

# 源於GAN的深度學習技術與網路精簡化在電腦視覺的應用

Tyng-Luh Liu 劉庭祿  
中央研究院 資訊科學研究所  
liutyng@iis.sinica.edu.tw



## ★ Introduction

The project is a four-year research effort that focuses on

- 發展源於GAN的深度學習技術
- 發展深層網路精簡化技術
- 開發電腦視覺關鍵新興應用

The research team comprises the following PIs:

- 劉庭祿 (中央研究院資訊所/研究員兼副所長)
- 鮑興國 (台科大資工系/教授兼電資學院副院長)
- 陳煥宗 (清大資工系/副教授)

## ★ Vision and Language

We aim to develop deep learning models to address emergent AI applications that require both computer vision and NLP techniques, including

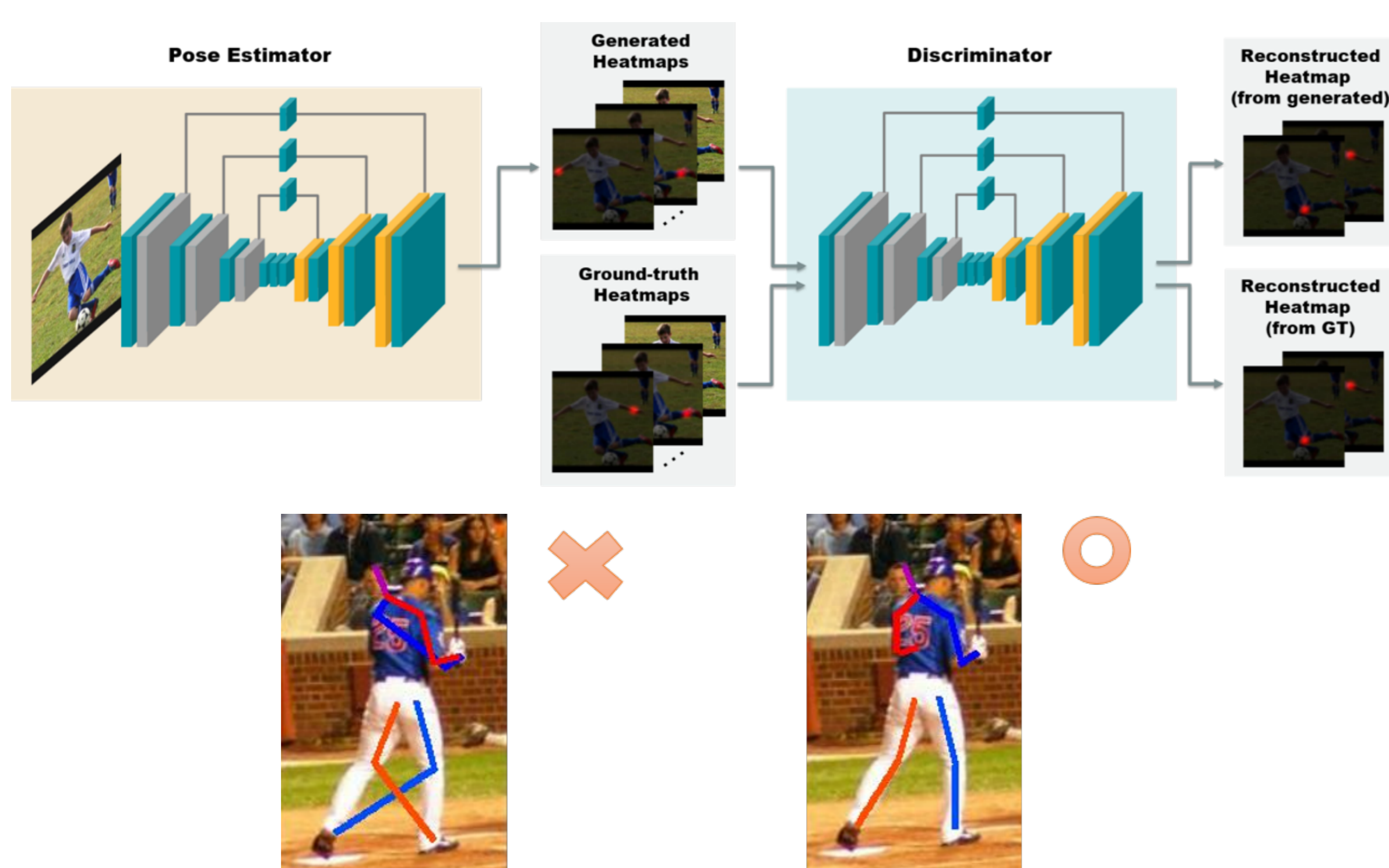
- Image/video annotations
- Image/video captioning
- Image/video/movie question generation
- Image/video/movie question answering
- NL-driven computer vision applications



## ★ Vision and Smart Environment

We also focus on designing deep learning models to improve a wide spectrum of computer vision techniques to tackle challenging applications, including

- Action recognition
- Activity recognition
- Event detection
- Event prediction



## ★ Vision and 360° Video

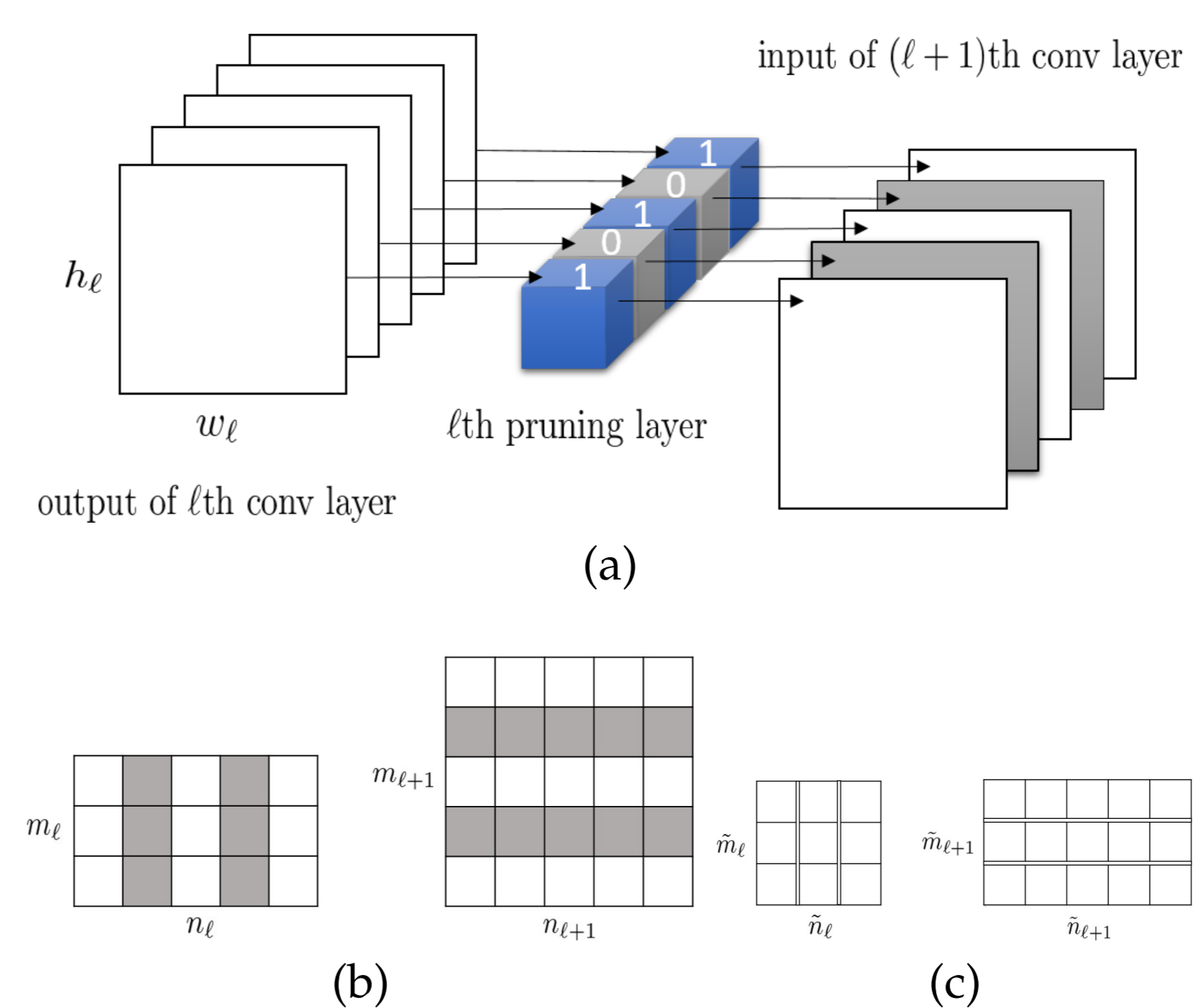
- 360° video object detection
- 360° video object recognition
- 360° video semantic segmentation
- 360° video saliency detection



## ★ Deep Network Compression

- Compression w.r.t. accuracy
- Compression w.r.t. computation efficiency
- Compression w.r.t. storage efficiency

Channel Selection:



Compression Result: FCN-8s on PASCAL VOC 2011

Layers	Kernel Size	Baseline #Channels	Pruned Model #Channels	↓ Ratio
Conv1-1	3	64	33	48.4%
Conv1-2	3	64	40	67.8%
Conv2-1	3	128	69	66.3%
Conv2-2	3	128	57	76.0%
Conv5-1	3	512	372	52.5%
Conv5-2	3	512	354	49.8%
Conv5-3	3	512	171	76.9%
Linear 6	7	4096	64	99.5%
Linear 7	1	4096	410	99.8%
score <sub>fr</sub>	1	21	21	0.0%
upscore <sub>2</sub>	4	21	21	0.0%
score <sub>pool4</sub>	1	21	21	0.0%
upscore <sub>pool4</sub>	4	21	21	0.0%
score <sub>pool3</sub>	1	21	21	0.0%
upscore <sub>8</sub>	16	21	21	0.0%
Total FLOPs reduced			63.7%	
Total #Params reduced			95.6%	

Semantic Segmentation:

