

# Deep Learning for Computer Vision

Fall 2022

<https://cool.ntu.edu.tw/courses/189345> (NTU COOL)

<http://vllab.ee.ntu.edu.tw/dlcv.html> (Public website)

Yu-Chiang Frank Wang 王鈺強, Professor

Dept. Electrical Engineering, National Taiwan University

# What to Cover Today...

- Self-Supervised Learning (SSL)
  - Pretext Tasks vs. Contrastive Learning
  - SSL Beyond Images
- Invited Talk
  - 10 Secrets You Need to Know About Software Engineering & Career Planning

Would you expect an Olympic athlete to compete for Gold medal without any preparation or practice? Of course not!

Career Planning is about choices, yet it is extremely difficult to make decisions about your future when you have been doing the same thing as everyone else for the past two decades. In this talk, we will discuss some of the fundamentals that are crucial to grooming a customized career path that no employers would ever tell you. We will discuss tactics and strategies to ace the interview for the software engineering space, as well as the hiring trend to keep up-to-date demanded skills.



**Linda Huang**  
Practice Leader  
Paul Wright Taiwan Limited

I specialize in executive search for c-level suites, and am the Practice Leader for SCM and Engineering recruiting at Paul Wright Taiwan. Holder of a Bachelor's degree in architecture from CMU, went onto studying for 2 more master's degrees at CMU and the University of Oxford in England, focusing on Urban Design and Corporate Social Responsibility. I have a genuine interest in improving humanity in general, from the time I studied architecture to my time now as a recruitment consultant/headhunter, it is "people" that's the center of my focus and fuels my passion for my work and life.

### 10 Secrets You Need To Know About Software Engineering & Career Planning

**Tuesday December 6<sup>th</sup> / BL-112 / 10:10am - 12pm**  
(Host: Prof. Frank Wang)

Do you wish to become the CTO of a corporation? Do you want to make more money than your colleagues? Are you looking for opportunities to set up your own startup? Or you just want to find a place to chill for the rest of your life? All your questions will be answered to crush competitors and land your dream jobs. Let's start building your own roller-coaster ride.



**Sharine Chen**  
Associate Director  
Paul Wright Taiwan Limited

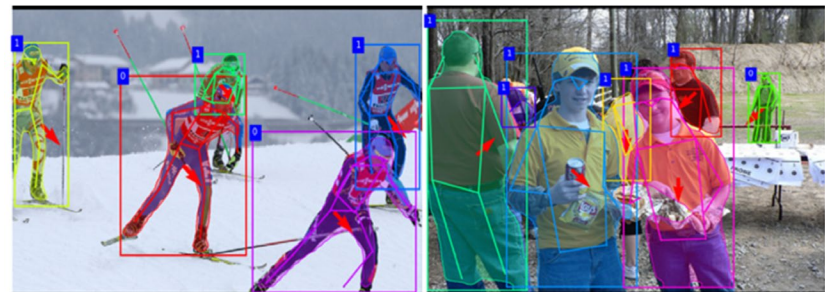
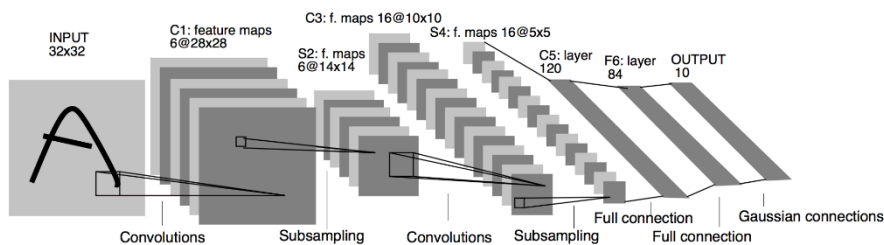
I am the Associate Director in Paul Wright Taiwan, specializing in mid-senior, executive-level search within the software engineering space. Having delivered numerous complex search mandates with domestic software start-ups to global organizations, I have an extensive network of candidates with skills across cutting-edge digital, FinTech & AI, data science, and mobile technology. Prior to a career in recruitment, I was working in account management roles in a Fortune 500 American IT service.



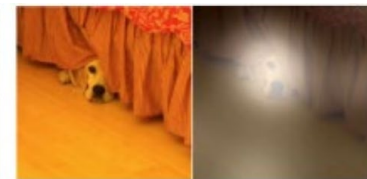
PAUL WRIGHT  
國立中央大學  
National Central University

# Supervised Learning

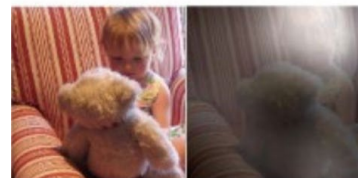
- Deep learning plus supervised learning are rocking the world ...



A woman is throwing a frisbee in a park.



A dog is standing on a hardwood floor.



A little girl sitting on a bed with a teddy bear.



A group of people sitting on a boat in the water.



- In real world scenarios, data-annotation is quite **time-consuming**
- Could one exploit supervised signals from **unlabeled** data?

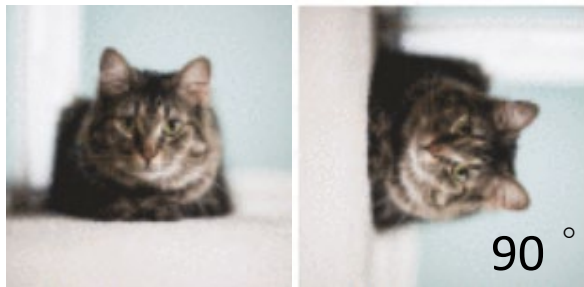


# Self-Supervised Learning (SSL)

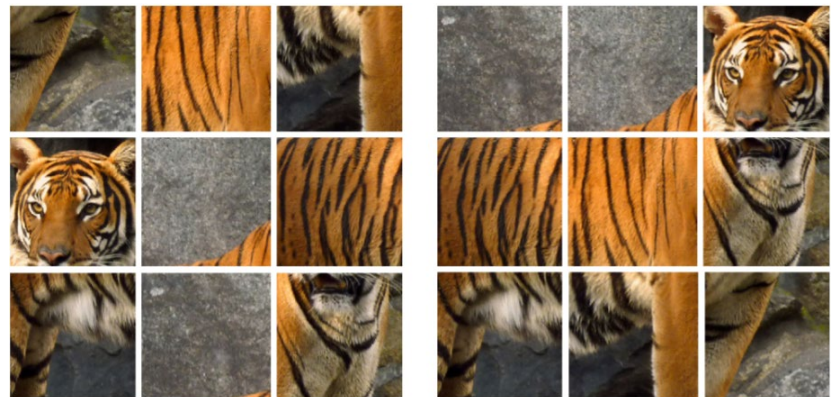
- Learning discriminative representations from **unlabeled** data
- Create self-supervised tasks via **data augmentation**



Colorization



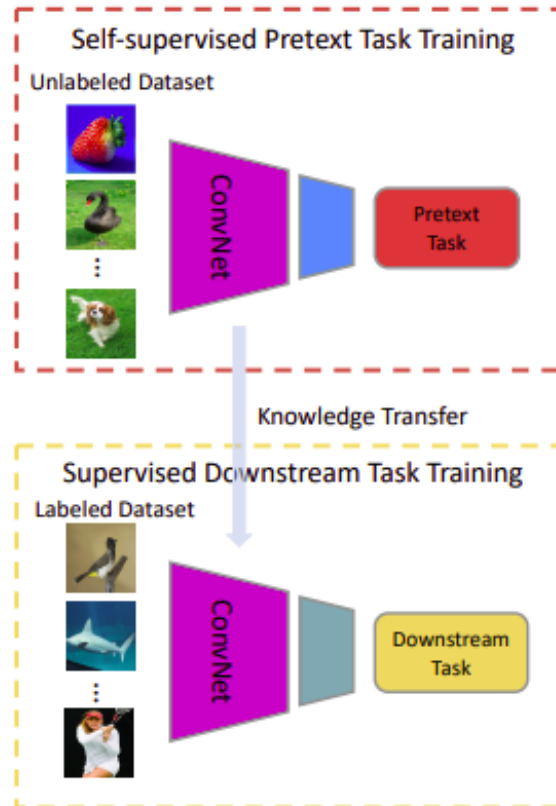
Rotation



Jigsaw Puzzle

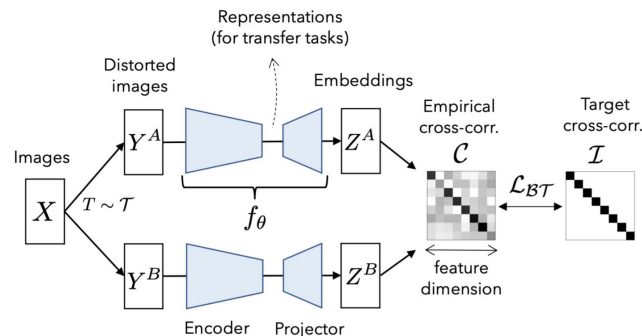
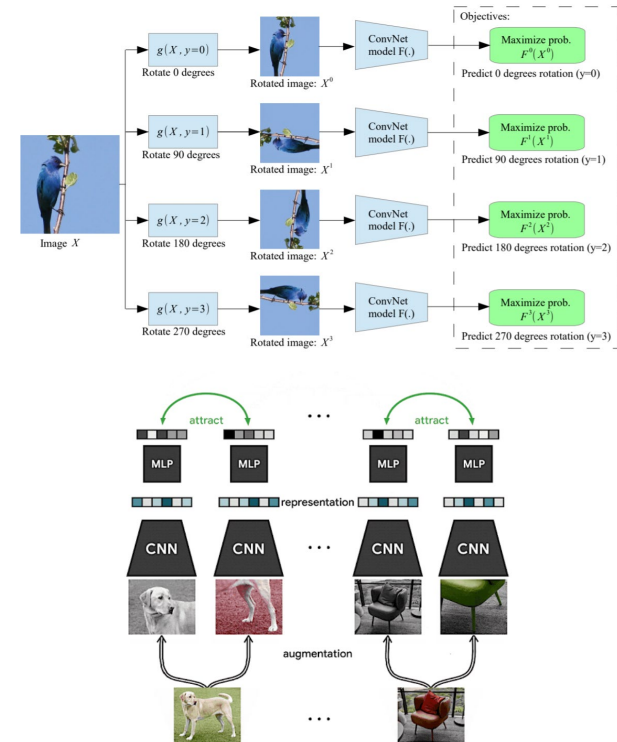
# Self-Supervised Learning (SSL)

- Self-Supervised Pretraining
- Supervised Fine-tuning



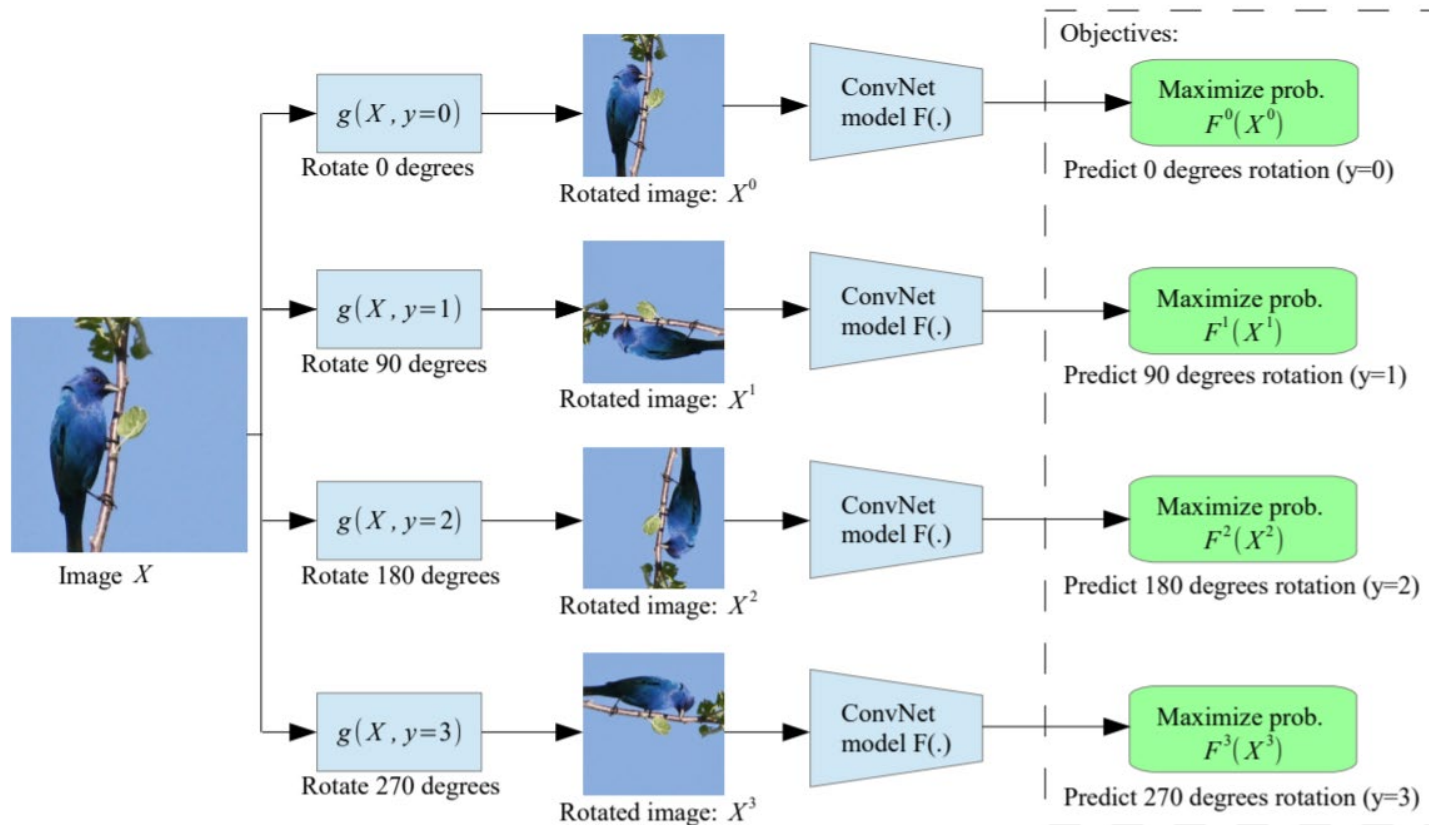
# Self-Supervised Learning (SSL)

- Pretext Tasks
  - Jigsaw (ECCV'16)
  - RotNet (ICLR'18)
- Contrastive Learning
  - CPC (ICML'20)
  - SimCLR (ICML'20)
- Learning w/o negative samples
  - BYOL (NeurIPS'20)
  - Barlow Twins (ICML'21)



# RotNet

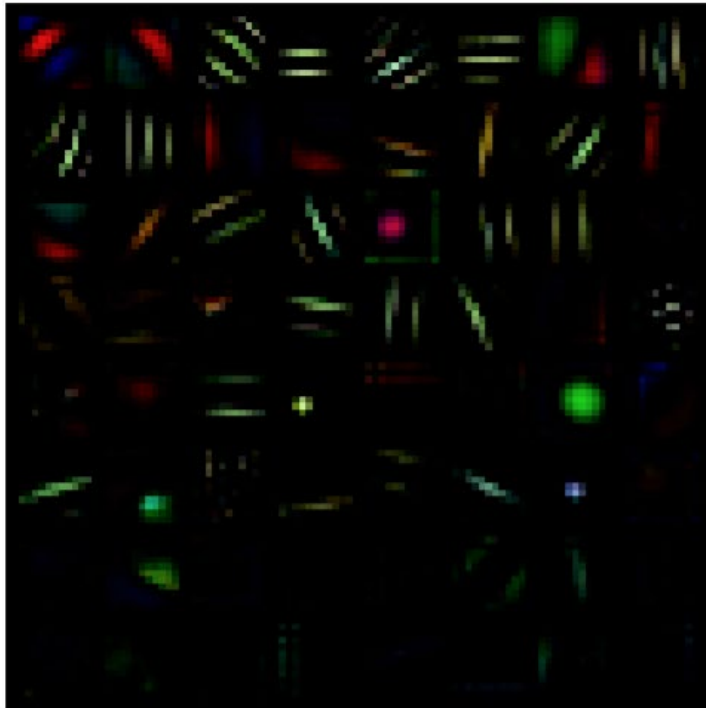
- Learning to predict the **rotation** angle



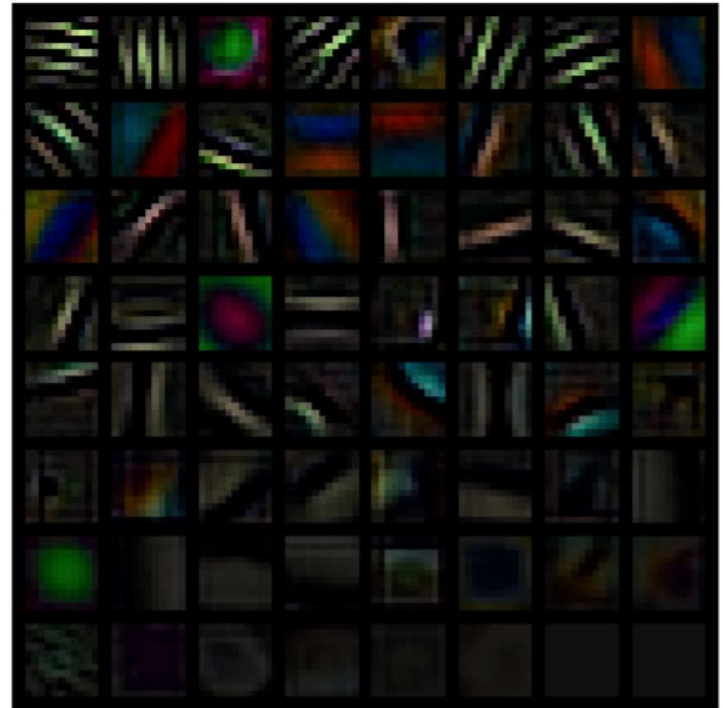


# RotNet

- Filters learned with SSL exhibit more **variety**



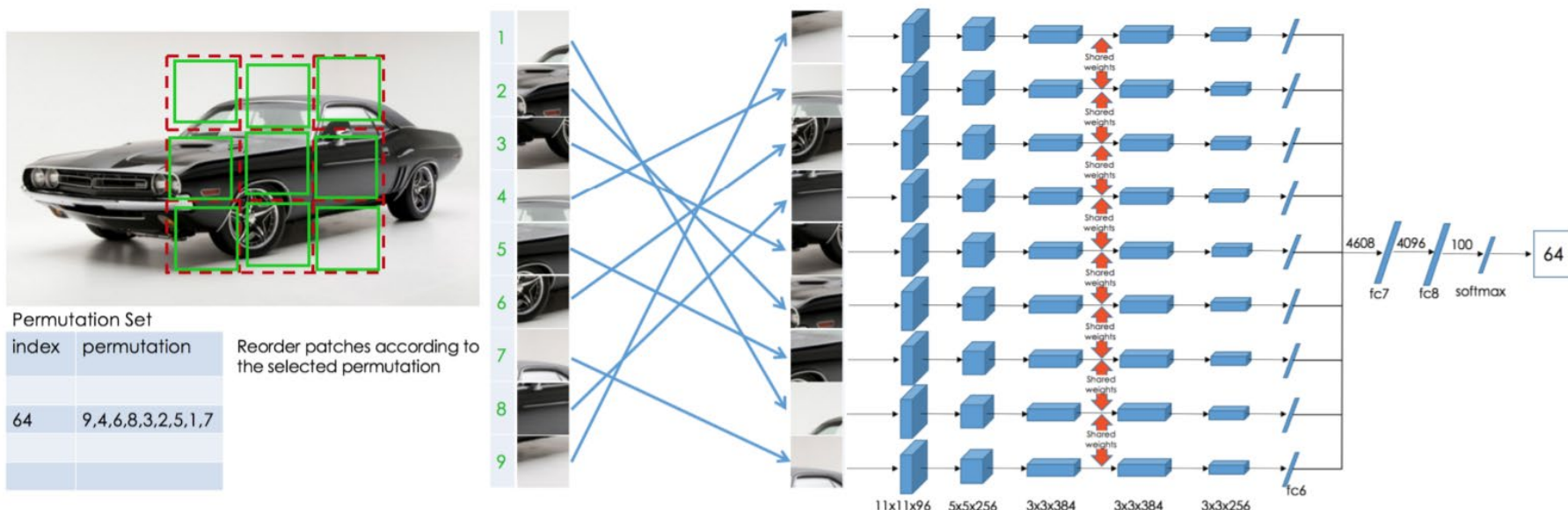
(a) Supervised



(b) Self-supervised to recognize rotations

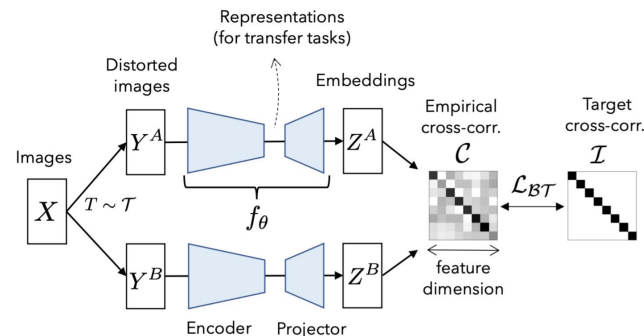
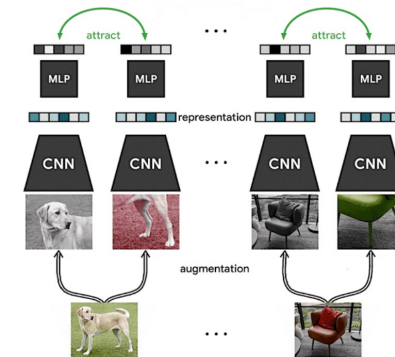
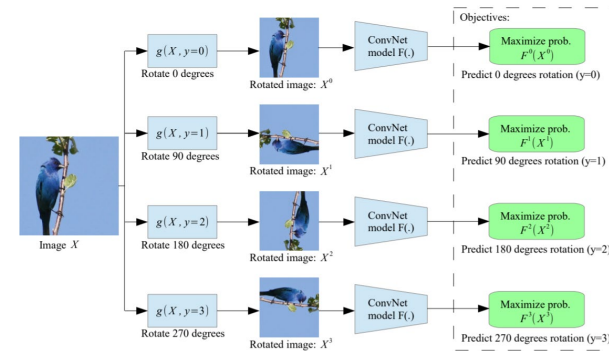
# Jigsaw Puzzle

- Assign the **permutation index** and perform augmentation
- Solve jigsaw puzzle by predicting the permutation index



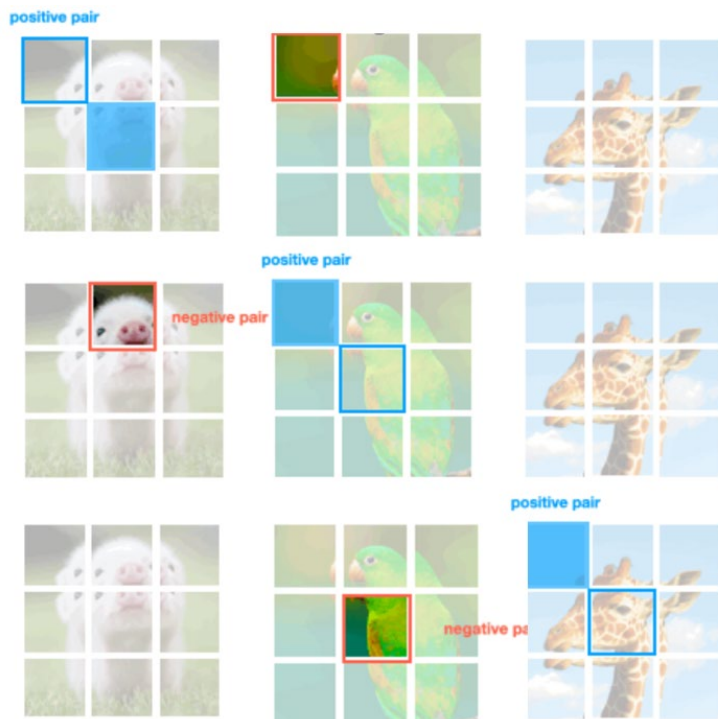
# Self-Supervised Learning (SSL)

- Pretext Tasks
  - Jigsaw (ECCV'16)
  - RotNet (ICLR'18)
- Contrastive Learning
  - CPC (ICML'20)
  - SimCLR (ICML'20)
- Learning w/o negative samples
  - BYOL (NeurIPS'20)
  - Barlow Twins (ICML'21)



# Contrastive Predictive Coding (CPC)

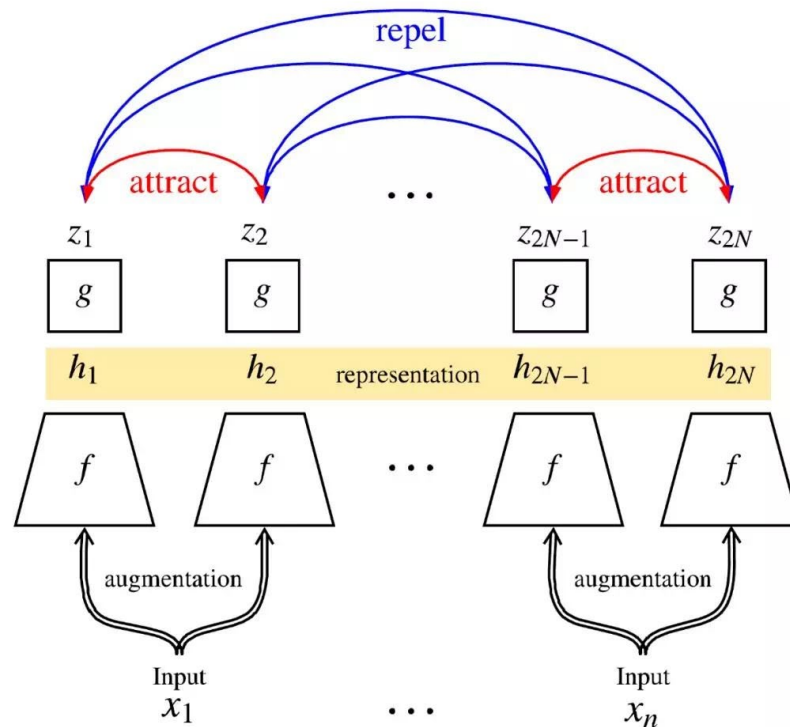
- Sample **positive** patches from itself and **negative** patches from other images
- **Maximize** positive similarities and **minimize** negative ones



$$\mathcal{L}_{\text{CPC}} = - \sum_{i,j,k} \log \frac{\exp(\hat{\mathbf{z}}_{i+k,j}^T \mathbf{z}_{i+k,j})}{\underbrace{\exp(\hat{\mathbf{z}}_{i+k,j}^T \mathbf{z}_{i+k,j})}_{\text{positive}} + \sum_l \underbrace{\exp(\hat{\mathbf{z}}_{i+k,j}^T \mathbf{z}_l)}_{\text{negative}}}$$

# SimCLR

- **Attract** augmented images and **repel** negative samples
- Improve the representation quality with **projection heads** ( $g$ )...why?





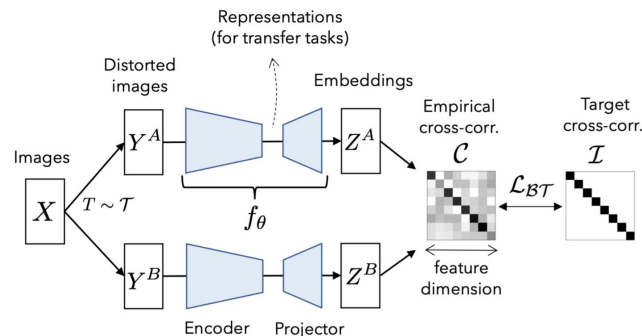
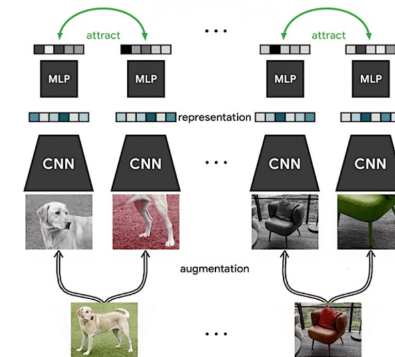
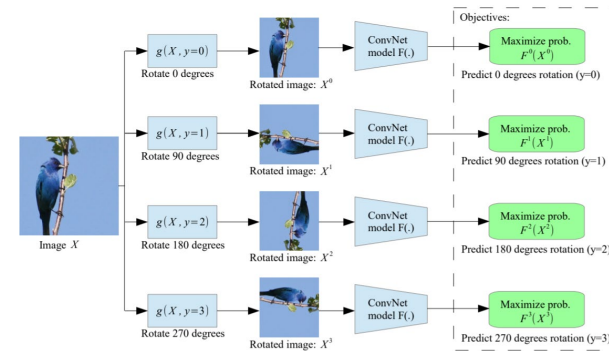
# SimCLR

- Experiments on semi-supervised settings

Method	Architecture	Label fraction	
		1%	10%
		Top 5	
Supervised baseline	ResNet-50	48.4	80.4
<i>Methods using other label-propagation:</i>			
Pseudo-label	ResNet-50	51.6	82.4
VAT+Entropy Min.	ResNet-50	47.0	83.4
UDA (w. RandAug)	ResNet-50	-	88.5
FixMatch (w. RandAug)	ResNet-50	-	89.1
S4L (Rot+VAT+En. M.)	ResNet-50 (4×)	-	91.2
<i>Methods using representation learning only:</i>			
InstDisc	ResNet-50	39.2	77.4
BigBiGAN	RevNet-50 (4×)	55.2	78.8
PIRL	ResNet-50	57.2	83.8
CPC v2	ResNet-161(*)	77.9	91.2
SimCLR (ours)	ResNet-50	75.5	87.8
SimCLR (ours)	ResNet-50 (2×)	83.0	91.2
SimCLR (ours)	ResNet-50 (4×)	<b>85.8</b>	<b>92.6</b>

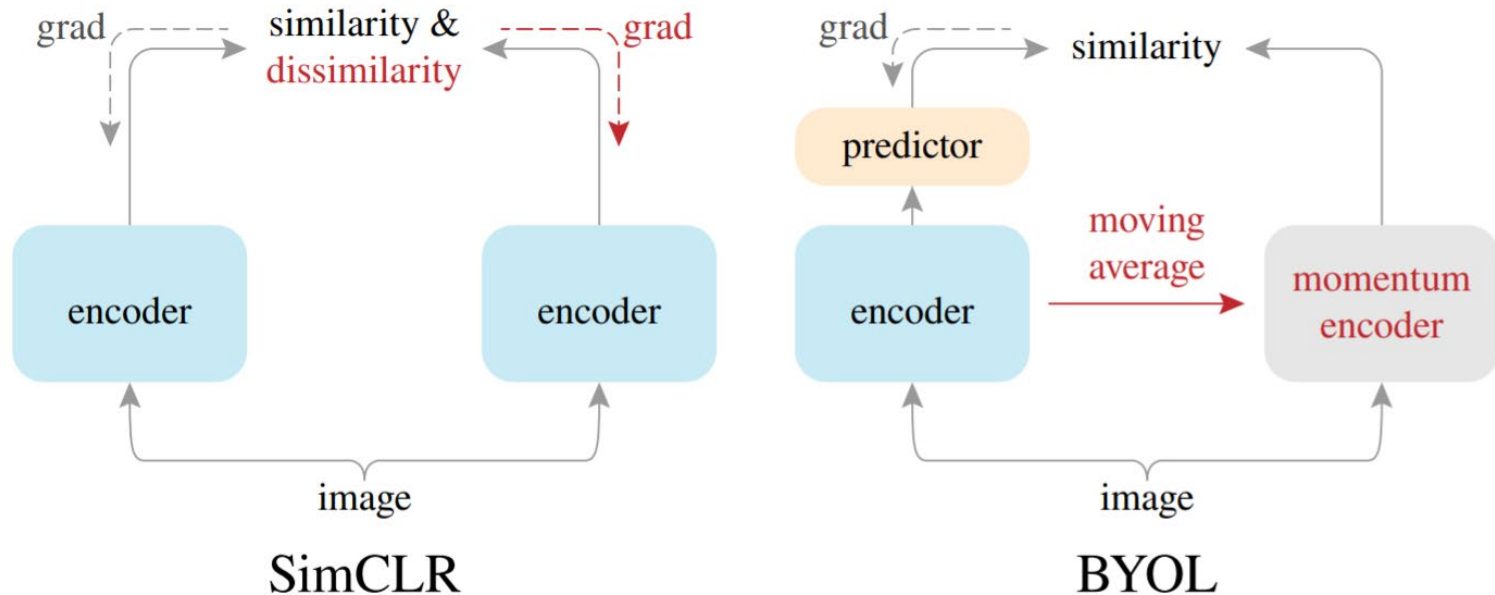
# Self-Supervised Learning (SSL)

- Pretext Tasks
  - Jigsaw (ECCV'16)
  - RotNet (ICLR'18)
- Contrastive Learning
  - CPC (ICML'20)
  - SimCLR (ICML'20)
- Learning w/o negative samples
  - BYOL (NeurIPS'20)
  - Barlow Twins (ICML'21)



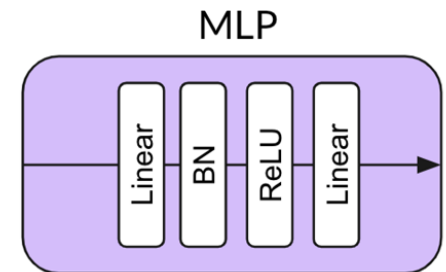
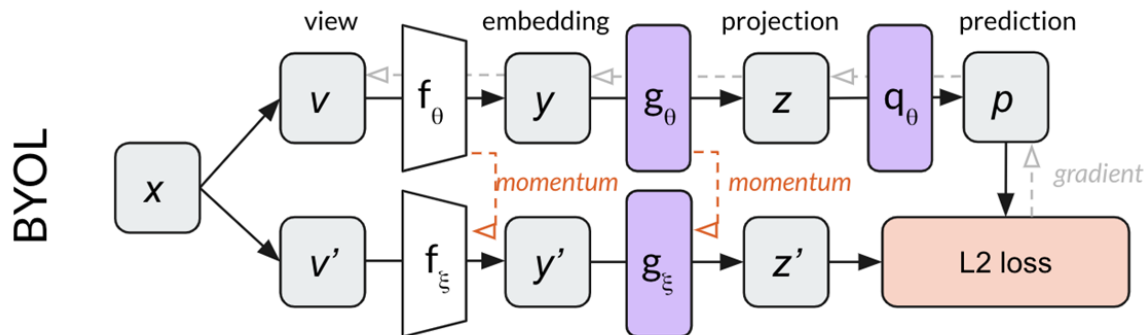
# Bootstrap Your Own Latent (BYOL)

- No need of negative pairs
- Introduce the **predictor** for asymmetry to avoid collapse
- Exponential Moving Average (**EMA**)  $\theta_T \leftarrow \tau\theta_T + (1 - \tau)\theta_S$



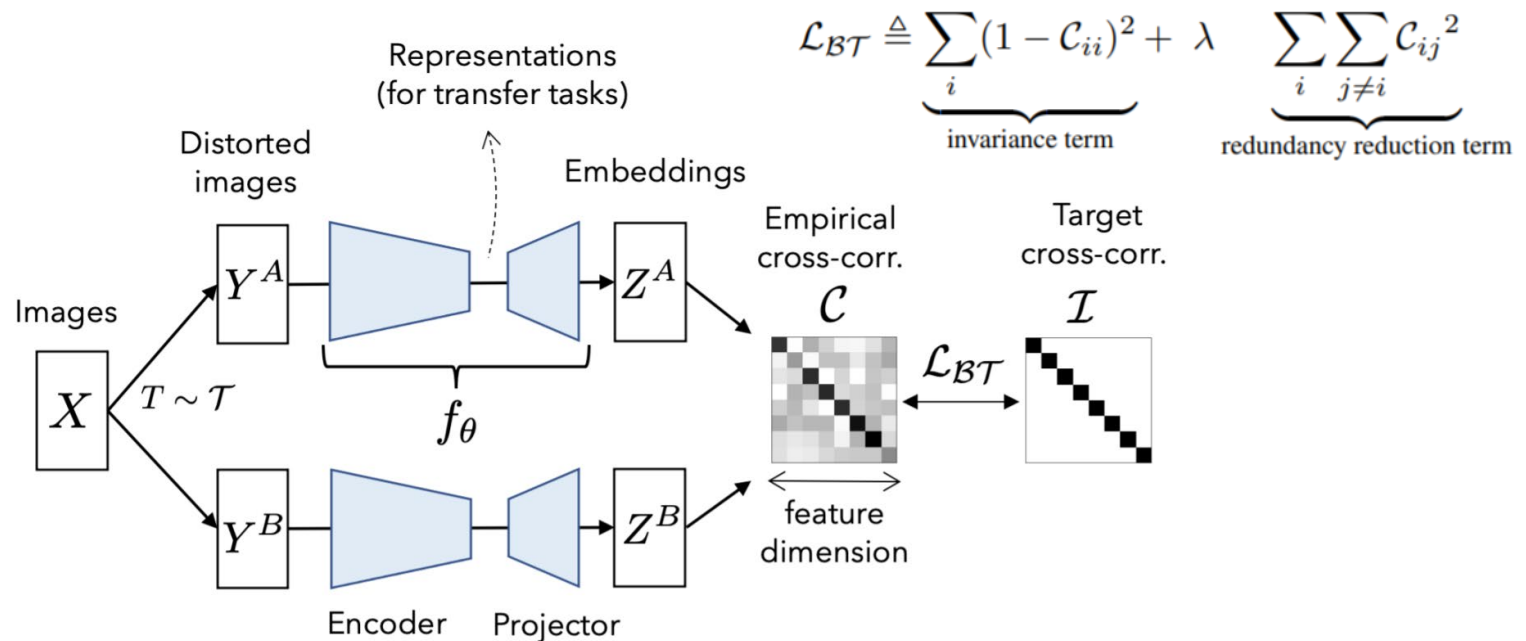
# BYOL

- No need of negative pairs
- Introduce the **predictor** for asymmetry to avoid collapse
- Exponential Moving Average (**EMA**)



# Barlow Twins

- Enforce **diversity** among **feature dimensions**
- Maximize diagonal terms and minimize off-diagonal ones
- No need of negative pairs, predictor network, gradient stopping or moving average techniques





# Barlow Twins

- Experiments on classification

Method	Top-1		Top-5	
	1%	10%	1%	10%
Supervised	25.4	56.4	48.4	80.4
PIRL	-	-	57.2	83.8
SIMCLR	48.3	65.6	75.5	87.8
BYOL	53.2	68.8	78.4	89.0
SwAV	53.9	<b>70.2</b>	78.5	<b>89.9</b>
BARLOW TWINS (ours)	<b>55.0</b>	69.7	<b>79.2</b>	89.3

# Barlow Twins

- Experiments on detection and segmentation

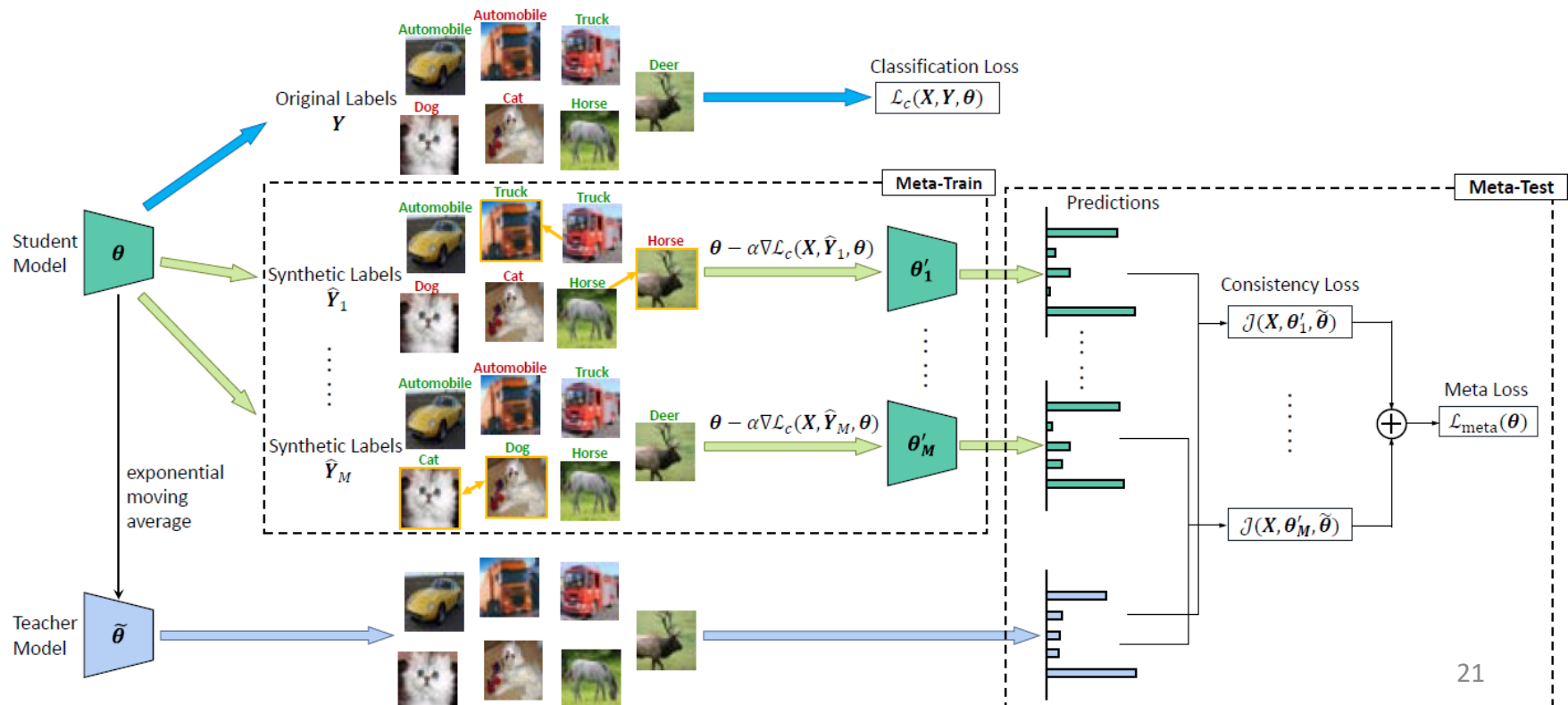
Method	VOC07+12 det			COCO det			COCO instance seg		
	$AP_{all}$	$AP_{50}$	$AP_{75}$	$AP^{bb}$	$AP_{50}^{bb}$	$AP_{75}^{bb}$	$AP^{mk}$	$AP_{50}^{mk}$	$AP_{75}^{mk}$
Sup.	53.5	81.3	58.8	38.2	58.2	41.2	33.3	54.7	35.2
MoCo-v2	<b>57.4</b>	82.5	<b>64.0</b>	<b>39.3</b>	58.9	<b>42.5</b>	<b>34.4</b>	55.8	36.5
SwAV	56.1	<b>82.6</b>	62.7	38.4	58.6	41.3	33.8	55.2	35.9
SimSiam	57	82.4	63.7	39.2	<b>59.3</b>	42.1	<b>34.4</b>	<b>56.0</b>	<b>36.7</b>
BT (ours)	56.8	<b>82.6</b>	63.4	39.2	59.0	<b>42.5</b>	34.3	<b>56.0</b>	36.5

# SSL Beyond Image Data

- What about videos?



- What about noisy data? J. Li et al., Learning to Learn from Noisy Labeled Data, CVPR 2019



# What's Next?

- Self-Supervised Learning (SSL)
  - Pretext Tasks vs. Contrastive Learning
  - SSL Beyond Images
- Invited Talk
  - 10 Secrets You Need to Know About Software Engineering & Career Planning



Would you expect an Olympic athlete to compete for Gold medal without any preparation or practice? Of course not!

Career Planning is about choices, yet it is extremely difficult to make decisions about your future when you have been doing the same thing as everyone else for the past two decades. In this talk, we will discuss some of the fundamentals that are crucial to grooming a customized career path that no employers would ever tell you. We will discuss tactics and strategies to ace the interview for the software engineering space, as well as the hiring trend to keep up-to-date demanded skills.



**Linda Huang**  
Practice Leader  
Paul Wright Taiwan Limited

I specialize in executive search for c-level suites, and am the Practice Leader for SCM and Engineering recruiting at Paul Wright Taiwan. Holder of a Bachelor's degree in architecture from CMU, went onto studying for 2 more master's degrees at CMU and the University of Oxford in England, focusing on Urban Design and Corporate Social Responsibility. I have a genuine interest in improving humanity in general, from the time I studied architecture to my time now as a recruitment consultant/headhunter, it is "people" that's the center of my focus and fuels my passion for my work and life.

### 10 Secrets You Need To Know About Software Engineering & Career Planning

**Tuesday December 6<sup>th</sup> / BL-112 / 10:10am - 12pm**  
(Host: Prof. Frank Wang)

Do you wish to become the CTO of a corporation? Do you want to make more money than your colleagues? Are you looking for opportunities to set up your own startup? Or you just want to find a place to chill for the rest of your life? All your questions will be answered to crush competitors and land your dream jobs. Let's start building your own roller-coaster ride.



**Sharine Chen**  
Associate Director  
Paul Wright Taiwan Limited

I am the Associate Director in Paul Wright Taiwan, specializing in mid-senior, executive-level search within the software engineering space. Having delivered numerous complex search mandates with domestic software start-ups to global organizations, I have an extensive network of candidates with skills across cutting-edge digital, FinTech & AI, data science, and mobile technology. Prior to a career in recruitment, I was working in account management roles in a Fortune 500 American IT service.



PAUL WRIGHT  
國立中央大學  
Graduate Institute of Communication Engineering