ConCL: Concept Contrastive Learning for Dense Prediction Pre-training in Pathology Images

Jiawei Yang[1,2,*], Hanbo Chen[1], Yuan Liang[2], Junzhou Huang[3], Lei He[2], and Jianhua Yao[1]

[1] Tencent AI Lab [2] University of California, Los Angeles [3] University of Texas at Arlington

ECCV TEL AVIV 2022

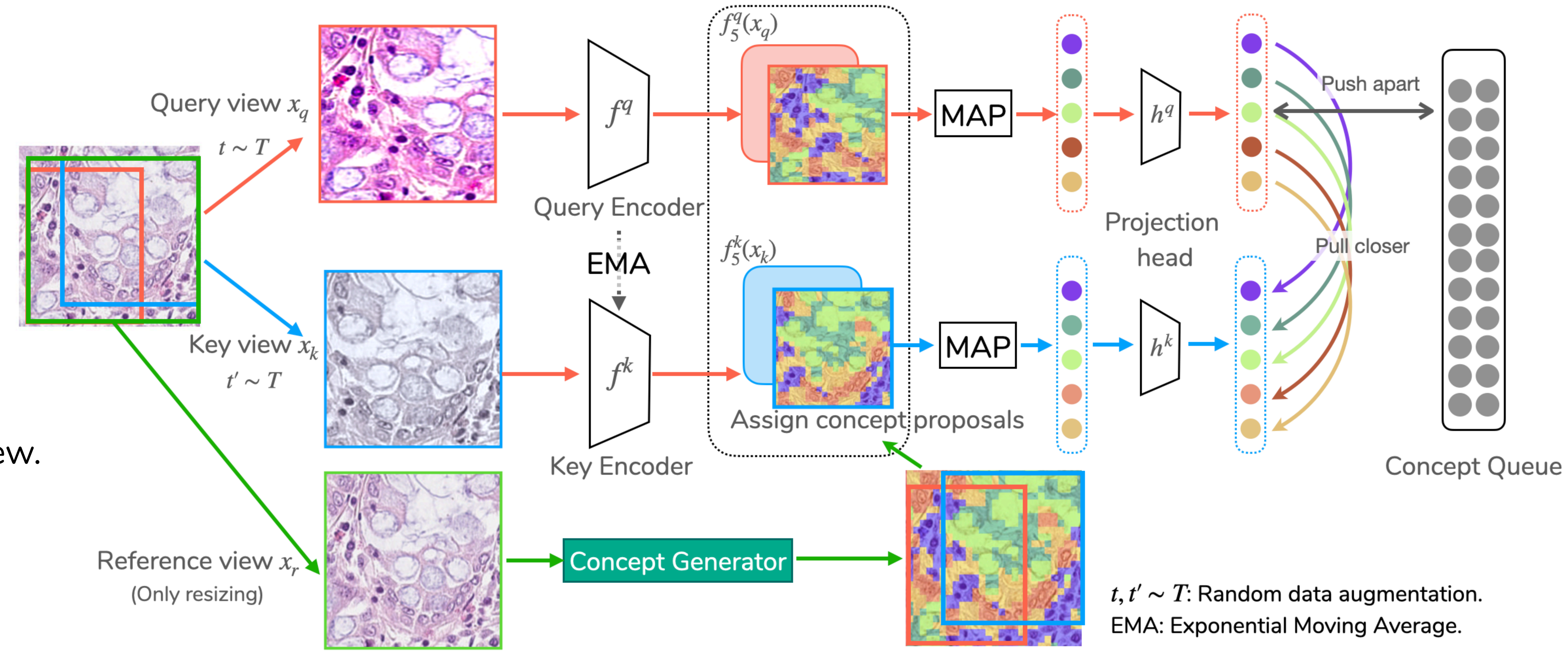
Introduction

- **Goals:**
  - Reduce the annotation burden for dense tasks in pathology images
  - Design a framework to pre-train a better backbone for dense tasks.
  - Benchmark existing arts in natural image for pathology images.
- **Motivation:**
  - Matched dense contrasting is extremely beneficial.
  - The plain contrastive learning framework can find semantically meaningful “concepts”.
  - Can we enhance it via bootstrapping?

Methods

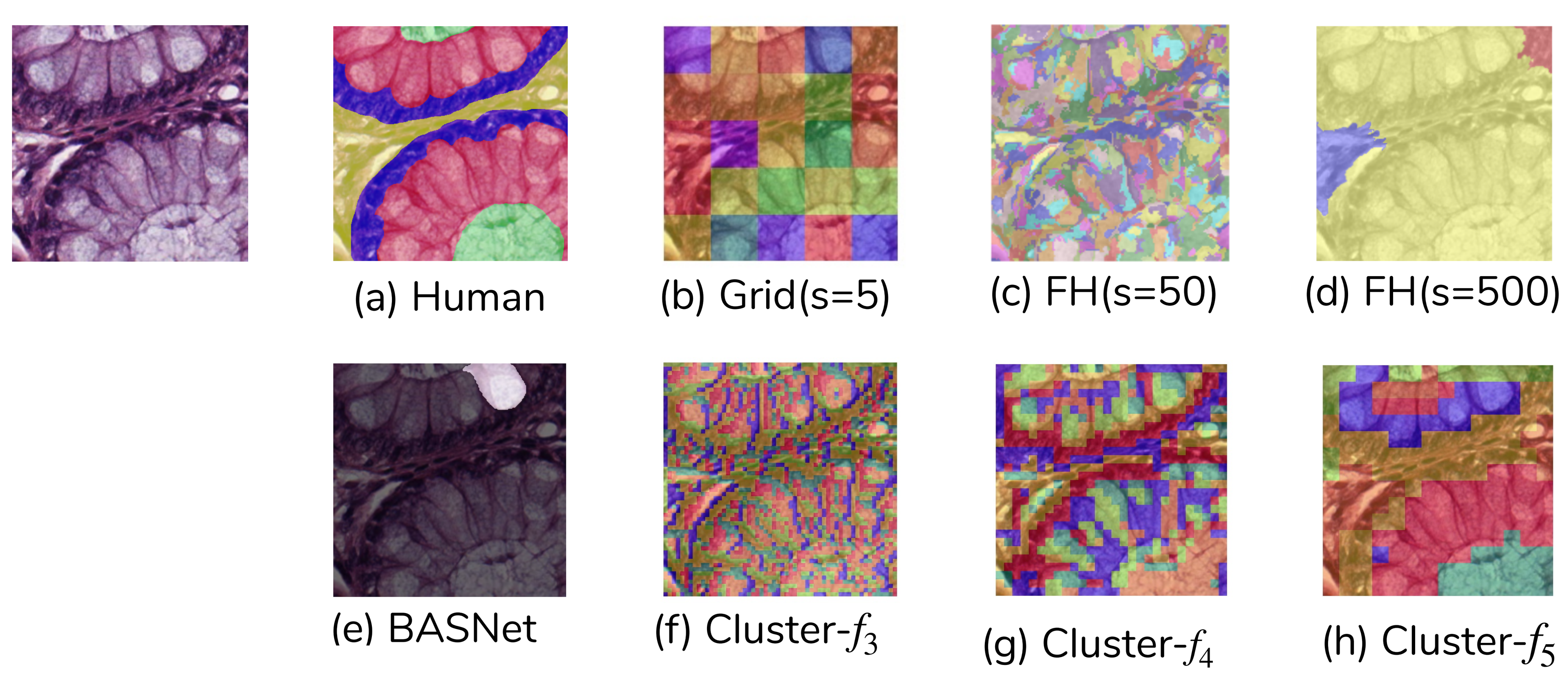
Pipeline Overview

- **Instance-level contrastive learning:**
  - Generate two different views from an image
  - Contrast views (attract pos. & repulse neg.)
- **Concept-level contrastive learning:**
  - Generate concepts from a reference view.
  - Contrast concepts
- **What differs is how concept is generated**



Concept Generator

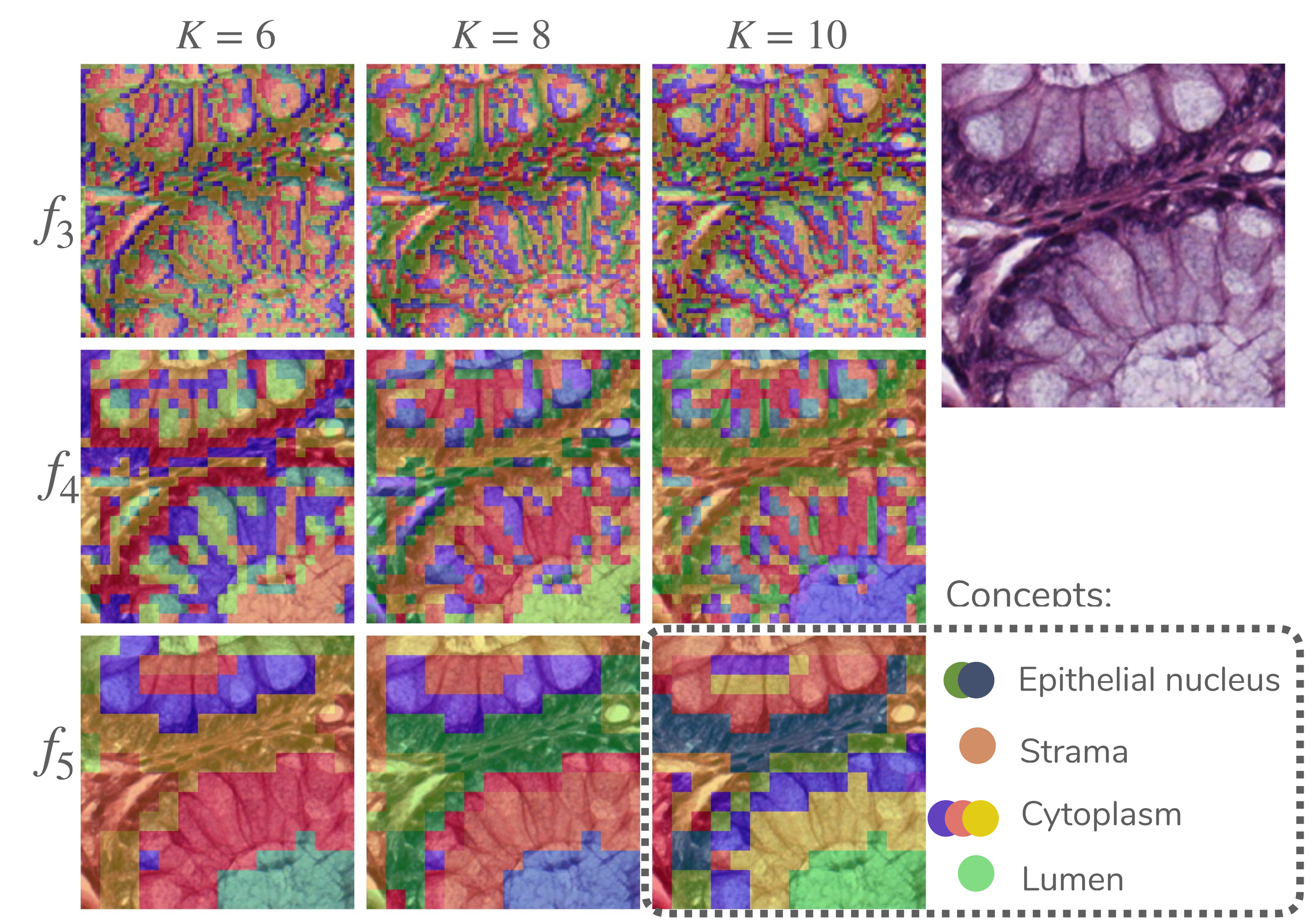
- **Different concepts we can use (inspired by DetCon, a DeepMind work):**
  - (a) Human annotation (not available)
  - (b) Grid-level concepts (not accurate)
  - (c-d) Graph-based segmentation (applicable to pathology images or not?)
  - (e) saliency detection model (trained on tailored natural image datasets)
  - (f-h) clustering feature arrays (ours)

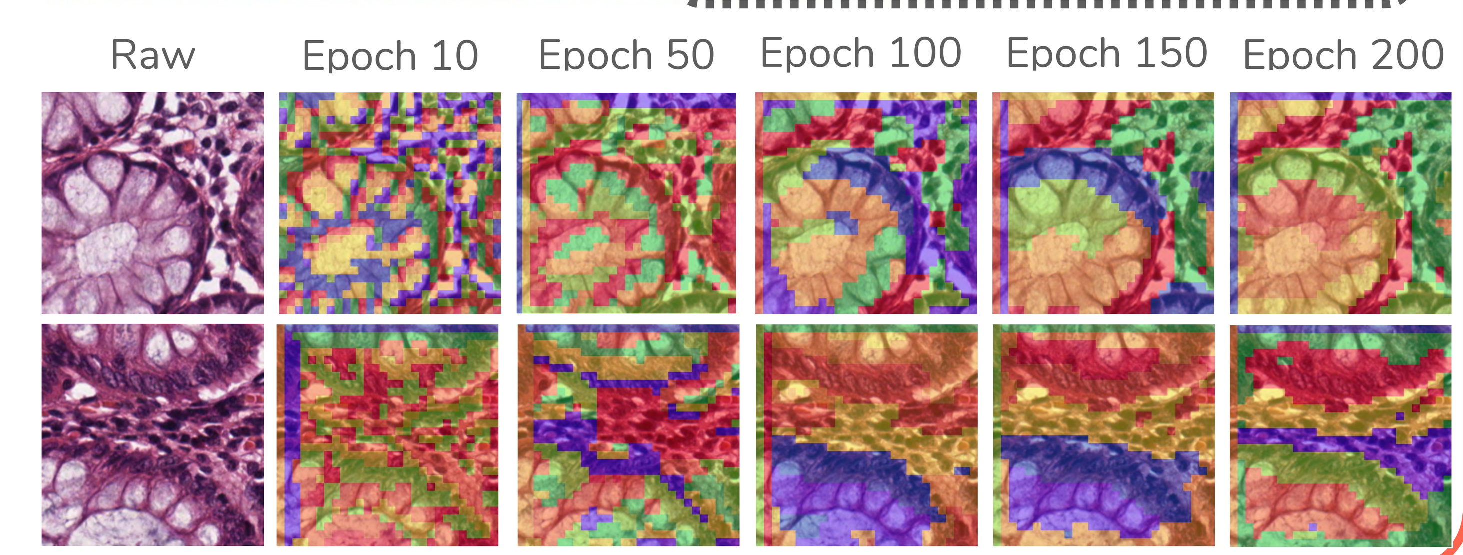


Main Results

Category	Methods	GlaS Detect AP^{bb}	GlaS Detect AP^{bb}_{75}	GlaS Segment AP^{mk}	GlaS Segment AP^{mk}_{75}	CRAG Detect AP^{bb}	CRAG Detect AP^{bb}_{75}	CRAG Segment AP^{mk}	CRAG Segment AP^{mk}_{75}
Baselines	Rand. Init.	49.8	57.3	52.1	60.7	51.1	57.0	50.6	57.3
	Supervised	50.2	56.9	53.2	62.1	49.2	55.2	49.4	55.0
Sec. 4.1 Prior SSL arts	SimCLR[4]	50.7	56.9	53.6	62.7	49.2	54.8	49.1	54.7
	BYOL[12]	50.9	57.7	53.9	62.6	49.9	55.8	49.3	55.3
	PCL-v2† [22]	49.4	55.9	51.9	61.0	51.0	56.6	50.5	56.7
	MoCo-v1[14]	50.0	56.2	52.1	59.9	47.2	51.1	47.5	52.0
	MoCo-v2[5]	52.3	60.0	55.3	65.0	50.0	55.7	50.3	56.8
	DenseCL[36]	53.9	62.0	56.5	66.2	52.3	58.2	52.2	59.8
Our differently instantiated ConCLs:									
Sec. 4.2 Grid concepts	(1) g-ConCL(s=3)	54.9	64.1	57.1	66.3	55.4	62.3	54.4	62.0
	(2) g-ConCL(s=5)	55.4	65.2	57.4	67.2	55.5	62.7	54.6	62.2
	(3) g-ConCL(s=7)	54.9	63.8	57.0	66.5	55.3	62.5	54.7	62.6
Sec. 4.3 Natural-image priors concepts	(4) fh-ConCL(s=50)	55.8	65.6	58.3	68.8	54.8	60.7	54.1	60.7
	(5) fh-ConCL(s=500)	56.2	65.9	57.7	67.9	54.7	61.9	53.8	60.5
	(6) bas-ConCL	56.1	66.1	58.1	68.1	54.2	61.1	53.4	60.8
Sec. 4.4 Bootstrapped concepts	(7) b-ConCL(f_4)	**56.8**	**66.2**	**58.7**	**68.9**	55.1	62.2	54.1	61.4
	(8) b-ConCL(f_5)	56.1	65.6	57.8	67.7	**56.5**	**63.3**	**55.3**	**62.9**

Table 1: **Main results of object detection and instance segmentation.** AP^{bb}: bounding box mAP, AP^{mk}: mask mAP.





- **Basic loop in ConCL:**
  - Learn to map similar concepts together via CL.
  - Cluster similar regions to generate meaningful concepts.
  - Contrast to improve & go back to the beginning.