



یادگیری خود-نظارت در بینایی ماشین

ارائه دهنده: شروین حالت

استاد درس: دکتر صفا بخش

اساتید راهنما: دکتر رحمتی، دکتر ناظر فرد

شهریور ماه 1400

فهرست

مقدمه

مرور
سوابق

یادگیری
خود-نظارت

روش‌های
بهبود

ارزیابی و
جمع‌بندی

منابع و
مراجع

مقدمه

مرور سوابق

یادگیری خود-نظارت

روش‌های بهبود

ارزیابی و جمع‌بندی

منابع و مراجع

مقدمه

انگیزه

- "اگر هوشمندی یک کیک باشد، **بدنه ی آن یادگیری خود-نظارت است**، روکش و تزیینات آن یادگیری با-نظارت و گیلای روی این کیک نیز یادگیری تقویتی است!" —یان لیکان (رییس هوش مصنوعی فیسبوک)

یادگیری تقویتی

یادگیری با-نظارت

یادگیری خود-نظارت



هدف

- مساله آموزش شبکه‌ی‌های عصبی عمیق
- معرفی شبکه‌های گسترده و عمیق تر
- تولید مجموعه دادگان گسترده تر

مجموعه دادگان برچسب-دار

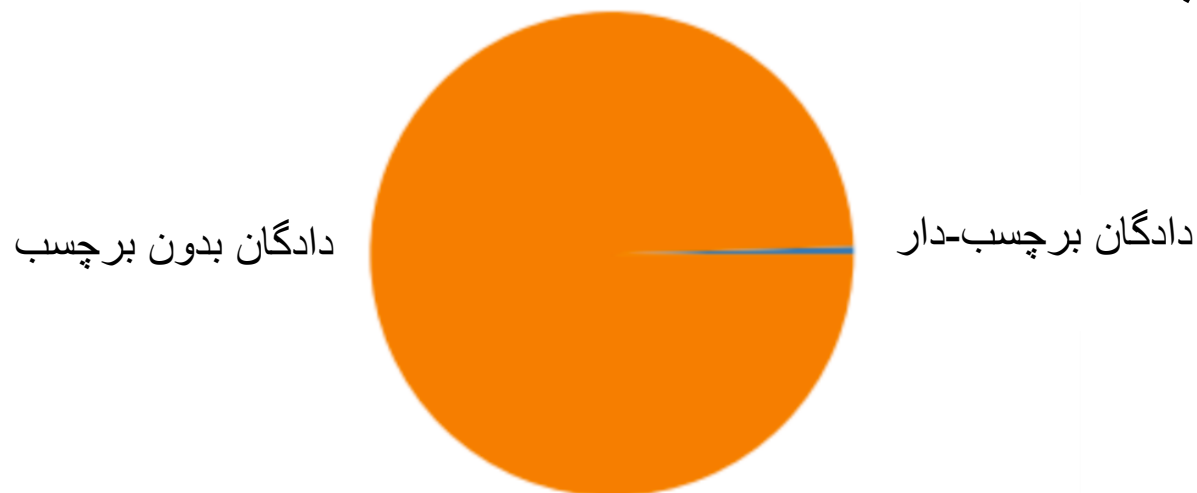


شبکه اصلی

وظیفه
اصلی

آموزش شبکه به روش با-نظارت

- مساله آماده سازی دادگان مناسب برای آموزش شبکه‌های عمیق
 - نیاز به دادگان فراوان
 - دشواری و هزینه‌های برچسب زنی به دادگان
 - بهره‌برداری از دادگان گسترده‌ی بدون برچسب
 - دادگان مصنوعی
 - دادگان موجود در وب



مقایسه حجم دادگان برچسب-دار و بدون برچسب

هدف

- راه حل آموزش شبکه‌ی‌های عصبی عمیق
- بکارگیری روش‌های یادگیری بدون نظارت

مجموعه دادگان بدون برچسب



شبکه پیش-آموزش



وظیفه
پوششی

مجموعه دادگان برچسب-دار



یادگیری
انتقالی



وظیفه
اصلی

شبکه اصلی

آموزش شبکه به روش یادگیری انتقالی

کاربردها

مرور
سوابق

یادگیری
خود-نظارت

روش‌های
بهبود

ارزیابی و
جمع‌بندی

منابع و
مراجع

- بینایی ماشین

- تمامی وظایف اصلی و فرعی

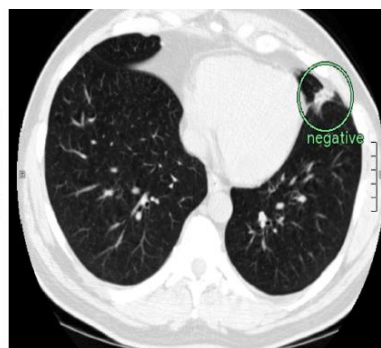
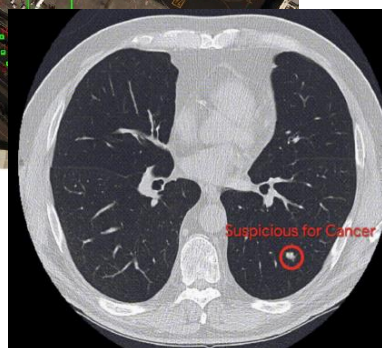
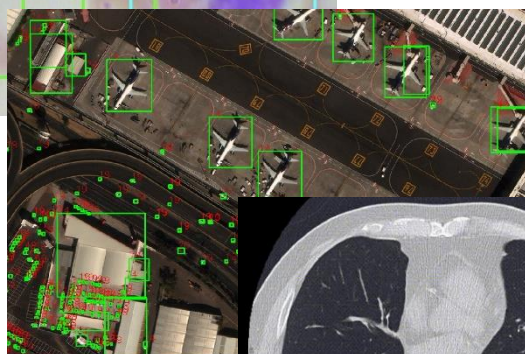
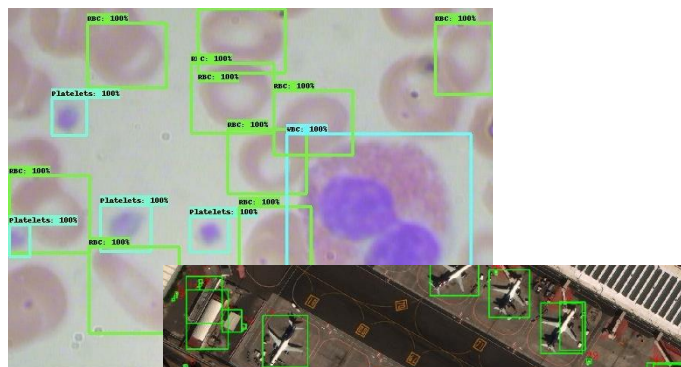
- حوزه‌هایی از بینایی ماشین که داده‌های آموزش کمی دارند، مانند:

- تصاویر پزشکی

- تصاویر ماهواره‌ای

- تصاویر میکروسکوپی

- ...



- پردازش زبان و سیگنال‌ها

مقدمه

مرور
سوابق

یادگیری
خود-نظارت

روش‌های
بهبود

ارزیابی و
جمع‌بندی

منابع و
مراجع

مرور سوابق

تاریخچه

1984

اولین مقاله که کلمه خود-نظارت را بکار برد
و در حوزه‌ی یادگیری تقویتی بود

1994

بکارگیری روش خود-نظارتی وجوه-متقابل
برای دسته بندی

2015
2016
2017

بکارگیری روش خود-نظارتی در پردازش
زبان و مشاهده‌ی نتایج مطلوب

2018

بهبود روش‌های خود-نظارتی در بینایی
ماشین و معرفی خلاقیت های جدید و
روش‌های خاص-منظوره

2020

معرفی روش‌هایی که به کمک یادگیری تبارینی

مقدمه

مرور
سوابق

یادگیری
خود-نظارت

روش‌های
بهبود

ارزیابی و
جمع‌بندی

منابع و
مراجع

یادگیری خود-نظارت

دسته بندی روش های یادگیری خود-نظارت

مقدمه

مرور
سوابق

یادگیری
خود-نظارت

روش های
بهبود

ارزیابی و
جمع بندی

منابع و
مراجع

- روش های عام-منظوره
- یادگیری انتقالی (Transfer Learning)

- روش های خاص-منظوره
- یادگیری چند-وظیفه ای (Multi-task Learning)
- بهبود مدل های موجود
- سایر روش ها

ساختار کلی یادگیری خود-نظارت (روش‌های عام-منظوره)

مقدمه

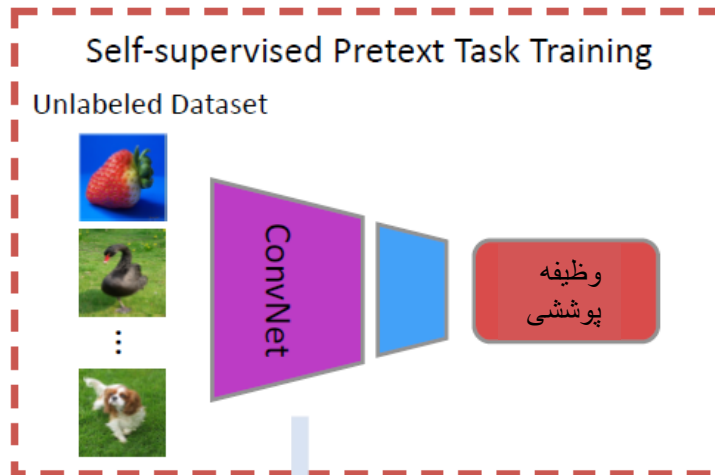
مرور
سوابق

یادگیری
خود-نظارت

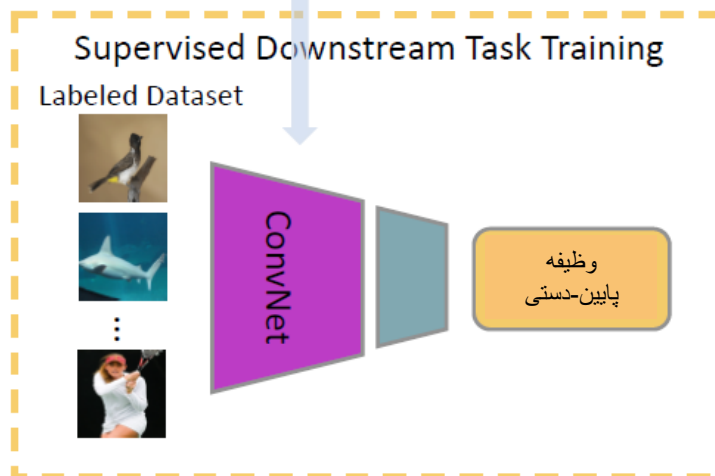
روش‌های
بهبود

ارزیابی و
جمع‌بندی

منابع و
مراجع



انتقال ویژگی



ساختار روش عام-منظوره یادگیری خود-نظارت

- آموزش شبکه وظیفه‌ی پوششی به روش خود-نظارت

$$loss(D) = \min_{\theta} \frac{1}{N} \sum_{i=1}^N loss(X_i, P_i)$$

- آموزش شبکه وظیفه‌ی پایین دستی به روش با-نظارت

$$loss(D) = \min_{\theta} \frac{1}{N} \sum_{i=1}^N loss(X_i, Y_i)$$

دسته بندی وظایف پوششی

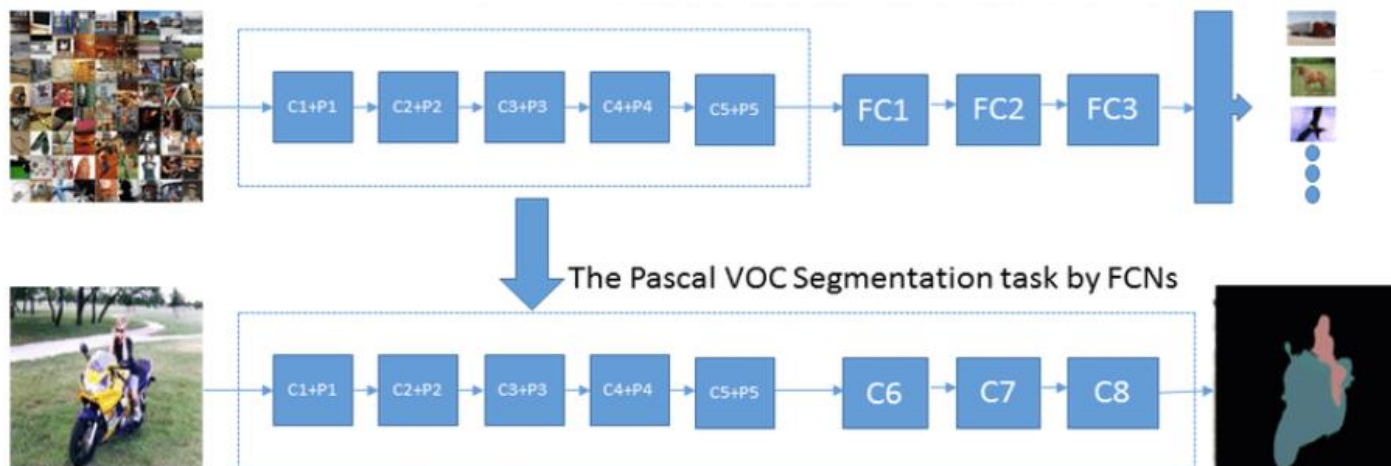
- وظایف پوششی ساخت-پایه
- نقاشی کردن، رنگ آمیزی و روش‌های مبتنی بر GAN
- وظایف پوششی زمینه-پایه
- پازل جیگسا، چرخش و پیشبینی زمینه
- وظایف پوششی معنایی-پایه
- وظایف پوششی وجوه متقابل-پایه

ارزیابی کیفی روش‌های یادگیری خود-نظارت

- روش‌های عام-منظوره
- مصوّر سازی هسته
- پیشبینی چرخش
- مصوّر سازی نقشه‌های ویژگی‌ها
- پیشبینی چرخش
- بازیابی نزدیک ترین همسایه
- نقاشی کردن، یادگیری جایگشت بصری

ارزیابی کمی روش‌های یادگیری خود-نظارت

- روش‌های عام-منظوره
- ارزیابی توسط وظایف پایین دستی اصلی (به کمک یادگیری انتقالی)



یادگیری انتقالی در روش‌های عام-منظوره

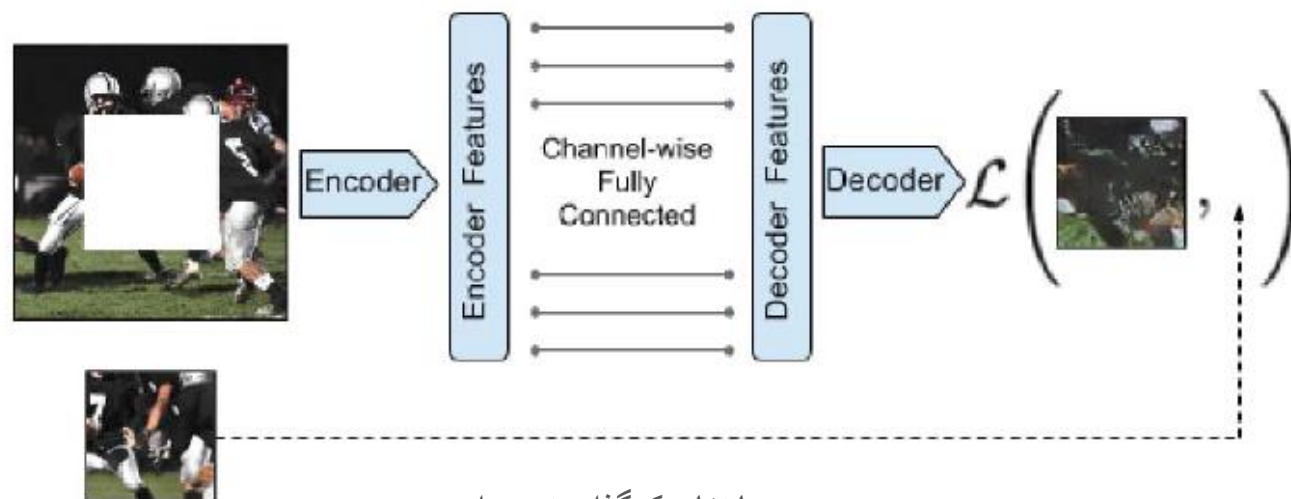
- روش‌های خاص-منظوره
- مقایسه با عملکرد روش‌های پیشین مرتبط

وظیفه پوششی نقاشی کردن

- نمونه ای از آموزش وظیفه‌ی پوششی



وظیفه پوششی نقاشی کردن



ساختار کدگذار زمینه‌ای

- کدگذار زمینه‌ای

چالش‌های وظیفه پوششی نقاشی کردن

- اهمیت تابع هزینه در کدگذار زمینه‌ای



(a) Input context

(b) Human artist

(c) Context Encoder
(L2 loss)(d) Context Encoder
(L2 + Adversarial loss)

مقایسه نقاشی کردن با توابع هزینه متفاوت

$$\mathcal{L}_{rec}(x) = \| \hat{M} \odot (x - F((1 - \hat{M}) \odot x)) \|_2$$

$$\mathcal{L}_{adv} = \max_D \mathbb{E}_{x \in \mathcal{X}} [\log(D(x)) + \log(1 - D(F((1 - \hat{M}) \odot x)))]$$

$$\mathcal{L} = \lambda_{rec} \mathcal{L}_{rec} + \lambda_{adv} \mathcal{L}_{adv}$$

چالش‌های وظیفه پوششی نقاشی کردن

مقدمه

مرور
سوابق

یادگیری
خود-نظارت

روش‌های
بهبود

ارزیابی و
جمع‌بندی

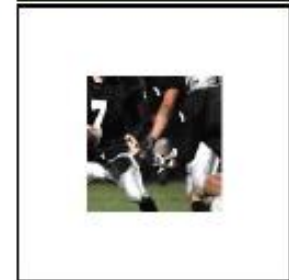
منابع و
مراجع

- نحوه ی ماسک کردن تصاویر

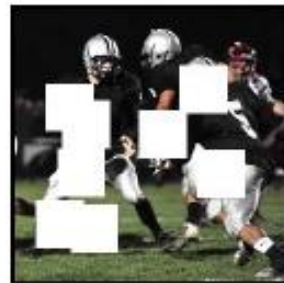
- ناحیه مرکزی

- بلوک تصادفی

- ناحیه تصادفی



(a) Central region



(b) Random block

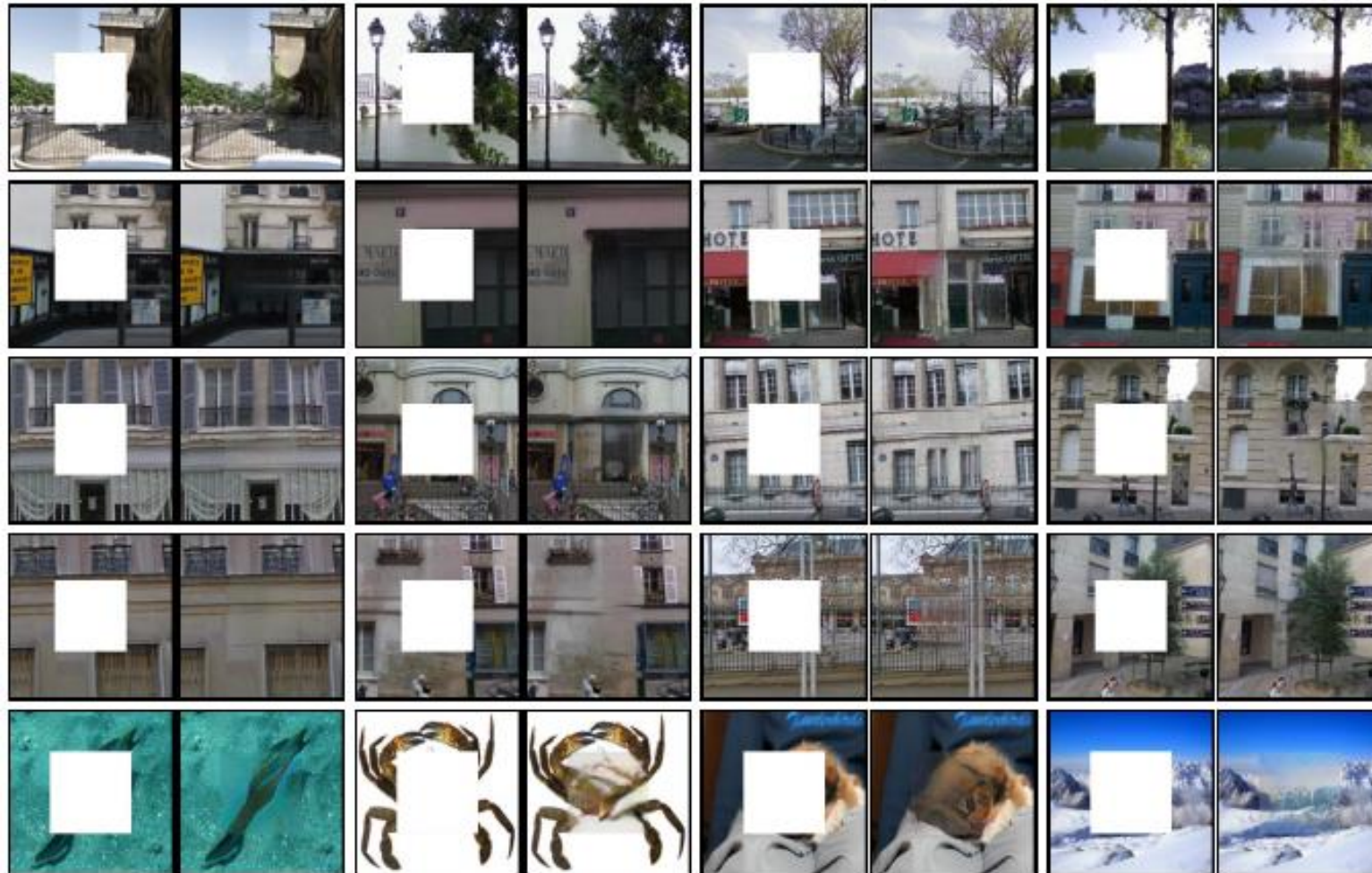


(c) Random region

روش‌های ماسک کردن های متفاوت

ارزیابی کیفی وظیفه پوششی نقاشی کردن

- نقاشی کردن معنایی



ارزیابی کیفی وظیفه پوششی نقاشی کردن

معرفی اجمالی برخی از وظایف پوششی

مقدمه

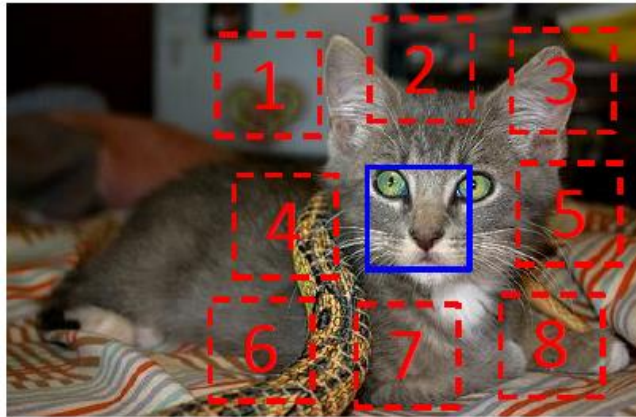
مرور
سوابق

یادگیری
خود-نظارت

روش‌های
بهبود

ارزیابی و
جمع‌بندی

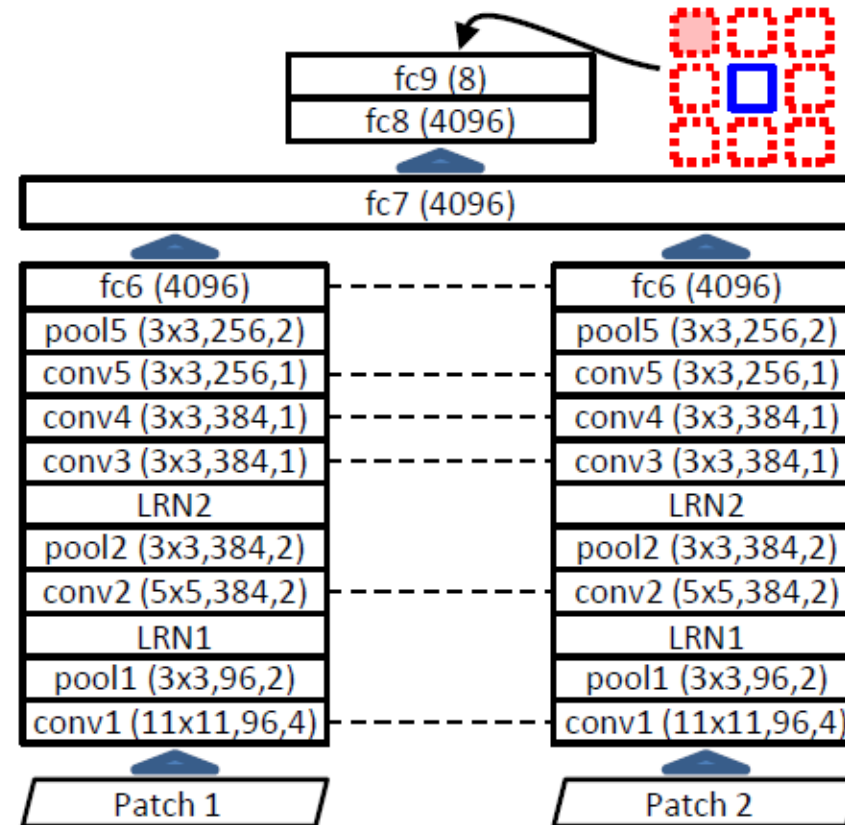
منابع و
مراجع



$$X = \left(\begin{array}{c} \text{cat face} \\ \text{cat ear} \end{array} \right); Y = 3$$

معرفی وظیفه پوششی پیشبینی زمینه

پیشبینی زمینه



ساختار شبکه وظیفه پوششی پیشبینی زمینه

معرفی اجمالی برخی از وظایف پوششی

مقدمه

مرور
سوابق

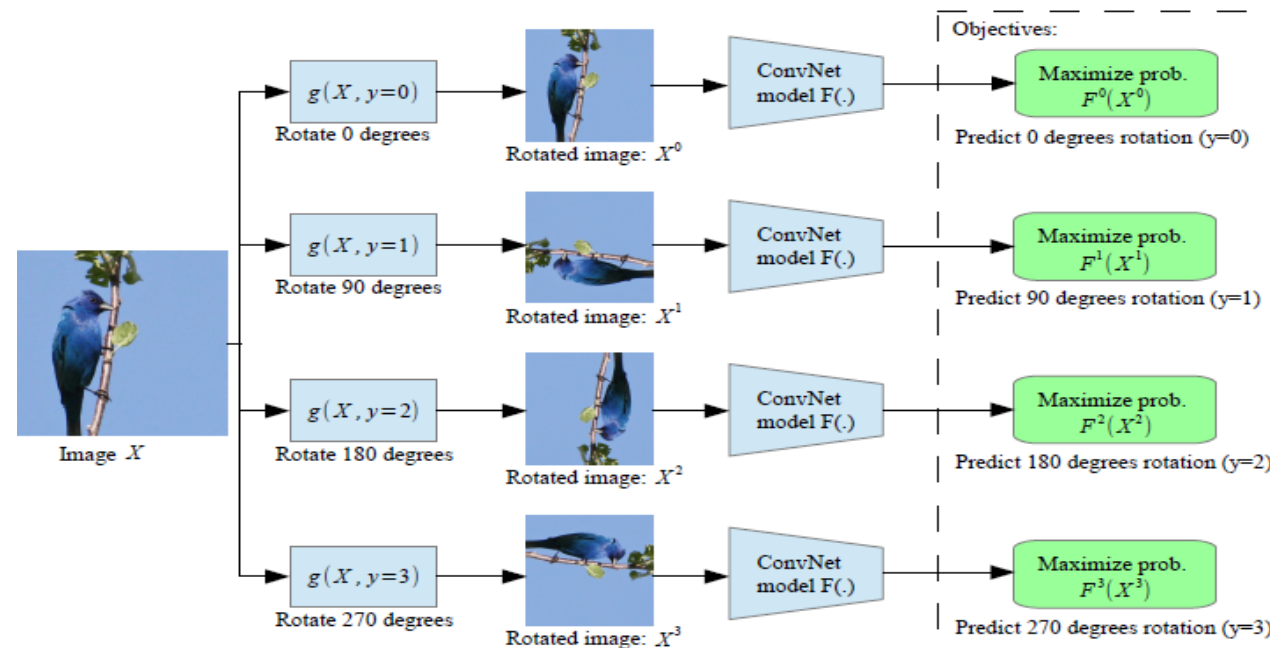
یادگیری
خود-نظارت

روش‌های
بهبود

ارزیابی و
جمع‌بندی

منابع و
مراجع

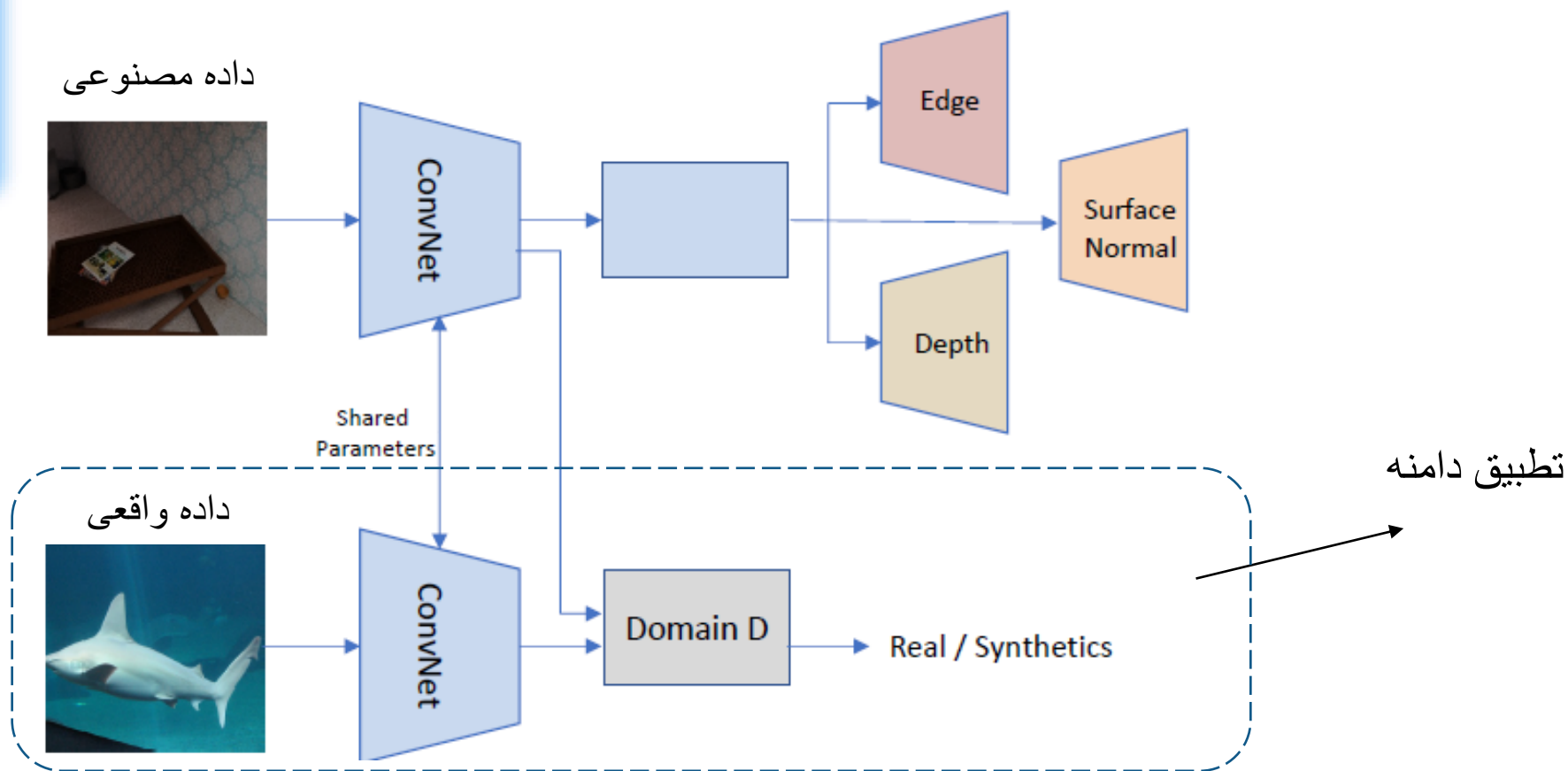
• پیشبینی چرخش



شبکه وظیفه پوششی پیشبینی چرخش

چالش‌ها

- چالش‌های موجود در یادگیری خود-نظارت
- تطبیق دامنه
- به روش یادگیری رقابتی



مثالی از نحوه تطبیق دامنه در یادگیری خود-نظارت

چالش‌ها

- چالش‌های موجود در یادگیری خود-نظارت
- عملکرد ضعیف‌تر نسبت به یادگیری با-نظارت
 - معرفی روش‌های خاص-منظوره
 - معرفی روش‌های چند-وظیفه‌ای
 - بکارگیری وظایف پوششی تبیینی
- عدم امکان بکارگیری ویژگی‌های استخراج شده در وظایف خاص
 - ویژگی‌های سطح بالا به جای ویژگی‌های سطح پایین

مقدمه

مرور
سوابق

یادگیری
خود-نظارت

روش‌های
بهبود

ارزیابی و
جمع‌بندی

منابع و
مراجع

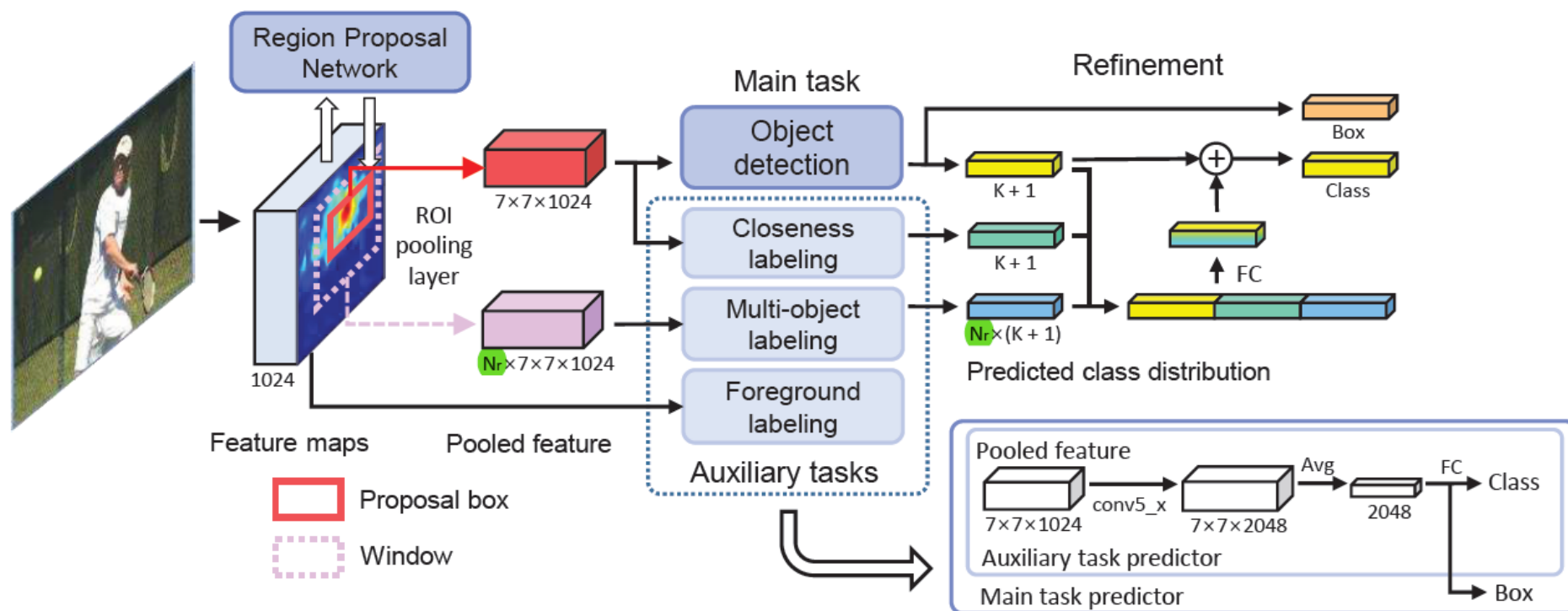
روش‌های بهبود

روش‌های بهبود

- اما به منظور رفع چالش‌های موجود در روش‌های عام-منظوره چه می‌توان کرد؟
 - معرفی وظایف پوششی کارآمدتر
 - وظایف پوششی به کمک یادگیری تبیینی
 - اعمال خلاقیت‌های نوین در مدل‌های موجود
 - نقاط سوپر، حذف نویز خود-نظارتی (NAC)، ازدیاد ویژگی آشکارسازی اشیا برای تصاویر حجیم
 - بکارگیری یادگیری چند-وظیفه‌ای (MTL)

آشکار سازی اشیا به روش خود-نظارت چند-وظیفه‌ای

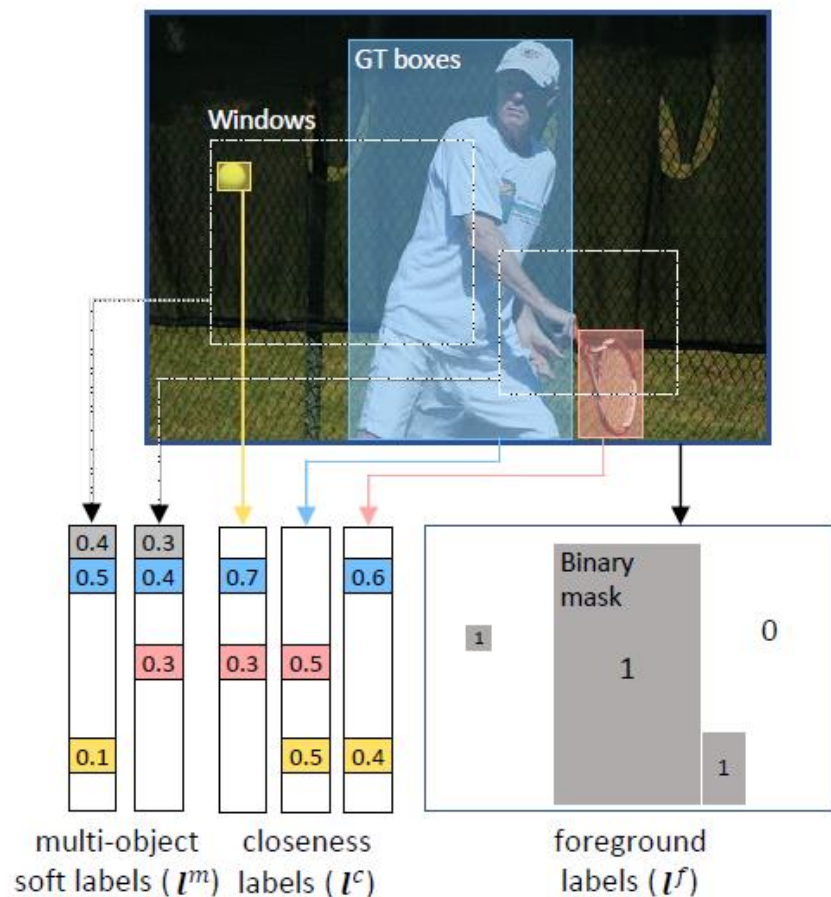
- بهبود آشکار سازی اشیا به کمک معرفی وظایف کمکی



معرفی شبکه آشکار سازی اشیا به روش خود-نظارت چند-وظیفه‌ای

آشکار سازی اشیا به روش خود-نظارت چند-وظیفه ای

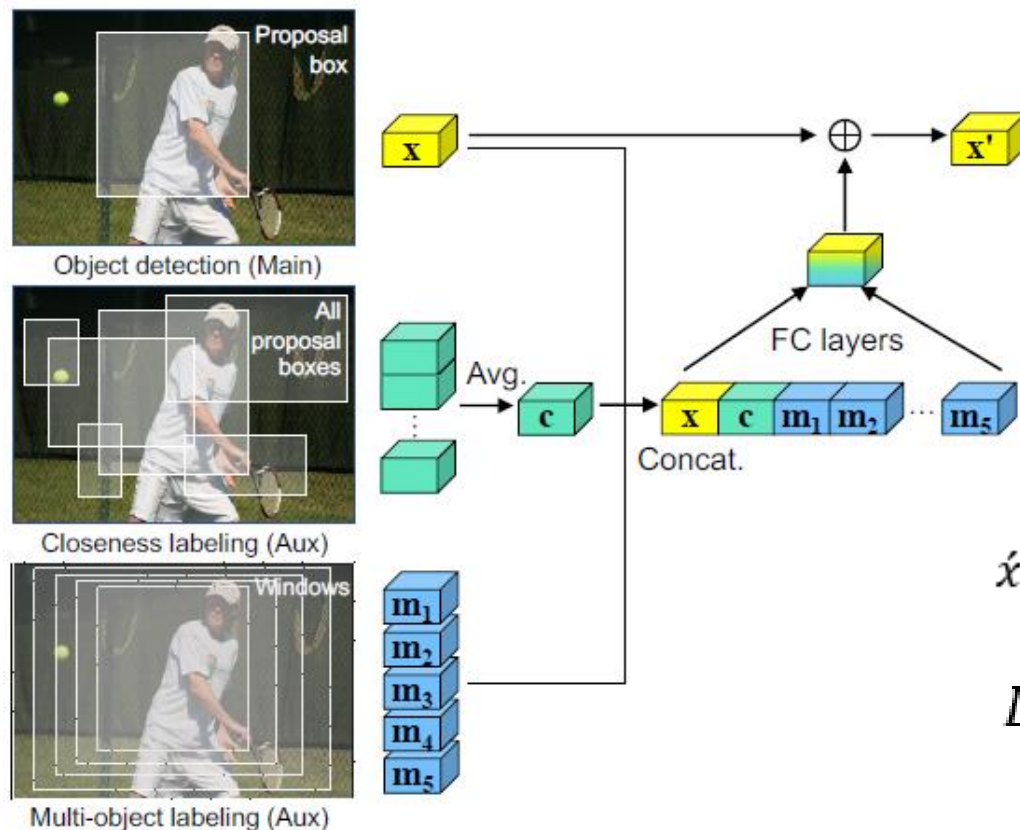
- وظایف کمکی معرفی شده:
 - برچسب زنی چند-شیء
 - برچسب زنی نزدیکی
 - برچسب زنی پیش‌زمینه



معرفی برچسب وظایف کمکی معرفی شده

آشکارسازی اشیا به روش خود-نظارت چند-وظیفه‌ای

- آموزش مدل پیشنهادی



خروجی‌های وظایف کمکی اعمال شده به مدل

$$\hat{x} = \text{refine}(x | c, m_1, \dots, m_n) = W_r [x, c, m_1, \dots, m_n] + x$$

$$L_{\{aux\}} = \lambda_{\{m\}} \cdot L_{\{m\}} + \lambda_{\{c\}} \cdot L_{\{c\}} + \lambda_{\{f\}} \cdot L_{\{f\}}$$

$$L_{\{total\}} = L_{\{aux\}} + L_{\{main\}} + \lambda_{\{r\}} \cdot L_{\{ref\}}$$

آشکارسازی اشیا به روش خود-نظارت چند-وظیفه‌ای

• ارزیابی کمی

مقایسه‌ی عملکرد مدل پیشنهادی در آشکارسازی اشیا با تنظیمات مختلف

Dataset	VOC			COCO	
Training	07	07+12		17 train	
Test	07	07	12	17 val	17 test-dev
Baseline	77.0	81.7	75.3	32.7	32.8
+ Task1	78.9	83.8	77.4	34.1	34.2
+ Task2	77.3	83.0	76.0	33.3	33.5
+ Task3	77.0	82.0	75.1	32.9	32.8
+ Task1,2	78.5	83.7	77.3	34.5	34.6
+ Task1,2,3	78.7	83.7	77.5	34.6	34.7

مقدمه

مرور
سوابق

یادگیری
خود-نظارت

روش‌های
بهبود

ارزیابی و
جمع‌بندی

منابع و
مراجع

ارزیابی و جمع‌بندی

ارزیابی

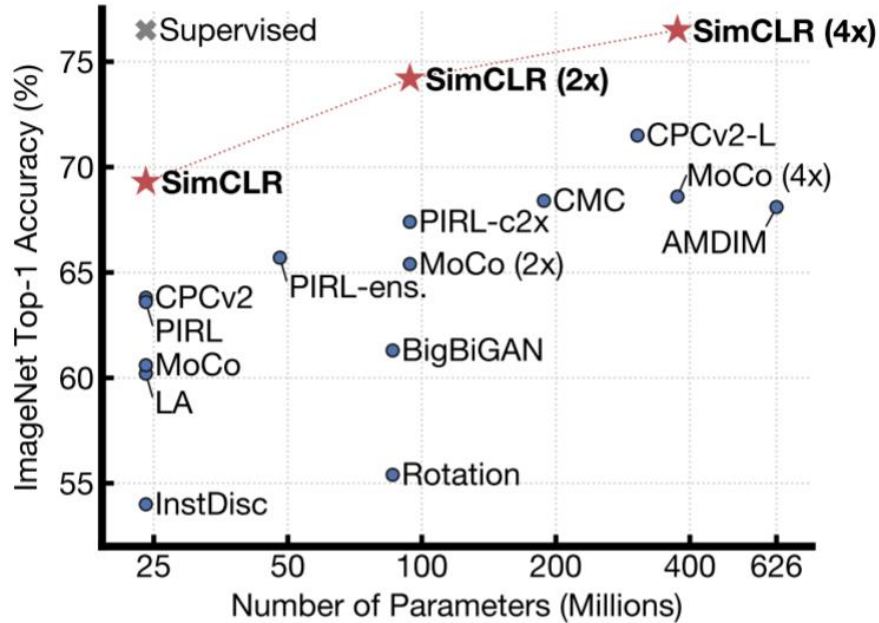
• مقایسه‌ی عملکرد روش‌های عام-منظوره یادگیری خود-نظارت

ارزیابی و مقایسه‌ی عملکرد یادگیری خود-نظارت و با نظارت در دسته بندی، قطعه بندی و آشکارسازی

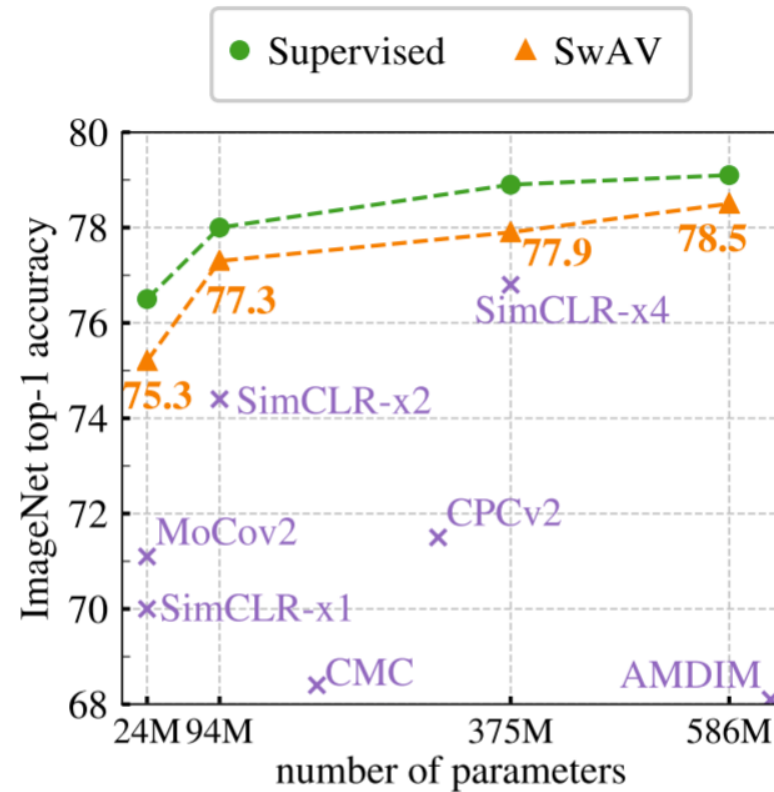
Method	Pretext Tasks	Classification	Detection	Segmentation
ImageNet Labels [8]	—	79.9	56.8	48.0
Random(Scratch) [8]	—	57.0	44.5	30.1
ContextEncoder [19]	Generation	56.5	44.5	29.7
BiGAN [122]	Generation	60.1	46.9	35.2
ColorfulColorization [18]	Generation	65.9	46.9	35.6
SplitBrain [42]	Generation	67.1	46.7	36.0
RankVideo [38]	Context	63.1	47.2	35.4 [†]
PredictNoise [46]	Context	65.3	49.4	37.1 [†]
JigsawPuzzle [20]	Context	67.6	53.2	37.6
ContextPrediction [41]	Context	65.3	51.1	—
Learning2Count [130]	Context	67.7	51.4	36.6
DeepClustering [44]	Context	73.7	55.4	45.1
WatchingVideo [81]	Free Semantic Label	61.0	52.2	—
CrossDomain [30]	Free Semantic Label	68.0	52.6	—
AmbientSound [154]	Cross Modal	61.3	—	—
TiedToEgoMotion [95]	Cross Modal	—	41.7	—
EgoMotion [94]	Cross Modal	54.2	43.9	—

ارزیابی

- بهبود وظایف پوششی با بکارگیری یادگیری تبارینی
- SwAV، SimCLR



مقایسه عملکرد روش SimCLR در مقایسه با
وظایف پوششی سابق

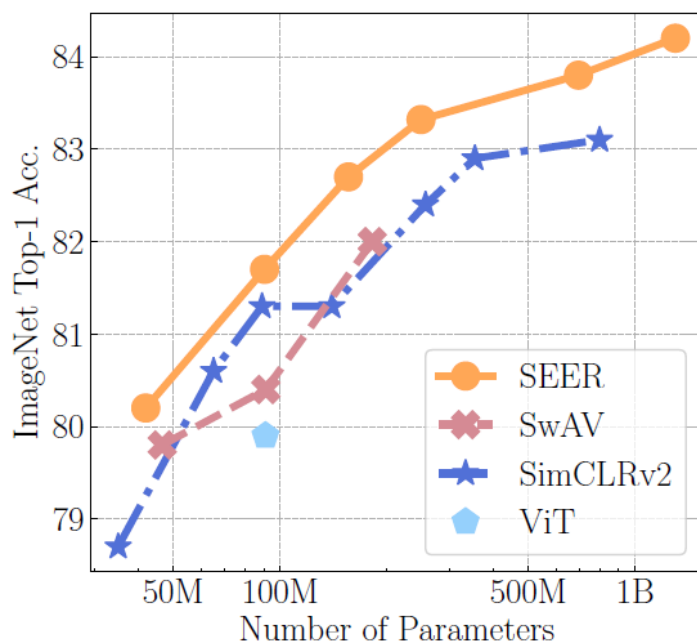


مقایسه عملکرد روش SwAV در مقایسه با
وظایف پوششی سابق

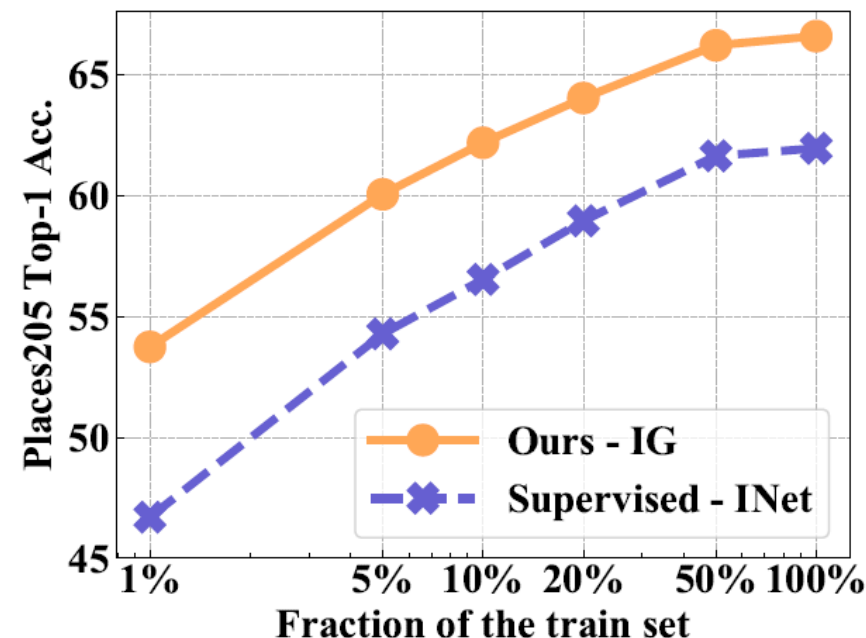
ارزیابی

SEER •

• عملکرد موفق روش خود-نظارت در دنیای واقعی



مقایسه عملکرد روش SEER در مقایسه با
وظایف پوششی سابق



مقایسه عملکرد روش SEER در
مقایسه با روش با-نظارت

جمع‌بندی

- مقدمه
- انگیزه، اهداف، کاربردها و مرور سوابق یادگیری خود-نظارت
- معرفی یادگیری خود-نظارت
- دسته‌بندی‌ها، معرفی ساختار، معرفی وظایف پوششی، نحوه‌ی ارزیابی و چالش‌های یادگیری خود-نظارت
- روش‌های بهبود
- روش‌های بهبود و معرفی یک نمونه از آن‌ها در یادگیری خود-نظارت
- ارزیابی یادگیری خود-نظارت

نتیجه‌گیری

- موفقیت‌های روز افزون یادگیری خود-نظارت
- اهمیت آن در کاهش هزینه‌ها
- زمینه‌ی فعالیت و نوآوری‌های گسترده

منابع و مراجع

- [1] Jing, Longlong and Yingli Tian. "Self-supervised Visual Feature Learning with Deep Neural Networks: A Survey." *IEEE transactions on pattern analysis and machine intelligence* (2020): n. pag.
- [2] Kolesnikov, Alexander, Xiaohua Zhai and L. Beyer. "Revisiting Self-Supervised Visual Representation Learning." *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2019): 1920-1929.
- [3] Pathak, Deepak, Philipp Krähenbühl, Jeff Donahue, Trevor Darrell and Alexei A. Efros. "Context Encoders: Feature Learning by Inpainting." *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2016): 2536-2544.
- [4] Gidaris, Spyros, Praveer Singh and N. Komodakis. "Unsupervised Representation Learning by Predicting Image Rotations." *ArXiv abs/1803.07728* (2018): n. pag.
- [5] Cruz, Rodrigo Santa, Basura Fernando, A. Cherian and Stephen Gould. "Visual Permutation Learning." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 41 (2019): 3100-3114.

منابع و مراجع

- [6] Doersch, Carl, A. Gupta and Alexei A. Efros. "Unsupervised Visual Representation Learning by Context Prediction." *2015 IEEE International Conference on Computer Vision (ICCV)* (2015): 1422-1430.
- [7] Noroozi, M., H. Pirsiavash and P. Favaro. "Representation Learning by Learning to Count." *2017 IEEE International Conference on Computer Vision (ICCV)* (2017): 5899-5907.
- [8] Liu, Xialei, Joost van de Weijer and Andrew D. Bagdanov. "Exploiting Unlabeled Data in CNNs by Self-Supervised Learning to Rank." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 41 (2019): 1862-1878.
- [9] Dosovitskiy, A., P. Fischer, Jost Tobias Springenberg, Martin A. Riedmiller and T. Brox. "Discriminative Unsupervised Feature Learning with Exemplar Convolutional Neural Networks." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 38 (2016): 1734-1747.
- [10] Xu, J., Yuan Huang, Ming-Ming Cheng, Li Liu, F. Zhu, Zhou Xu and L. Shao. "Noisy-as-Clean: Learning Self-Supervised Denoising from Corrupted Image." *IEEE Transactions on Image Processing* 29 (2020): 9316-9329.

منابع و مراجع

- [11] DeTone, Daniel, Tomasz Malisiewicz and Andrew Rabinovich. "SuperPoint: Self-Supervised Interest Point Detection and Description." *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (2018): 337-33712.
- [12] Pan, Xingjia, Fan Tang, Weiming Dong, Yang Gu, Zhichao Song, Yiping Meng, P. Xu, O. Deussen and Changsheng Xu. "Self-Supervised Feature Augmentation for Large Image Object Detection." *IEEE Transactions on Image Processing* 29 (2020): 6745-6758.
- [13] Lee, Wonhee, Joonil Na and Gunhee Kim. "Multi-Task Self-Supervised Object Detection via Recycling of Bounding Box Annotations." *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2019): 4979-4988.
- [14] Ren, Zhongzheng and Yong Jae Lee. "Cross-Domain Self-Supervised Multi-task Feature Learning Using Synthetic Imagery." *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2018): 762-771.

منابع و مراجع

[15] Goyal, Priya, Mathilde Caron, Benjamin Lefaudeaux, Min Xu, Pengchao Wang, Vivek Pai, Mannat Singh, Vitaliy Liptchinsky, Ishan Misra, Armand Joulin and Piotr Bojanowski. "Self-supervised Pretraining of Visual Features in the Wild." *ArXiv abs/2103.01988* (2021): n. pag.

[16] Chen, Ting, Simon Kornblith, Mohammad Norouzi and Geoffrey E. Hinton. "A Simple Framework for Contrastive Learning of Visual Representations." *ArXiv abs/2002.05709* (2020): n. pag.

[17] Caron, Mathilde, Ishan Misra, J. Mairal, Priya Goyal, Piotr Bojanowski and Armand Joulin. "Unsupervised Learning of Visual Features by Contrasting Cluster Assignments." *ArXiv abs/2006.09882* (2020): n. pag.



از توجه شما
سپاسگزارم