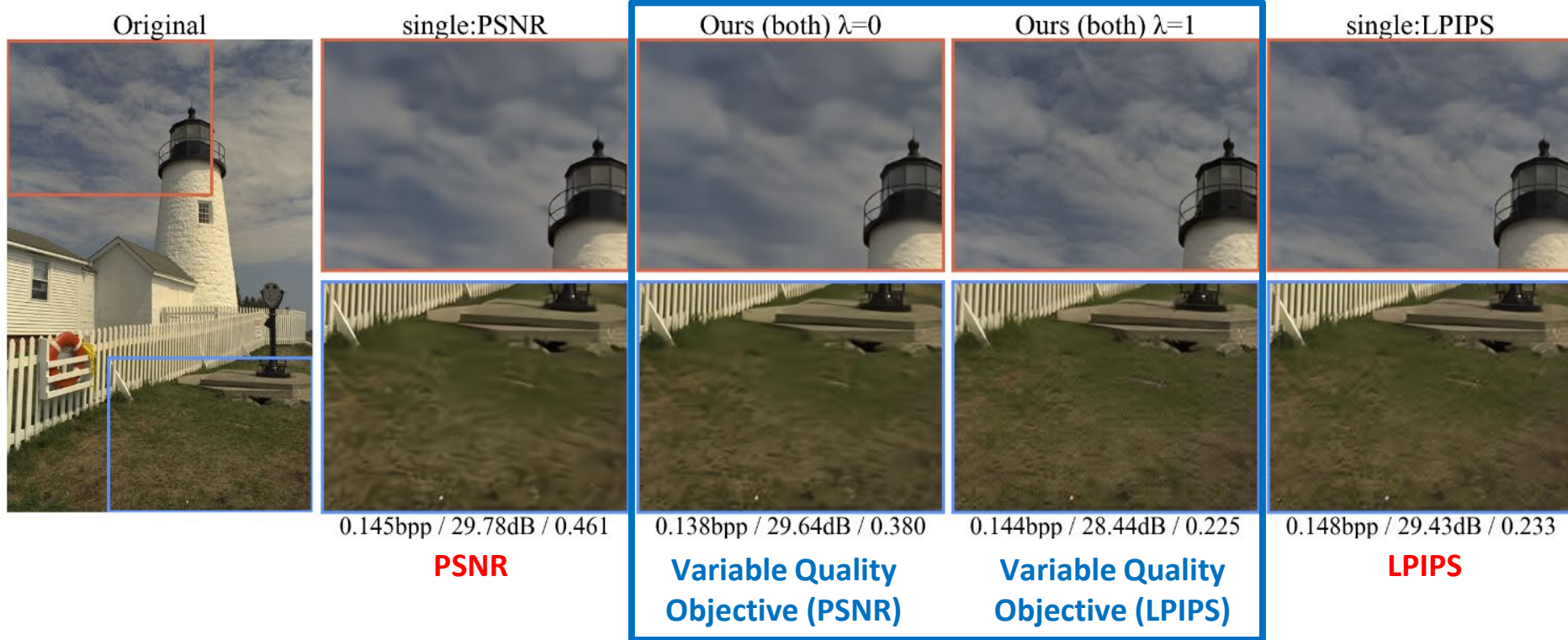
Codecs trained for <u>ONE</u> quality metric
Codecs trained for <u>MULTIPLE</u> quality metrics

Learned Image Compression Challenge in CVPR

- **Challenge on Learned Image Compression (CLIC)**
- CLIC 2018 targeted image coding @ **0.15bpp**
- CLIC 2019 included **transparent track** (PSNR>40dB)
- CLIC 2020 introduced **P-frame** track (**1 P-frame@0.075 bpp**)
- CLIC 2021 introduced **video coding** track (2-sec videos@30Hz) and **multi-rate image coding** track (0.075, 0.15, 0.3 bpp)
- CLIC 2022 introduced **video coding** track (**1mbps & 0.1 mbps** for 720p/1080p@15-60fps) and **multi-rate image coding** track (0.075, 0.15, 0.3 bpp)

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

Sample publications from a Ph.D. student

- Journal (**1**)
  - **Y. H. Ho**, C. C. Chan, W. H. Peng, H. M. Hang, and M. Domanski, "ANFIC: Image Compression Using Augmented Normalizing Flows," **IEEE Open Journal of Circuits and Systems (OJCAS)**, 2021
- Conference (**9**)
  - C.-P. Chang, P.-Y. Chen, **Y.-H. Ho**, and W.-H. Peng, "Deep Incremental Optical Flow Coding for Learned Video Compression," **IEEE International Conference on Image Processing (ICIP)**, Oct. 2022.
  - **Y.-H Ho**, C.-P Chang, P.-Y. Chen, A. Gnutti, and W.-H Peng, "CANF-VC: Conditional Augmented Normalizing Flows for Video Compression," **European Conference on Computer Vision (ECCV),** Oct. 2022.
  - **Y. H. Ho**, C. H. Lin, P. Y. Chen, M. J. Chen, C. P. Chang, W. H. Peng, H. M. Hang, "Learned Video Compression for YUV 4:2:0 Content Using Flow-Based Conditional Inter-Frame Coding," **IEEE International Symposium on Circuits and Systems (ISCAS)**, 2022.
  - **Y. H. Ho**, C. C. Chan, W. H. Peng, H. M. Hang, "End-to-End Learned Image Compression with Augmented Normalizing Flows," **IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)**, 2021
  - **Y. H. Ho**, G. L. Jin, Y. Liang, W. H. Peng, X. Li, "A Dual-Critic Reinforcement Learning Framework for Frame-Level Bit Allocation in HEVC/H.265," **Data Compression Conference (DCC)**, 2021.
  - **Y. H. Ho**, C. C. Chan, D. Alexandre, W. H. Peng, C. P. Chang, "P-frame Coding Proposal by NCTU: Parametric Video Prediction through Backprop-based Motion Estimation," **IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW),** 2020.
  - **Y. H. Ho**, C. C. Chan, W. H. Peng, "Deep Video Prediction Through Sparse Motion Regularization," **IEEE International Conference on Image Processing (ICIP),** 2020
  - **Y. H. Ho**, C. Y. Cho, G. L. Jin, W. H. Peng, "SME-Net: Sparse Motion Estimation for Parametric Video Prediction Through Reinforcement Learning," **IEEE/CVF International Conference on Computer Vision (ICCV)**, 2019
  - **Y. H. Ho**, C. Y. Cho, W. H. Peng, "Deep Reinforcement Learning for Video Prediction," **IEEE International Conference on Image Processing (ICIP),** 2019
- Patent (**1**)
  - W. H. Peng, **Yung-Han Ho**, G. L. Jin, Y. Liang, "Reinforcement Learning Method for Video Encoder" United States Patent, 2021

Awards & Jobs

- Received the **IPPR outstanding Ph.D. thesis award**

- Received job offers from **MediaTek, Qualcomm, etc.** with a very competitive package

To Join MAPL (1/2)

- We are recruiting **Ph.D./Master students**

  - Image processing

  - Machine & deep learning

  - Computer vision

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

To Join MAPL (2/2)

- Read one of our recent publications (after 2020)

- Send me (wpeng@cs.nctu.edu.tw) 10 slides to summarize the ideas